

A Course in Quantitative Ecology

L.L. Eberhardt

Rev. 19 Dec. 2007

The attached pdf files are chapters for a course in quantitative ecology. Drafts of all 14 chapters are now available. A date is included in the chapter designator to make it possible for readers to determine whether they have previously downloaded the chapter. The chapters are simply text files so that students can print them out as needed. Modifications to take advantage of various web features will be added later. The first order of business has been to make the textual material readily available. Inasmuch as there are over 400 pages, some users may prefer to look up a topic in the Table of Contents and scan it in the pdf files without printing them all out. For classroom use, students will no doubt prefer to print out assigned and relevant chapters.

The text is intended to be essentially self-contained, but quite a few references have been included. These serve several purposes. One is to supply more information on a topic, and further sources. A second is to provide support for the text. Textbooks produced by the major publishing houses are usually reviewed by independent authorities in the subject matter field before being published (and are often reviewed in journals after publication). Also, editors and proofreaders normally go over the text quite carefully before publication. This text has not benefited from such services, so readers may want to check back to sources on occasion (I would appreciate being informed of any apparent errors, etc., at the email address linked to my name above).

The text stems originally from notes used in teaching a course in 1983 at the Center for Graduate Studies in Richland, Washington (now the Tri-Cities Branch of Washington State University). Before the most recent (1996) use, I reviewed current issues of the journal, *Ecology*, in an effort to determine just what statistical procedures might be most prominent in ecological papers. This turned out to be the analysis of variance by a very wide margin, so we spent a great deal of time on that subject in the course that year. While a chapter is devoted to ANOVA in the present version, I do not now believe that it should be a major component of the course. I think frequency of analyses of variance in ecological publications has more to do with editorial insistence on evidence of “statistical significance” than it does with utility as a research tool.

In my opinion, a major weakness in ecological work today stems from the use of “canned” programs without adequate understanding of the underlying techniques. Certainly the computer saves a huge amount of time and effort, and makes it possible to do computations that were not possible in the past. But it is essential to understand the basis for an analysis before applying it. Consequently the present course depends on EXCEL spreadsheets for exercises, and students are required to work through problems by direct computations so that they see how the equations “work”. There is, of course, no objection on my part if you use software to check your answers! The most tedious aspect of this course is doing bootstrapping in EXCEL. Anyone who wants to use bootstrapping extensively will quickly realize the advantages of learning a programming language. Because many ecology students do not know how to program, it has seemed to be necessary to use spreadsheets to explore bootstrapping. After students learn how to bootstrap, they may want to use the Visual Basic for Applications (VBA) programs attached as EXCEL files, and described in the Appendix, for further applications of bootstrapping.

Two final comments: I suspect that most graduate programs in Ecology and Wildlife Management will not accommodate more than 4 semester hours of a “quantitative ecology” course. I have assumed that such a course will have as prerequisite one course in elementary statistics (but experience shows that most students require a refresher, hence Chapter 1). I don’t believe that 4 semester hours is nearly enough background if students are to be grounded in the quantitative techniques that they will need. I suspect that this is why there are so many rather dubious analyses in the literature that appear to be based on blind use of sophisticated commercial software. I am thus uncertain how the course should be taught. In earlier efforts I tried to cover a very large amount of material. Subsequently, I felt that bootstrapping is such an important and useful development for ecologists that it “had” to be included. Doing so is likely to make it impossible to cover all the material included here in one semester.

The second comment is that the course has much about large mammals in consequence of my own recent experience. I hope it will be possible to use more examples dealing with other components of ecosystems in future revisions, and would welcome suggestions and general comments.

Dr. D. P. DeMaster kindly arranged for inclusion of the text in the National Marine Mammal Web Pages (National Marine Fisheries Service), and Chris Boucher accomplished the actual installation. Keith Brenden has been handling revisions. Anyone is welcome to download and use the pdf version, but the usual copyright restrictions apply insofar as further duplication is concerned.

L. L. Eberhardt
2528 W. Klamath Avenue
Kennewick Washington 99336
November 7, 2002

CONTENTS

1.0 STATISTICAL BACKGROUND

- 1.1 Introduction
- 1.2 Some basic statistical concepts
 - Example 1.1 Coin-tossing
- 1.3 The Central Limit Theorem
 - Example 1.2 Frequency distributions
- 1.4 Simple linear regression
- 1.5 Multiple regression
- 1.6 Test for significant deviations from regression
- 1.7 Testing for curvilinearity without replications
- 1.8 Basic models for population growth
- 1.9 Testing for differences between regression lines
- 1.10 The analysis of covariance
- 1.11 ANOVA as a regression model
- 1.12 Stepwise regression
- 1.13 Logistic regression
 - Example 1.3 An Example of Logistic regression
- 1.14 Locally weighted regression
- 1.15 Non-linear least-squares
- 1.16 Exercises (17)

2.0 INTRODUCTION TO BOOTSTRAPPING 17pp

- 2.1 Introduction
- 2.2 The mechanics of bootstrapping
- 2.3 Empirical probability distributions
 - Example 2.1 Calculations for survival time data set.
- 2.4 Sample sizes for bootstrapping
- 2.5 Percentile confidence limits
- 2.6 Regression models and parametric bootstrapping
 - Example 2.2 Parametric regression bootstrapping
 - Example 2.3 A regression estimate of survival rate
 - Example 2.4 The correlation coefficient
- Exercises (4)

3.0 FURTHER TOPICS IN BOOTSTRAPPING 18pp

- 3.2 Estimates of bias
 - Example 3.1 Bias of variance for normal distribution
- 3.3 An improved bias estimate
- 3.4 Cross-validation
 - Example 3.2 Cross-validation error.
- 3.5 Bootstrapping for predictions
- 3.6 Improved confidence intervals
- 3.7 The Jackknife

3.8 The Monte Carlo method

3.9 The delta method

Example 3.3 Application of bootstrapping to a complex function.

Exercises (7)

4.0 SAMPLING METHODS 27pp

4.1 Introduction

4.2 Simple random sampling

4.3 The finite population correction

4.4 Confidence intervals

Example 4.1 Calculating confidence limits

4.5 Determining sample size

Example 4.2 Determining sample size

4.6 Stratified random sampling

4.7 Mechanics of stratification

4.8 Estimates from a stratified sample

4.9 Confidence limits

Example 4.3 A caribou census

Example 4.4 A mortality survey

4.10 Allocating the sample to strata

Example 4.5 A deer population estimate

4.11 Further remarks on stratified sampling

4.12 Ratio Estimation

4.13 Variance of ratio estimates

Example 4.6 ratio corrections for variable plot sizes

4.14 Double sampling

4.15 Cluster sampling and subsampling

Example 4.7 A cluster sample example

Example 4.8 Cluster sampling involving proportions

4.16 Exercises (10)

5.0 TRANSECT METHODS 38pp

5.1 Introduction

5.2 A classification of transect methods

Example 5.1 Censusing prairie-dog dens

Example 5.2 "Needle" sampling

5.3 The line-intercept method

5.4 Length-biased sampling

Example 5.3 Auxiliary measurements

5.5 Flushing-distance line transects

Example 5.4 Errors of measurement in line transects

5.6 Right-angle distance line transects

Example 5.5 Testing flushing-angles

5.7 Density Estimation

Example 5.6 Density estimation for line transects

5.8 A "distribution-free" method

- Example 5.7 The "Cox" method
 - Example 5.8 Emlen's method
- 5.9 Assumptions underlying line transect methods
- 5.10 Strip transects
- 5.11 Modified strip transects
- 5.12 Survey design
- 5.13 Exercises (3)

- 6.0 THE ANALYSIS OF VARIANCE 22pp
- 6.1 One-way analysis of variance
- 6.2 Two-way analysis of variance
- 6.3 Randomized blocks designs
- 6.4 Two-way analysis of variance with replication
- 6.5 Assumptions for the analysis of variance
- 6.6 Comparisons in ANOVA
- 6.7 Type I and II errors and "power"
- 6.8 Other aspects of ANOVA
- 6.9 Exercises (15 Exercises)

- 7.0 INTRODUCTION TO POPULATION ASSESSMENT 21 pp
- 7.1 Viable populations
 - Check program for Fig. 6.2
- 7.2 Methods for population assessment
- 7.3 Population estimation
- 7.4 Survival estimation
- 7.5 Reproductive rate estimation
- 7.6 Projection models
- 7.7 Some historical features of population analysis
- 7.8 A classification of methods for estimating abundance

- 8.0 CAPTURE-RECAPTURE METHODS 21pp
- 8.1 Introduction
- 8.2 Petersen's method (Lincoln Index)
 - Example 8.1 Petersen method
- 8.3 The Schnabel method
 - Example 8.2 The Schnabel method
 - Example 8.3 Estimation from frequency of capture
 - Example 8.4 Mean Petersen method
- 8.4 Methods for "open" populations
 - Example 8.5 The Jolly-Seber method
- 8.5 Tests of assumptions
- 8.6 Exercises (5)

- 9.0 INDICES OF ABUNDANCE 56pp
- 9.1 Introduction

- Example 9.1 The pellet-group count
- Example 9.2 Conversion factors for pellet-counts
- Example 9.3 Roadside counts
- Example 9.4 A pocket-gopher index
- 9.2 Trends in abundance
- 9.3 Testing significance of trend lines
- 9.4 Assessment of trend data
- 9.5 A test for significant deviations from regression using replicate points.
- 9.6 A test for linearity based on the lowess method
- 9.7 Confidence limits from bootstrapping
- 9.8 Alternative estimates of rate of population change
 - Example 9.5 A census using ratio estimation
- 9.9 Criteria for regression fits
- 9.10 Using auxiliary variables with trend data
 - Example 9.6 Trend indices with auxiliary variables
 - Example 9.7 An alternative approach to index models
- 9.11 Catch-effort methods
- 9.12 Models for catch-effort methods
 - Example 9.8 Calculations for catch-effort methods
 - Example 9.9 Variable effort models
 - Example 9.10 Catch-effort data on an "open" populations
- 9.13 Catch-effort models with non-harvest losses
- 9.14 Sampling for indices
- 9.15 Transformations
- 9.16 Combining indices
- 9.17 Converting indices
- 9.18 Comments on the use of index data
- 9.19 Exercises(16)

- 10.0 ESTIMATING SURVIVAL 24pp
- 10.1 Introduction
- 10.2 A single-recovery model
 - Example 10.1 Chapman-Robson survival estimate
- 10.3 Further estimators
- 10.4 Analysis of incomplete data
 - Example 10.2 Analysis of incomplete data
- 10.5 Testing assumptions
 - Example 10.3 Testing compatability of first recoveries
 - Example 10.4 Combining years
- 10.6 Life tables
- 10.7 Catch-curves and kill-curves
- 10.8 Combining estimates from several years
 - Example 10.5 An example of Haldane's method
- 10.9 Regression methods
- 10.10 Effect of changes in survival
- 10.11 Use of telemetry to estimate survival

10.12 Kaplan-Meier estimates of survival

10. Exercises (0)

11.0 THE LOTKA-LESLIE MODEL 33pp

11.1 Introduction

11.2 Discrete and continuous population growth

11.3 Lotka's equation and the Leslie matrix model

11.4 The stable age structure

11.5 Augmented life tables

11.6 An example of an augmented life table

11.7 Reproductive and survival curves

11.8 Some useful approximations

11.9 Approximate variance of lambda

11.10 Examples of bootstrapping and the delta method

11.11 Delta method examples

11.12 Comparisons with direct estimates of lambda

11.13 Development of stable age-structure

11.14 Lotka's equations

11.15 Net rate of increase

11.16 Practical applications

11.17 Exercises (7)

12.0 POPULATION MODELS 15 pp

12.1 Introduction

12.2 Curve-fitting models

12.3 Some growth curves

12.4 Projection models

12.5 Predator-prey models

12.6 A wolf-ungulate model

12.7 Exercises (7)

13.0 Dealing with density-dependence

13.1 Introduction

13.2 Some historical aspects

13.3 Testing for density-dependence

- Models for density dependence

- Testing for density dependence

- Check simulations of autocorrelation models

- Further work on density-dependence; look over Bulmer's papers on d-d

- check test now in program

- Four stage hypothesis

- role of dispersal

- Projection models and density dependence

- Reality of asymptotes -- fluctuations around asymptote

14.0 SOME CASE HISTORIES

14.1 Introduction

14.2 The Florida manatee

14.2 The Pacific walrus

14.3 Bowhead whales

14.4 Grizzly bears

14.5 Hawaiian monk seals

14.6 Fur seals

14.7 Sea otters

14.8 Feral horses

14.9 The Madison-Firehole elk herd

14.10 Isle Royale moose and wolves

15.0 APPENDICES

Some basic math

Maximum likelihood

Matrices

1.0 STATISTICAL BACKGROUND

1.1 Introduction

In most of the following chapters, it will be assumed that readers have at least taken an introductory course in statistical methods. Some basic concepts will nonetheless be reviewed in this chapter to provide background for the following material. Some essential definitions are listed here. Students should look these up in any introductory statistics textbook, but preferably in a text that they have used in the past. An effort has been made to keep the introductory terminology to a minimum, and it will be supplemented as we go along, and by auxiliary reading.

1.2 Some basic statistical concepts

Random variables

In even quite simple situations, we need to be able to distinguish between an abstract label for an observation, and the observations that we actually make in some real-world situation. Statisticians do this by using capital letters ($X_1, X_2, X_3, \dots, X_N$) for the abstract label and lower case letters (x_1, x_2, \dots, x_n) for the observations we make in practice. Note that the ellipsis (...) means that some letters are left out -- from the first three given, we can infer that these are X_4, X_5 , etc., thru X_{N-1}). More importantly, note that this is a series of finite length -- N random variables in all. In some cases, we need to consider an indefinitely long series of numbers, and write X_1, X_2, X_3, \dots to indicate that fact. Also, note that the random variables run from X_1 to X_N , but that the observations end in x_n . This is because we often want to sample a large population and thus only record n of the N possible observations.

Example 1.1 Coin-tossing Consider a simple coin-tossing example. Put 10 coins in a jar, shake well, spill them out and count the number of heads. You will get observations like the following table (note that the individual observation, x_i , is the total number of heads out of 10 coins and that the table is based on 100 tosses of 10 coins):

```

5,5,2,4,3,4,5,6,5,6
4,6,3,7,4,6,5,3,5,4
6,5,5,2,5,5,3,3,6,7
5,8,4,3,4,5,6,5,5,3
5,6,7,5,8,8,7,3,7,7
5,4,6,5,3,6,4,6,5,4
3,3,6,4,7,5,6,6,3,4
6,4,5,6,6,4,3,4,8,3
6,2,8,5,7,4,6,4,5,6
1,6,6,7,5,3,5,6,7,3

```

One can continue this process indefinitely, so we may have to consider an infinite sample space. In many cases, we will be considering finite sample spaces, although we often will not know N . In this case, we do know that $N = 100$, but if we are considering some natural population over a large area, we likely will not know N , and we may in fact have estimating N as our objective. There is some ambiguity in notation here in that N can be considered to be a fixed population of the outcomes of 100 tosses, or a sample (n) of the infinite number of possible tosses.

Much of statistical methodology consists of describing the outcomes of "experiments" like coin-tossing, and making inferences about the process that led to the set of observations. Most of the theory underlying statistical methods depends on having a model for the underlying process. Such models are described as probability density functions (abbreviated as pdf). Such a model for the coin-tossing example is the binomial distribution, often written as $Bi(n,p)$ which says that the probability that a randomly obtained observation denoted as x_j takes the value k is:

$$\text{Prob}\{x_i = k\} = f_k = \binom{n}{k} p^k (1-p)^{n-k} \quad (1.1)$$

where $\binom{n}{k}$ is evaluated as $\frac{n!}{(n-k)!k!}$, in which, for example, 5! (read as "five factorial") is calculated as $5 \times 4 \times 3 \times 2 \times 1 = 120$.

This equation gives the pdf for a binomial having n trials (10 in the coin-tossing example). In the example, the random variable can take 11 possible values 0,1,2,3, ..., 10, but in the 100 trials listed above, we observed no zeros and no 9's or 10's. In many practical examples, we won't know the value of p , and want to estimate it from the observed data. If we can somehow establish that it is appropriate to assume the model of eq.(1.1), then we can calculate its expected value, defined as:

since we are here considering a discrete random variable that is only defined on the sample space 0,1,2,...,10, the integral can be replaced with a summation, and this can be evaluated with some algebra to find that $E(x) = np$. We can then turn this around to estimate p from the mean value of our sample, which is

$$E(x) = \int_{x=0}^{x=\infty} x f_x dx = \sum_{x=0}^{x=10} x \binom{n}{x} p^x (1-p)^{n-x} \quad (1.2)$$

calculated as the sum of the observations (496) divided by the number of observations (100) or $E(x) = np = 4.96$. Since $n=10$, we estimate p as:

$$\hat{p} = \frac{4.96}{10} = 0.496.$$

The "hat" over p denotes that it is an estimate of the parameter, p , of the binomial pdf. From the structure of the experiment we can infer that the value of p should be about 0.5, that is, if the coin is "unbiased", the probability that it turns up heads should be 1/2.

The sample mean, $\bar{x} = \sum x_i / n$ is often described as a "statistic" derived from a set of observations. Other commonly used statistics are the sample variance:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

and the standard error of the mean, $s.e. = \left[\frac{s^2}{n} \right]^{1/2}$. Note that statistics are functions of the data. The mean

can be written as $\bar{x} = \frac{1}{n} (x_1 + x_2 + x_3 + \dots + x_n)$, which is a linear function of the random variables x_1, x_2, \dots, x_n . There are some simple rules from probability theory about linear functions of random variables that make it easy to derive useful results about means.

No doubt the most important probability density function (pdf) in statistics is the normal distribution, which is written as:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1.3)$$

The parameter μ is the mean of the distribution and σ the standard deviation. Tables of the frequency distribution ($f(x)$) of this distribution are available in almost any statistics text, but with parameters $\mu = 0$

and $\sigma = 1$, which is described as the unit normal distribution or standard normal distribution, often represented by the notation $N(0,1)$, while observations drawn from eq.(1.3) are described as $N(\mu, \sigma^2)$.

1.3 The Central Limit Theorem

A very useful result from mathematical statistics is the Central limit theorem:

"Let X be a random variable with mean μ and variance σ^2 , then the random variable Z :

$$Z = \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma}$$

has a distribution that approaches the standard normal distribution as n approaches infinity.

This says that, if n is large, then we are virtually guaranteed that the sample mean will have nearly a normal distribution. Inasmuch as the great bulk of modern statistical methods depend on the normal distribution, this result is very reassuring. The important question then is "how large must n be for approximate normality?", and the answer depends very much on the frequency distribution underlying the observed x_i .

Example 1.2 Frequency distributions Consider the data from the coin-tossing experiment (Example 1.1). The random variable tabulated is the number of heads in 10 tosses. We can tabulate the frequency of each outcome (0,1,2,3,...,10 heads) and compare it with the expected frequency calculated from eq.(1.1), giving the following result:

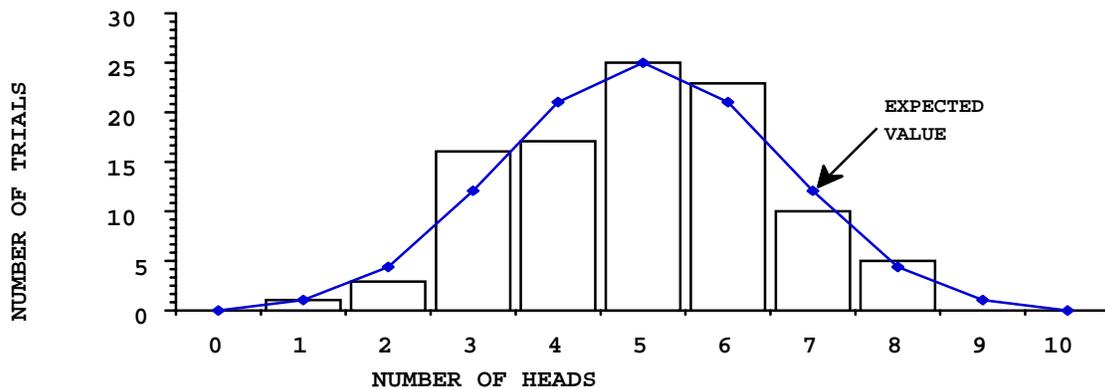


Fig. 1.1 Frequency distribution of number of heads observed in 100 tosses of a coin compared to number expected from eq.(1.1), $Bi(10,0.5)$.

The observed data are not as symmetrical as the expected binomial distribution, but the variance (2.34) is a reasonably good approximation to the variance from the theoretical binomial (2.5) and the mean (4.96) of 100 trials is very close to the theoretically expected value (5). The expected binomial variance of the random variable x , the number of heads in 10 tosses, is readily calculated as $np(1-p) = 5(.5)(.5) = 2.5$. It is worthwhile to compare (Fig. 1.2) the expected binomial distribution with a normal distribution with the theoretical mean and variance, as calculated from eq.(1.2).

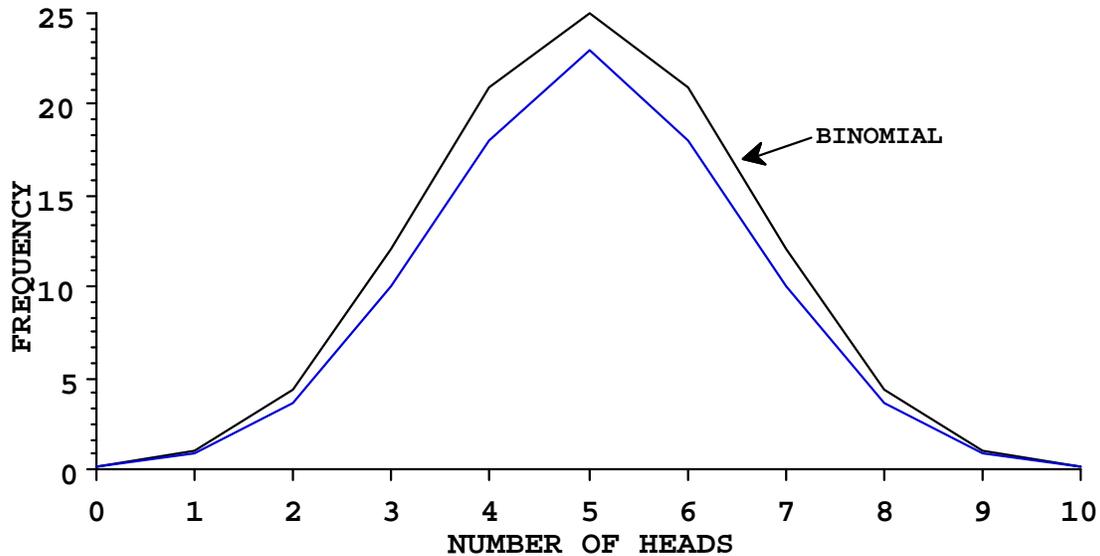


Fig. 1.2 Expected values from a binomial distribution of outcomes of 10 tosses of coins compared to frequencies calculated from a normal distribution (broken line) with the theoretical mean (5) and variance (2.5) for the binomial distribution.

Note that the normal distribution is continuous, i.e. that it takes on all values over the interval considered and is thus only an approximation to the discrete distribution of the results of coin-tossing, in which only integer values can be observed ($i = 1, 2, 3, \dots, n$ heads). Hence the points representing the binomial distribution in Fig. 1.1 properly should not be connected by lines. Because the normal distribution has an infinite range it isn't strictly proper to use it in Fig. 1.2 because there is only a finite possible range of outcomes (0 to 10). However, it is often used as an approximation. Note, too, that there is less area under the normal distribution in Fig. 1.2 because theoretically some observations will be greater and lesser than the range plotted.

1.4 Simple linear regression

Simple linear regression follows the model:

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (1.4)$$

where y_i is the dependent variable and x_i the independent variable and the error term (ε) is a deviation from the "true" relationship. Estimates of α and β are frequently written as a and b , giving the estimated or fitted relationship as:

$$y_i = a + bx_i \quad (1.5)$$

Estimates of regression parameters, α and β do not require any assumptions, and can be calculated from any set of x, y pairs. However, tests of significance and confidence limits require adding some assumptions, which center around the ε_i being normally distributed with mean zero and variance σ^2 . The assumptions will be discussed after we consider the "machinery" of regression analysis.

The estimates are obtained by the method of least-squares, an important and useful tool that traces back to Legendre and Gauss (known also for the normal distribution) in the early 1800's. Other ways of fitting a straight line to data are available, but seldom used. The approach is based on minimizing a sum of squared deviations, written as:

$$S = \sum [y_i - (\alpha + \beta x_i)]^2 \quad (1.6)$$

where the summation runs from 1 to n. This is accomplished by the methods of calculus, finding the partial derivatives:

$$\frac{\partial S}{\partial \alpha} = 2 \sum (y_i - \alpha - \beta x_i) = 0 \quad (1.7)$$

$$\frac{\partial S}{\partial \beta} = 2 \sum x_i (y_i - \alpha - \beta x_i) = 0$$

these give the normal equations (α and β are replaced by the symbols for estimates, a and b):

$$\sum y_i = na + b \sum x_i \quad (1.8)$$

$$\sum y_i x_i = a \sum x_i + b \sum x_i^2$$

and these can be solved jointly to give the estimates:

$$a = \bar{y} - b\bar{x} \quad (1.9)$$

$$b = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

Note that the deviations of eq.(1.4) are in the vertical plane, being deviations of y_i from the fitted line. Fig. 1.3 shows two of the deviations from a regression line fitted to some counts of deer. The fitted line appears on the graph along with a measure of the fit, R^2 , which will be defined below.

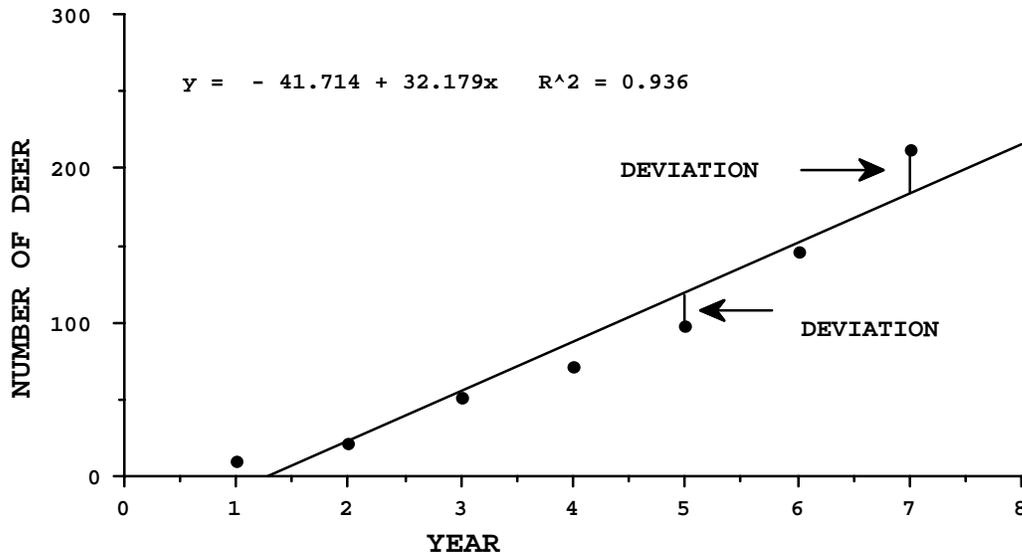


Fig. 1.3. Simple linear regression fitted to successive counts of the number of deer on a study area by the method of least squares.

Table 1.1 gives the analysis of variance results for the deer data from EXCEL, in ANOVA format (the analysis of variance is discussed in Chapter 6). Figure 1.4 shows the deviations from the mean of the y -values, and a comparison with Fig. 1.3 shows why the reduction in Sum of Squares from regression is so substantial (compare Total SS with Residual SS). The residual S.S. is computed from the residuals from the fitted regression line, i.e.:

$$\begin{aligned} \text{Residual S. S.} &= \sum_{i=1}^n [y_i - (a + bx_i)]^2 = \sum_{i=1}^n [y_i - ((\bar{y} - b\bar{x}) + bx_i)]^2 & (1.10) \\ &= \sum_{i=1}^n [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 \end{aligned}$$

$$= \sum_{i=1}^n [y_i - (a + bx_i)]^2 = \sum_{i=1}^n (y_i - \bar{y})^2 - b^2 \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1.11)$$

$$\text{Residual SS} = \text{Total SS} - \text{Regression SS}$$

Eq.(1.11) can be obtained by introducing the definition of b after squaring the intermediate step above.

Table 1.1 Analysis of variance in regression of deer data of Fig. 1.3 as obtained in EXCEL.

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P value</i>
Regression	1	28992.89	28992.89	73.25	0.0004
Residual	5	1979.11	395.82		
Total	6	30972.00			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Statistic</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	-41.71	16.81	-2.48	0.06	-84.94	1.51
Slope (b)	32.18	3.76	8.56	0.00	22.51	41.84

EXCEL gives the slope coefficient (b) as "X Variable 1" because the regression program is also set up to handle multiple regression, where there will be 2 or more independent variables. ANOVA is discussed in detail in Chapter 6.

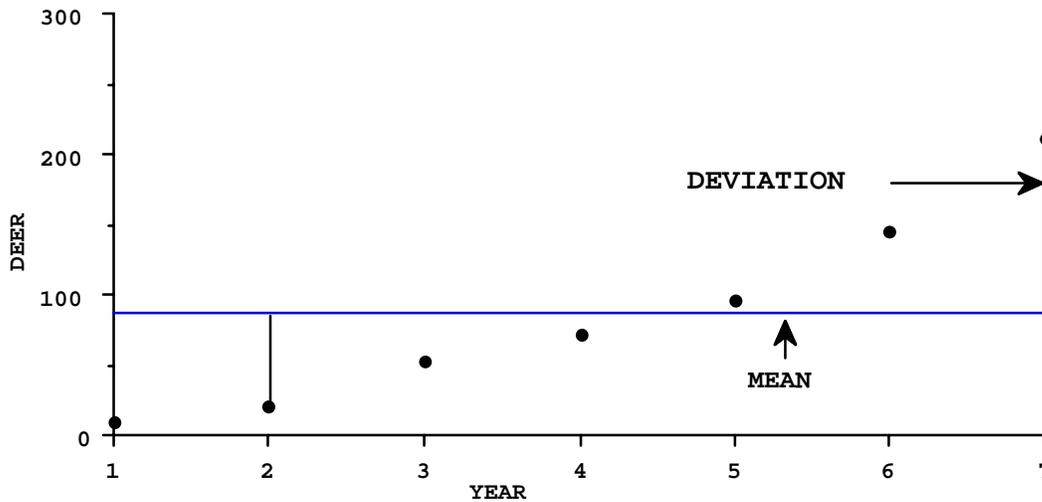


Fig. 1.4 Deer data as in Fig. 1.3 but showing deviations from the mean of the y-values, \bar{y} . This shows why the Residual S.S. is ordinarily much smaller than the Total Sum of Squares, which is calculated from the deviations illustrated here.

If the F-value is not significant, there clearly is not much to be gained from the regression line. For simple linear regression, the square root of R-squared (R) is Pearson's product-moment correlation, usually simply referred to as "the" correlation coefficient (but written as a lower-case r), and calculated as follows:

$$r = \frac{\sum(y_i - \bar{y})(x_i - \bar{x})}{[\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2]^{1/2}} \quad (1.12)$$

The correlation coefficient is related to the slope of the regression line (b) by the following expression:

$$b = \left[\frac{\sum(y_i - \bar{y})^2}{\sum(x_i - \bar{x})^2} \right]^{1/2} r \quad (1.13)$$

this is sometimes expressed by $\frac{s_y}{s_x}$ r, i.e., the ratio of the sample standard deviation of y to that of x times r.

R^2 is also used for multiple regression (described below), where the square root is not the ordinary correlation coefficient, so it is useful to have another expression for R^2 . This is:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1.14)$$

The quantity R^2 is often described as measuring the "percent of variance accounted for by regression", in consequence of the fact that it is the ratio of the Regression SS to the Total SS.

Another valuable expression is that of the estimated variance of the slope:

$$s_b^2 = \frac{s^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1.15)$$

This expression is particularly useful because it makes it possible to suggest how the estimate of b with smallest variance might be obtained. Concentrating the selection of values of x_i at which to observe y_i at the ends of the possible range of x will evidently give the smallest obtainable variance on b (by giving the largest possible value of the denominator in eq.(1.15)). However, such a course is recommended only when one can be virtually certain that the underlying relationship is linear. We will consider ways to test for nonlinearity in the regression line in a section below. Note, for example, that the data of Fig. 1.3 seem clearly to follow a curved relationship. Concentrating the observations at x -values at the ends of the range of observable y would make it impossible to detect such curvature. Whether we can concentrate observations depends, of course, on the nature of the data. In the case of the counts of deer, we normally make only one observation per year, if the data are an actual census (i.e., a complete count of the deer on an area). In the case of a sample estimate of the number present, it may be possible to take repeated, independent samples and thus get several estimates per year (replicates).

A confidence interval for the slope, b , uses the t -distribution:

$$b \pm t_{\alpha, d.f.} S_b \quad (1.16)$$

Note that α now represents the significance level for the t -distribution, and not the parameter of a regression line. Additional confidence intervals for values predicted from the regression line of y or \bar{y} for a given x are given in standard references (e.g., Snedecor and Cochran). Much more detail on regression analysis is given in texts on the subject. An extensive treatment is given by Draper and Smith (Applied Regression Analysis, J. Wiley and Sons Third Edition, 1998). The main parts of the book are presented in matrix algebra notation, but the authors do give a short introduction to the matrix algebra that is adequate to let one follow their presentation of regression topics, and not difficult to understand.

In order to justify any significance tests in regression analysis we must consider the assumptions. The model now becomes:

$$y_i = \alpha + \beta x_i + \epsilon_i \quad (1.17)$$

where, as with the ANOVA model, we now assume that the ϵ_i are normally distributed with mean zero, variance σ^2 , and are uncorrelated (independent). An important additional assumption is that the x_i values are all measured without error. If the x_i are subject to measurement ("sampling") variation, then the regression line can still be calculated as given above, but its interpretation changes, as do the tests of significance. For the most part, the assumptions for linear regression are somewhat less troublesome than for ANOVA in general. However, we usually need large numbers of replicates to do any testing of the assumptions. Possibly the most important precaution is to be sure that any replicate values of y are indeed obtained independently. In much ecological data it appears likely that the variances of sets of y -values may be proportional to the x_i at which they are taken, or that the coefficients of variation of the replicate y -values may be approximately constant. The F -tests will then be less-reliable. However, simple linear regression is quite "robust" to uncertainties about the assumption of normal errors, so long as the x -values are not subject to error.

A simulation is useful in appraising the assumptions for simple linear regression. Using eq. (1.17) as

$$y_i = 2 + 0.30x_i + \epsilon_i$$

with the x_i as 1,2,3, ...,10 and the ϵ_i generated as observations from a normal distribution with mean 0 and variance = $\sigma^2 = 1$, one can generate a table of "data" as before. This was done to produce a set of data for 20 regression lines.

The first 5 data sets are as follows:

x	True y	Simulated y_i				
		1	2	3	4	5
1	2.30	2.86	1.28	2.61	3.26	1.88
2	2.60	0.90	2.08	2.58	1.68	1.11
3	2.90	1.56	3.35	1.35	4.28	3.50
4	3.20	3.85	2.84	2.02	3.67	2.46
5	3.50	1.62	4.20	4.87	1.49	4.68
6	3.80	4.39	5.78	5.22	2.77	3.62
7	4.10	3.66	2.61	4.70	4.24	4.99
8	4.40	3.95	3.90	5.98	2.59	3.81
9	4.70	4.45	6.15	6.41	5.53	3.72
10	5.00	4.50	5.53	6.09	3.54	4.32

Note that the simulated data vary appreciably from the "true values" computed from $y_i = 2 + 0.30x_i$, which appear in the second column above. The simulated data points should follow a normal distribution around the true regression line. Plotting the data (Fig1.5) suggests a certain amount of clumping near the center in some cases, but also shows considerable variability around the true line. If we plot all 200 deviations used to construct the simulations (20 simulations for each of 10 x-values (Fig. 1.6) then it does appear that the underlying distribution is roughly symmetrical, but it should be apparent that one cannot do much testing for normality with smaller samples (say 10 or 20) of deviations from a regression line.

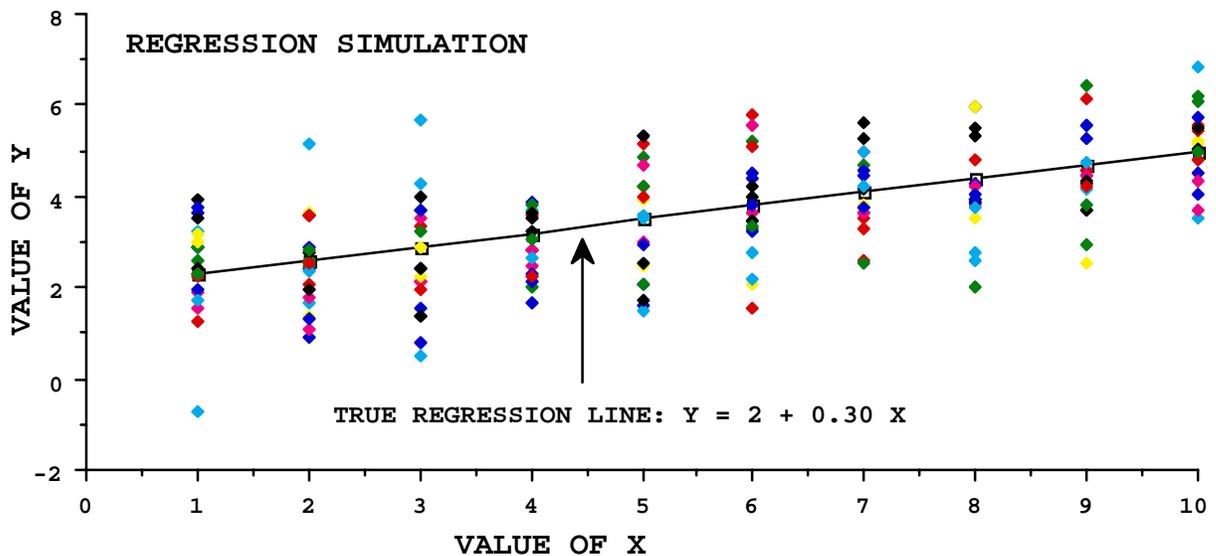


Fig. 1.5 Simulated regression data plotted with the true regression line from which the data were simulated by adding normal deviates with mean zero and unit variance.

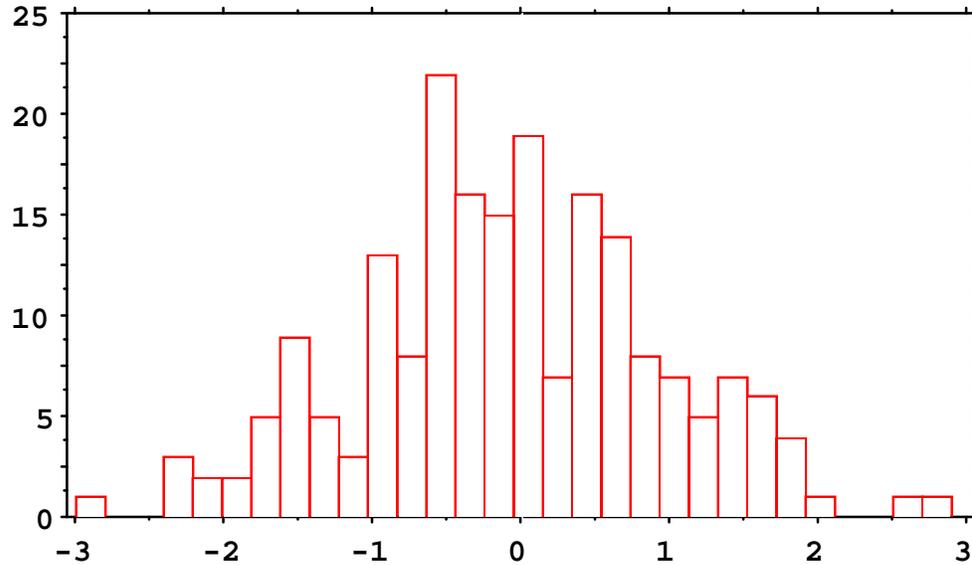


Fig. 1.6 Plot of 200 normal deviations with mean zero and unit variance used to obtain 20 regression simulations.

The regression program in EXCEL used to produce Table 1.1 was run on all 20 sets of generated data and the estimates of intercept (a) and slope (b) were tabulated along with the residual M.S. (s^2) and the confidence limits for b. The error M.S. estimates ranged from 0.5 to 2.41, but averaged 1.05, very close to the expected 1.0. Estimates of the intercept (a; true value 2.0) ranged from 0.2 to 2.83, averaging 2.02, while slope estimates (b; true value 0.30) ranged from 0.11 to 0.54, averaging 0.31. The 95% confidence limits (Fig. 1.7) for the 20 regression estimates of the slope (b) vary considerably, but include the true value in 19 of 20 cases, as expected ($0.95(20)=19$). It should be noted that this was a fortuitous outcome -- much larger simulations would be needed to be sure that the confidence limits actually include the true β in 95% of cases.

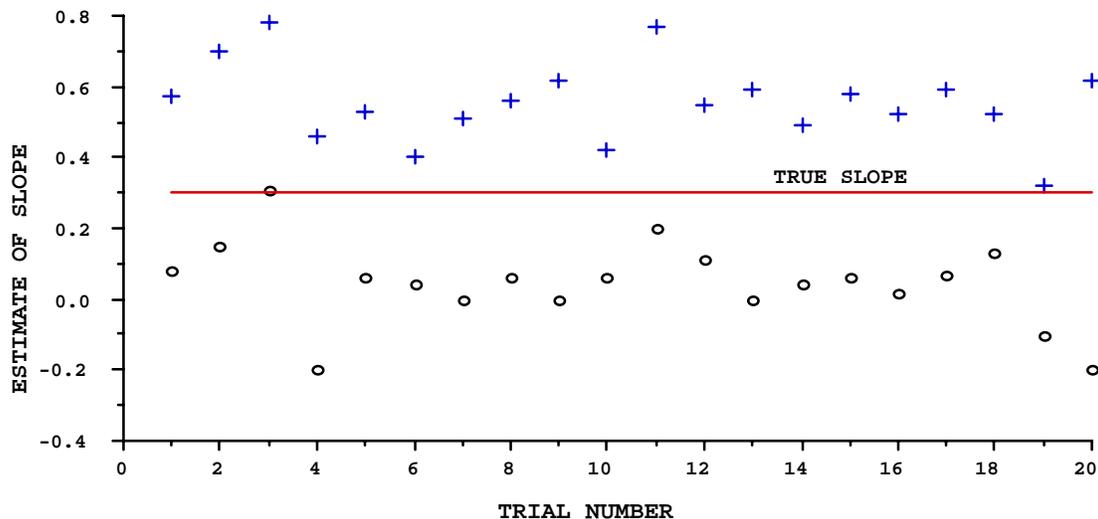


Fig. 1.7 Confidence limits (95%) for slope of 20 simulated regression lines, shown with the true value (0.30). Note that confidence limits for the 3rd data set do not include (are above) the true value.

1.5 Multiple regression

Multiple regression is somewhat of a risky proposition for ecologists, inasmuch as relationships between several ecological variables tend not to be linear. However, it can be used to explore curvilinear relationships (which we will do below and in the Exercises) and there are various circumstances where a linear model may be useful. It is also true that a multiple regression model is behind many other kinds of analyses. The analysis of variance can be obtained through a multiple regression model, but with a different structure than that used here.

The general model is like that for simple linear regression, but adds more independent variables. We will use 2 here, but EXCEL will compute models with many x-variables. The basic model is:

and the same assumptions are made. We again minimize the sum of squares leading to normal equations in

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \quad (1.18)$$

three variables and the following solutions for the parameters (a, b₁, b₂):

$$a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 \quad (1.19)$$

$$b_1 = [\Sigma(x_{2i} - \bar{x}_2)^2 \Sigma(y_i - \bar{y})(x_{1i} - \bar{x}_1) - \Sigma(x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2) \Sigma(y_i - \bar{y})(x_{2i} - \bar{x}_2)] / D$$

$$b_2 = [(\Sigma(x_{1i} - \bar{x}_1)^2 \Sigma(y_i - \bar{y})(x_{2i} - \bar{x}_2) - \Sigma(x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2) \Sigma(y_i - \bar{y})(x_{1i} - \bar{x}_1))] / D$$

where:

$$D = \Sigma(x_{1i} - \bar{x}_1)^2 \Sigma(x_{2i} - \bar{x}_2)^2 - [\Sigma(x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)]^2$$

Our first use of the above equations will be with $x = x_1$ and $x_2 = x^2$, which may look suspicious, but the purpose is legitimate inasmuch as we can now fit a second-degree polynomial (a "quadratic" to many statisticians) as an aid in studying curvature in regression data. To illustrate, we use the deer data of Fig. 1.3 getting the curve of Fig. 1.8. Snedecor and Cochran (1967) show how to do the Analysis of variance in regression in stages, fitting first x_1 and then x_2 to see whether there is any gain in adding a second variable. In the present case, we know that the second variable is necessary to yield a curve.

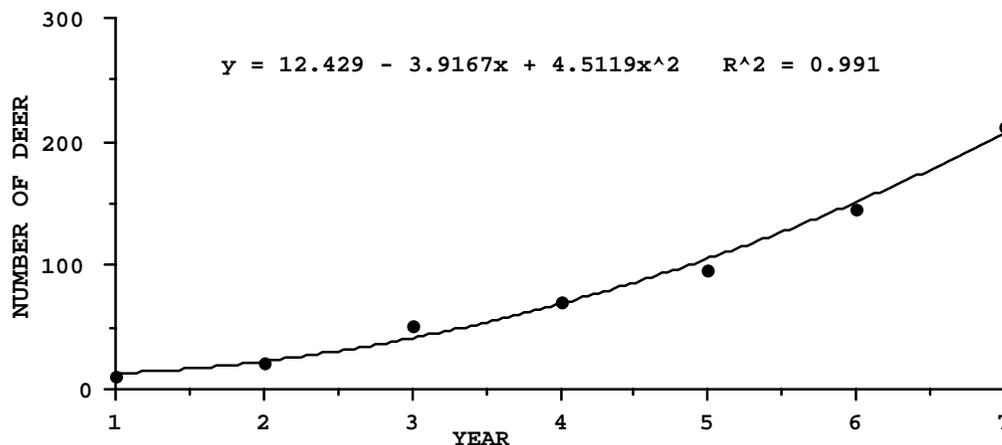


Fig. 1.8. Second degree polynomial fitted to deer data of Fig. 1.3, using multiple regression with $x_1 = x$, and $x_2 = x^2$.

Multiple regression can be used for a wide variety of analyses. For example, the analysis of variance can be represented and computed in a multiple regression format. A wide range of analyses based on multiple regression equations are described in some statistic texts under the heading of "General Linear Hypotheses".

1.6 A test for significant deviations from regression using replicate points.

A test for significant deviations from linearity depending on fitting a curve and testing to see whether the improvement in fit might simply be due to chance will be discussed in the next section. In some cases, however, replicate counts may be available, so that one can use the variability within years to test significance of deviations from linearity. This is the preferred approach, when available. The advantage is that we do not need to specify an alternative model like the quadratic or cubic, which may very well be the wrong model. Note, for example, that population growth data such as that of Fig. 1.8 are known to follow an exponential or geometric curve rather than the second degree polynomial used in Fig. 1.8. Some counts of brown bears at spawning streams provide an example for the test (Fig. 1.9). In this case, the test consists of making the usual analysis of variance to test for significance of the linear regression (Table 1.2), and then using the pooled variance of individual observations within years to estimate "pure error" (Draper and Smith 1998:49). The data for calculation of pooled error appear in Table 1.3. A sum of squares of deviations from the mean is calculated for the data in each year where there are two or more observations and these values are summed to give an overall sum of squares, which is subtracted from the "residual" sum of squares in Table 1.2 to yield the "lack of fit" sum of squares (i.e., the variability not accounted for by "pure error"). The number of counts used to calculate pure error (32) is similarly subtracted from the degrees of freedom for residual error to get the degrees of freedom used to calculate a mean square for "lack of fit". The resulting F-test indicates significance at the 0.05 level, but there does not seem to be much evidence of a consistent pattern of change in Fig. 1.9.

Table 1.2 Test of significance for deviations from regression

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Prob.</i>
Regression	1	0.127	0.127	6.258	0.016
Residual	47	0.954	0.020		
Total	48	1.081			
Lack of fit	15	0.479	0.032	2.150	0.034
Pure error	32	0.475	0.015		

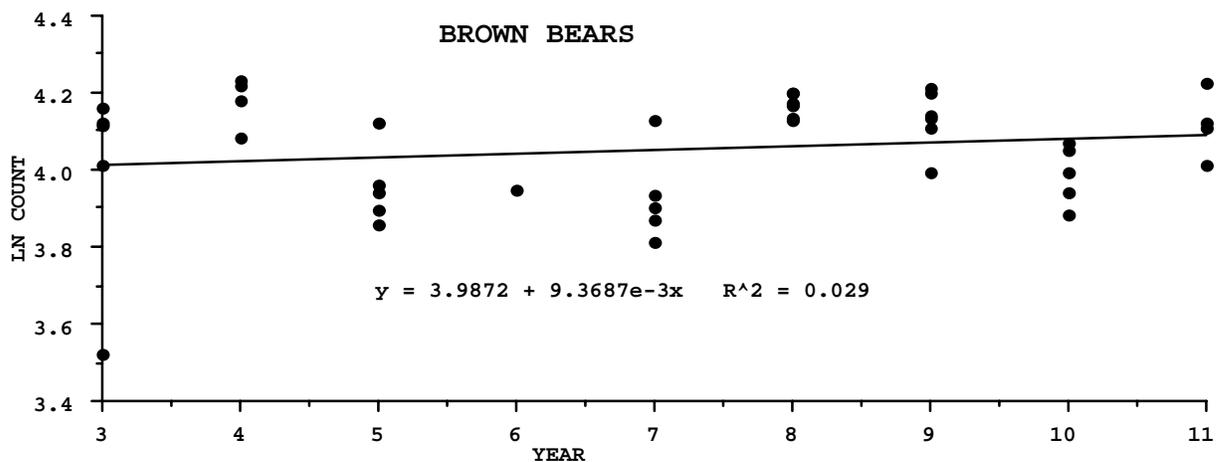


Fig. 1.9 Logarithms of counts of brown bears on salmon spawning streams.

Table 1.3 Data for computation of "pure error" for brown bear counts.

Year	Bears/hour	ln (bears/hr)	Sum of squares	d.f.
3	39.85	9.5219		
3	64.04	4.1595		
3	61.88	4.1252		
3	61.2	4.1141		
3	55.24	4.0117	0.2819	4
4	68.7	4.2297		
4	59.3	4.0826		
4	67.9	4.2180		
4	65.3	4.1790	0.0134	3
5	49.4	9.9000		
5	51.4	9.9396		
5	61.6	4.1207		
5	47.4	9.8586		
5	52.45	9.9599	0.0400	4
6	51.88	9.9489		
7	45.14	9.8098		
7	62	4.1271		
7	48.13	9.8739		
7	49.58	9.9036		
7	51.21	9.9359	0.0572	4
8	62.06	4.1281		
8	66.59	4.1986		
8	62.32	4.1323		
8	66.88	4.2029		
8	65.03	4.1748		
8	64.58	4.1679	0.0051	5
9	54.17	9.9921		
9	67.49	4.2120		
9	66.67	4.1998		
9	62.8	4.1400		
9	61	4.1109		
9	62.42	4.1339	0.0311	5
10	48.68	9.8853		
10	51.47	9.9410		
10	58.51	4.0692		
10	57.65	4.0544		
10	54.08	9.9905	0.0238	4
11	61.12	4.1128		
11	55.15	4.0101		
11	68.29	4.2238		
11	61.52	4.1194	0.0229	3
Sums	2386.08	166.2194	0.4753	32

1.7 Testing for curvilinearity without replications

Trend data are often collected without replications. Occasionally this is because an absolute count is made annually of individuals on an area; more often it is because the investigators cannot afford to make replicate sample counts (seasonal changes limit the time that such "replicates" are likely to be valid, too). In such circumstances, checking for nonlinearity of regression depends on fitting a straight line and a curved line, and appraising the improvement, if any, provided by the curve. The simplest curve available is the second degree polynomial ("quadratic") considered in the section (1.5) on multiple regression above. Sometimes it may be worth trying a third-degree polynomial ("cubic"), which is readily computed by multiple regression in EXCEL. The model is:

where $x_1 = x$, $x_2 = x^2$, $x_3 = x^3$. If a graphics program that fits polynomials is available, it is worthwhile to use it for a quick preliminary check. Often the 3rd degree polynomial has too much curvature, and the

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i \quad (1.20)$$

practical approach is to stick with the quadratic.

The procedure is straightforward. One first fits the simple linear regression model, obtaining the ANOVA of Table 1.1. Then fit a quadratic, and obtain the fit illustrated in Fig. 1.8, along with the corresponding regression ANOVA (Table 1.4).

Table 1.4 Analysis of variance in regression based on a multiple regression fit of a second degree polynomial (Fig. 1.8) to the deer data.

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	2	30702.90	15351.45	228.19	0.0001
Residual	4	269.10	67.27		
Total	6	30972.00			

From the linear regression table (Table 1.1), extract the residual sum of squares and use it as the first entry in a new table (Table 1.5). From the ANOVA table giving the multiple regression fit (Table 1.4) also extract the residual sum of squares and make it the second entry in the new table. Use the corresponding degrees of freedom in both cases. Subtract the S. S. for curvilinear regression from the S.S. for linear regression. This quantity, with 1 degree of freedom, represents the improvement in fit provided by curvilinear regression and is tested against the M.S. for curvilinear regression by an F-test. Table 1.5 gives the new arrangement for the deer data.

Table 1.5 Test for curvilinearity of regression using the difference between Residual Sum of Squares in linear regression and multiple regression.

TEST FOR CURVILINEARITY-ORIGINAL SCALE				
SOURCE		d.f.	S. S.	M. S.
Dev. from linear regr.	5	1979.11		
Dev. from curvilinear regr.	4	269.10	67.27	
Difference		1	1710.01	1710.01
	F-RATIO		25.42	
	SIGNIFICANCE LEVEL	0.0073		

Note that the F-ratio in this table is reversed from the usual regression case. Previously we calculated the F-ratio from $M.S._{regr}/M.S._{resid}$, with 1 and $n-2$ degrees of freedom. Now we use $M.S._{diff}/M.S._{dev}$ from curvilinear regr., with 1 and $n-3$ d.f. ($n-3$ because the 3 parameters of eq.(1.19) are fit to the data). In the rare case of using a cubic model, one would use the same procedure, but $n-4$ d.f. because a 4th parameter is fitted in the cubic (3rd degree polynomial) model.

It is worthwhile to look at the equations for residual S.S. on which the test is based. From eq. (1.10) the S.S. for linear regression is:

$$\text{S. S. Residual (linear regr.)} = \sum [y_i - (\alpha + \beta x_i)]^2$$

The corresponding S.S. for the quadratic (2nd degree polynomial) would be:

$$\text{S.S. Residual (quadratic)} = \sum [y_i - (\alpha + \beta_1 x_{1i} + \beta_2 x_{2i})]^2 \quad (1.20)$$

From these equations, it can be seen that an improved fit with the quadratic model should reduce the S.S. considerably. If not, then the F-ratio should be small and non-significant.

There are various difficulties in using this test on real data, mostly associated with the inaccuracies of censusing animals and the very real prospect that a population may cease to grow for a variety of reasons. Ecological data are like that! Some statistics books and editors advise checking assumptions before analyses are published. As noted previously here, such tests require more data than are ordinarily available, and may thus be misleading and contradictory.

Following the advice to test assumptions, one might well use the above test to see whether population growth data are linear or non-linear. It is worthwhile to conduct such a test as a way to explore the data. An Exercise asks the student to conduct these tests on actual data on growth of a number of populations. Theoretically, the outcome should be that the test will show nonlinearity and thus lead to using a transformation. In the real world, the results are confusing. The moral is that experience and accepted theory dictate the advisability of a transformation.

1.8 Basic models for population growth

Most ecology textbooks describe population growth by the familiar exponential model:

$$N_t = N_0 e^{rt} \quad (1.22)$$

Where N_t is population size at time t , N_0 the starting population size, and r the "instantaneous" rate of population growth. It is worth pointing out that a great many populations do not follow the commonly assumed model, inasmuch as they reproduce only during a short annual period, and thus follow what has been called a "birth-pulse" model, spurting up in numbers at the time of reproduction, and then decreasing through the rest of the year due to mortality. Eq. (1.22) describes continuous change, with reproduction and mortality assumed to be going on constantly in any short time period. A model closer to the truth is of the "compound interest" type:

$$N_t = N_0(1 + r)^t \quad (1.23)$$

Thus, where equation (1.22) describes a smoothly ascending continuous curve, eq.(1.23) describes a "step function" jumping up at specific times and then staying flat in the interim. Neither model is correct at all times, but they do agree at specific times. Figure 1.10 sketches out the likely actual time trend of a population, along with the results of eqs.(1.22) and (1.23). Either model can be described by $N_t = N_0 \lambda^t$, with λ representing e^r or $(1+r)$. When we use a log transform to represent population growth data as a straight line (thus performing "log-linear" regression), it is important to have in mind this interpretation of the slope of the regression represented by the two models. Note that eq. (1.23) is actually only defined at the time of reproduction or recruitment, but the plot (dashed line) connects these "jumps" by a straight line.

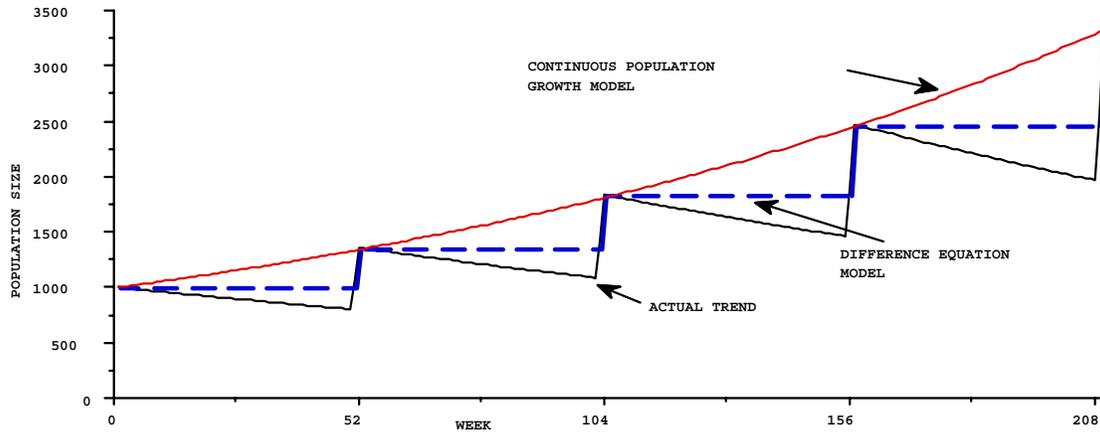


Fig. 1.10 The two population growth models of eqs. (1.22) and (1.23).

1.9 Testing for differences between regression lines

An essential feature of regression analysis is the ability to determine whether a number of fitted regression lines differ. We start out by considering whether the slopes (β_j) of several lines are significantly different. If not, then it is logical to test whether the intercepts (α_j) are different. This leads to the Analysis of Covariance, discussed in the next section.

Most of the data for testing equality of slopes comes from the calculations presented in Section 1.4. The main new feature lies in estimating a common slope. In order to compare the several slopes, we will first need to combine individual slopes to obtain a "pooled" value to compare with the individual values. This also can be obtained by weighting the individual slopes inversely by their variances. The weights come from the variance estimate for individual slopes, eq. (1.15). A basic assumption in assessing regression lines is that they all have the same variance about regression, as estimated by the residual (error) mean square of eq. (1.10). As always, if there is enough data it is worthwhile to test that assumption. Usually only gross differences can be detected with small to moderate sized data sets. If we assume a common variance (s^2), then the weights can be taken as:

$$w_i = \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1.24)$$

Thus the slope based on the widest spread of x-values gets the most weight, and the pooled slope becomes:

$$\bar{b} = \frac{\sum w_i b_i}{\sum w_i} \quad (1.25)$$

where we have k regression lines to analyze so the summations run from 1 to k. In the analysis, we pool familiar sums of squares for the k regression lines, namely:

$$SS_y = \sum (y_i - \bar{y})^2, SS_x = \sum (x_i - \bar{x})^2, \text{ and } SS_{xy} = \sum (y_i - \bar{y})(x_i - \bar{x})$$

and use these to arrive at a pooled estimate of the residual (error) sum of squares. The resulting mean square is then used as the denominator in an F-test, where the numerator is:

$$S.S.\text{diff} = \sum w_i (b_i - \bar{b})^2 \quad (1.26)$$

with k-1 degrees of freedom.

Table 1. 6. Calculations for eq. (1.26).

	Weights	b_i	$w_i b_i$	$(b_i - \bar{b})^2$	$w_i(b_i - \bar{b})^2$
Horse	17.5	0.193	3.376	0.0087	0.1525
Deer	28.0	0.488	13.664	0.0407	1.1394
Elk	42.0	0.191	8.008	0.0091	0.3838
	87.5	Sum	25.048	S.S.	1.6757
		b-bar	0.286		

Another example concerns a situation where it seems likely that the regression intercept (α) should be zero. The data come from a study of Hawaiian monk seals. These seals occupy 5 sites spread over about 1300 miles northwest of the main Hawaiian Islands, and are classified as Endangered under the Endangered Species Act. To monitor their abundance, "beach counts" are conducted annually on most of the sites. These amount to tallying all seals seen in covering all beaches on a site. Only a fraction of the seals using a site are ashore at any given count. However, individual seals can be identified by tags, scar patterns, and the use of temporary bleach marks. In those instances where many counts can be made over 6 weeks or so, it becomes possible to achieve a virtually complete tally of the population using the site. A further description of monk seal dynamics appears in Section 14.5 (Case Histories).

The analysis in this example thus contrasts the mean beach counts against population totals for 3 sites, using regression through the origin. Because α is now assumed zero, the regression model becomes $y_i = \beta x_i + e_i$. The least-squares estimate of β is;

$$b = \frac{\sum y_i x_i}{\sum x_i^2} \quad (1.29)$$

which is eq. (1.9) without the means, e.g., $\sum(x_i - \bar{x})^2$ is now $\sum x_i^2$. Apart from this change in definitions, the analysis (Table 1. 7) proceeds as in the previous example, with one other exception. Inasmuch as α is not included in the model, we use n-1 d.f. where regression analyses with 2 parameters (α and β) use n-2 d.f.

Table 1. 7 Data for a test of equality of slopes for 3 regression lines relating mean beach counts to total abundance for Hawaiian monk seals at 3 sites.

Source	SSx	SSxy	SSy	Slope	d.f.	S.S.	M.S.
KURE	63431	26670.2	11782.3	0.420	11	568.50	51.68
LAYS	6576511	98198.4	60015.1	0.301	9	283.36	31.48
FFS	1338292	410254.5	126842.6	0.307	4	1078.76	269.69
						1930.62	
Pooled	2059374.0	635123.1	198639.9	0.308	24	2764.196	115.17
		Difference between slopes		1	833.580	833.58	
							F 7.238
							Prob. 0.013

It thus appears that there is a significant difference among sites, with one site (Kure) having a significantly greater slope (b) than the other two, where the slopes are virtually identical. The two relationships appear in Fig. 1.12. The site with the largest total counts (French Frigate Shoals) contains many small islands, some of which are small enough that it has been difficult to approach seals for identification. The "total" counts at that site have thus not been considered complete, but the data for the recent 5 years (1991-1995) considered here now suggest that the apparent total counts do agree with the relationship between beach counts and totals at Laysan, suggesting that the FFS data may now approximate actual totals.

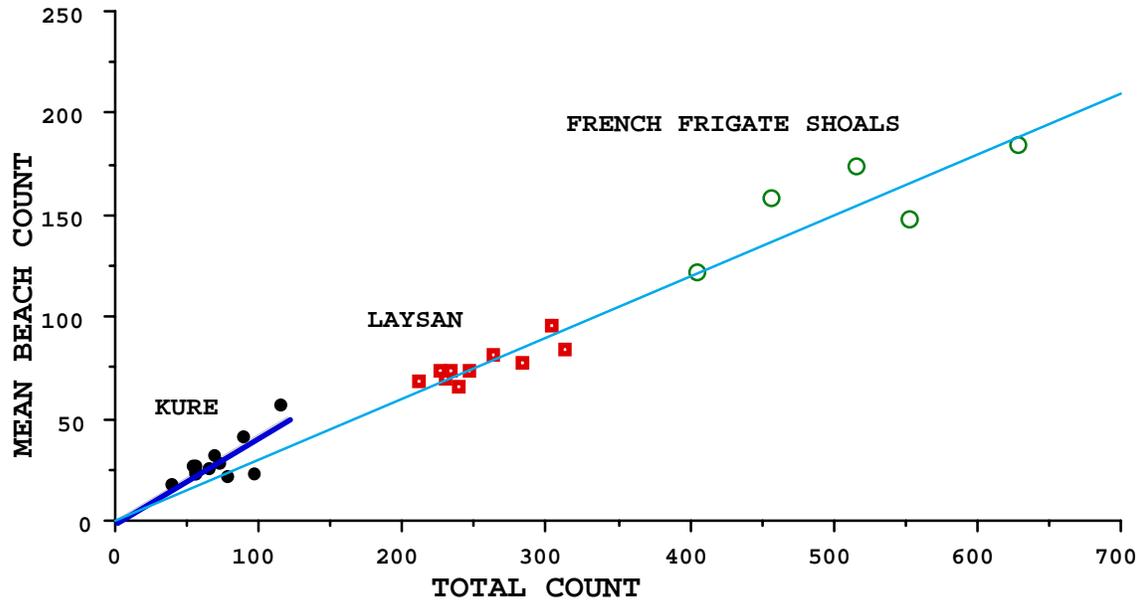


Fig. 1.12. Relationship between mean beach counts and total counts at three monk seal population sites. Regression through the origin for Kure is shown by a solid line, while the same regression for Laysan and French Frigate Shoals appears as a broken line.

One other issue illustrated by the monk seal data should be discussed here. This is the aggravating question of "outliers". In some data sets, there are points that seem evidently to lie well away from a trend evident in the bulk of the points. This is the case with the Laysan data. There are two years (Fig. 1.13) that are well away from the trend line (and were not used in the analysis of Table 1.7). Simple and direct methods are not available for deciding to exclude "outliers". However, in extreme cases like this one, we can simply consider the probability of such a deviation. The standard deviation of the distribution of points around the regression line for Laysan is the square root of the Mean Square of Table 1.7, which is $31.48^{1/2} = 5.6$. Deviations of the two suspect points from the regression line are 65 and 62 units, or about 10 standard deviations away from the line. Clearly these two deviations have an extremely low probability of arising by chance alone. Corroboration is also available in that the two points (they occurred in successive years) represent an increase in population size that is simply not feasible, and a subsequent decrease that surely would have been detected (dead seals) if it occurred.

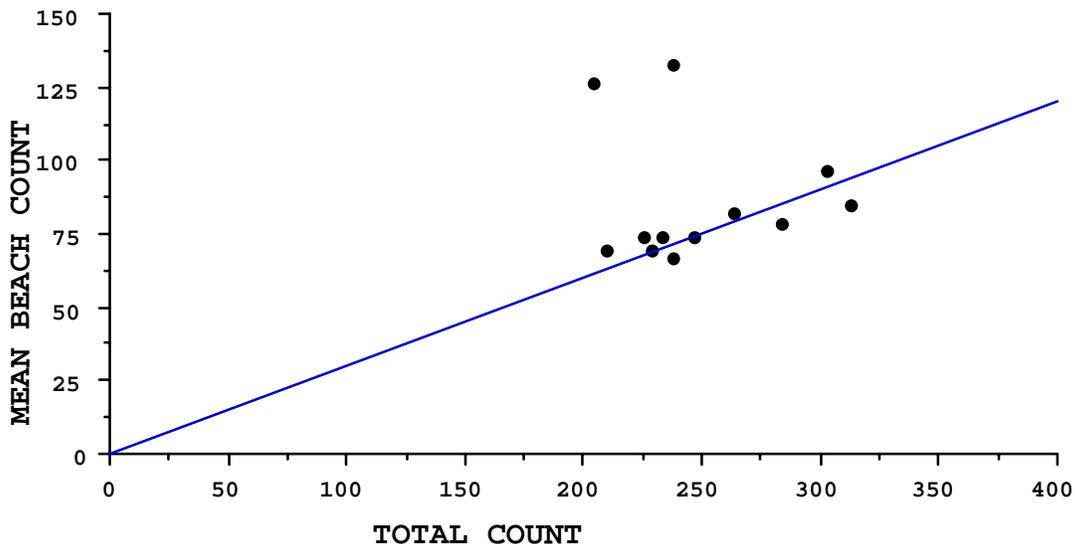


Fig. 1.13. Position of two aberrant counts at Laysan Island relative to the regression line and data from which it was calculated.

1.10 The Analysis of Covariance

The analysis of covariance depends on the availability of an auxiliary measurement linearly related to the variable of interest. Consider a one-way analysis of the yield (y_i) of fruit trees subjected to several different treatments (different types of fertilizer or perhaps insecticides) that presumably will increase yield. Yield of individual trees may vary with the size and location of the tree, so a useful auxiliary variable may be the yield (x_{ij}) of a given tree in the year before the treatments were applied. Hence, a one-way model without information from the auxiliary variable is:

$$y_{ij} = \mu_i + \epsilon_{ij}$$

but the auxiliary variable can be introduced by:

$$y_{ij} = \mu_i + \beta(x_{ij} - \bar{x} \dots) + \epsilon_{ij} \tag{1.28}$$

so that the adjusted mean for a given treatment becomes:

$$\bar{y} \dots_i = \mu_i + \beta(\bar{x} \dots_i - \bar{x} \dots) + \bar{\epsilon} \dots_i$$

The ANOVA for a covariance adjustment then tests whether adjusted means are significantly different. The dot notation is used with multiple subscripts to indicate which subscript is involved in averaging. Thus $\bar{x} \dots_i$ denotes the average over j for the i th group.

A key assumption in the analysis of covariance is that the same linear relationship holds in all of the treatment groups. Thus we need to use the methodology of Section 1.9 to test the hypothesis that β_i within treatment groups are not significantly different. Some investigators may proceed with the analysis without testing homogeneity of the slopes. This is not wise unless there is a good deal of prior experience on which to base such a decision. Inasmuch as both analyses depend on much the same computations, prudence calls for computing the results given in Table 1.5 and 1.7 in any case.

The data are arranged in the same way as in the previous section, but we here assume the same number of observations in each treatment group, giving a table like the following:

A		B		C	
x	y	x	y	x	y
x ₁₁	y ₁₁	x ₂₁	y ₂₁	x ₃₁	y ₃₁
x ₁₂	y ₁₂	x ₂₂	y ₂₂	x ₃₂	y ₃₂
⋮	⋮	⋮	⋮	⋮	⋮
x _{1j}	y _{1j}	x _{2j}	y _{2j}	x _{3j}	y _{3j}
⋮	⋮	⋮	⋮	⋮	⋮
x _{1n}	y _{1n}	x _{2n}	y _{2n}	x _{3n}	y _{3n}

To provide an example, a table of data from Snedecor and Cochran follows:

A		D		F	
x	y	x	y	x	y
11	6	6	0	16	13
8	0	6	2	13	10

5	2	7	3	11	18	
14	8	8	1	9	5	
19	11	18	18	21	23	
6	4	8	4	16	12	
10	13	19	14	12	5	
6	1	8	9	12	16	
11	8	5	1	7	1	
3	0	15	9	12	20	
Means	9.3	5.3	10	6.1	12.9	12.3

From the ANOVA for simple regression we had the following results (eq.(1.11):

$$\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Total S.S. - Regression S.S. = Error (Residual S. S.)

The error term can be written in various ways

$$\begin{aligned} \sum_{i=1}^n (y_i - \hat{y}_i)^2 &= \sum_{i=1}^n [y_i - (a + bx_i)]^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - b^2 \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - \frac{[\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})]^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{aligned} \quad (1.29)$$

with the last result being most useful here. It is obtained by using the definition of b in developing eq. (1.29) from equation for the Residual (Error) Sum of Squares above. The above calculations are expressed for one group of data, so in dealing with several groups below, a subscript for the jth observation in the ith group needs to be added.

The calculations proceed by computing the 3 components of eq.(1.29) and arranging them in an ANOVA type of table in which the Total S.S. is calculated from the entire set of data, using overall means of x and y, e.g. with the other values SSy, and SSxy calculated in the same manner. Thus,

$$SSy = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})(x_{ij} - \bar{x}_{i.}).$$

The Error line is calculated by using the group means, e.g.,

$$SSx = \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2$$

The Between S.S. are readily obtained by subtracting the line for Error S.S. from Total S. S. These calculations then give the following table from the data above:

Source	d.f.	SSx	SSxy	SSy	M.S.
Total	29	665.9	731.2	1288.7	
Between	2	72.9	145.8	293.6	
Error	27	593.0	585.4	995.1	36.86

The M.S. due to error is calculated from $SSy/d.f. = 995.1/27 = 36.86$ in the Error line of S.S. just as it would be done without the auxiliary variable. The other entries in the table are needed to obtain a reduction in the error sum of squares as shown below.

The "reduction due to regression" is obtained from $(SSxy)^2/SSx$ in the Error line, and is subtracted from the Error sum of squares as computed without the auxiliary variable, giving an estimate of error mean square adjusted by the regression data. The complete calculation of an adjusted error mean square is thus:

Source	d.f.	SSx	SSxy	SSy	M.S.
Total	29	665.9	731.2	1288.7	
Between	2	72.9	145.8	293.6	
Error	27	593.0	585.4	995.1	36.86
Reduct. due to regr	1			577.9	
Dev. from regr	26			417.2	16.05

An estimate of a common slope is also obtained from the error line,

$b = SSxy/SSx^2 = 585.4/593.0 = 0.987$. This value then can be used to get adjusted values of \bar{y} from the following:

$$\bar{y}_{i,adj} = \bar{y}_i - b(\bar{x}_i - \bar{x}..)$$

The adjusted mean for the first group of data (group A in the table above) is thus:

$$5.3 - 0.987(9.3 - 10.73) = 6.71 = \bar{y}_{i,adj}$$

The results of the covariance adjustment can then be assembled to produce a covariance-adjusted F-test, as in the following table:

Table 1.8 Covariance F-test in one-way classification

	d.f.	SSx	SSxy	SSy	Deviations from regression				
					Reduc.	d.f.	S.S.	M.S.	
Treatments	2	72.867	145.8	293.6	577.9	26	417	16.05	
Error	27	593.000	585.4	995.1	802.94	28	486		
T+E	29	665.867	731.2	1288.7			2	68.6	34.28

The F-ratio is $34.28/16.05 = 2.14$ with 2 and 26 d.f. and does not suggest a significant treatment effect ($P = 0.14$).

The whole purpose of the exercise is to get a more sensitive F-test of main effects than would be possible without the auxiliary variable. Such an improvement depends, of course, on the presence of a significant linear relationship between the variable of interest (y_i) and the auxiliary variable (x_i), and this relationship needs to be checked out first (i.e., do regressions on the data in each group (A, D, and F) first).

1.11 ANOVA as a regression model

To sketch out a basis for doing an analysis of variance with a regression model, we need the concept of a "dummy variable" which is simply a variable that takes only values of 0 or 1. Consider the multiple regression model:

$$y_i = \mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i}$$

and let $x_i = 1$ if y_i belongs to a particular group in a one-way ANOVA and 0 otherwise. Then we can write:

$$\begin{aligned} y_1 &= \mu + \beta_1 \\ y_2 &= \mu + \beta_1 \\ y_3 &= \mu + \beta_1 \\ y_4 &= \mu + \beta_2 \\ y_5 &= \mu + \beta_2 \\ y_6 &= \mu + \beta_2 \\ y_7 &= \mu + \beta_3 \\ y_8 &= \mu + \beta_3 \\ y_9 &= \mu + \beta_3 \\ y_{10} &= \mu + \beta_4 \\ y_{11} &= \mu + \beta_4 \\ y_{12} &= \mu + \beta_4 \end{aligned}$$

and thus have a regression model conforming to a one-way ANOVA with three observations in each of 4 groups, giving the general model of $E(y_i) = \mu + \beta_i$, as is appropriate for one-way analysis of variance. Draper and Smith (1998) give extensions to two-way and higher analysis and methods of fitting. The approach is likely not of much importance here, but is mentioned to emphasize an earlier remark that models of the multiple regression type can be used for a wide variety of purposes, often subsumed under the heading of "General Linear Hypotheses".

1.12 Stepwise regression

This is an approach to regression that permits adding variables one step at a time while searching for the "best" model for a given data set. Consider the test for curvilinearity of Section 1.7. We first fitted a linear regression of the form $y_i = \alpha + \beta_1 x_{1i}$ and then extended the model to become a second degree polynomial $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i}^2$, using multiple regression to fit the model. We then tested for a significant "improvement of fit" by comparing the reduction in Sum of Squares obtained by subtracting the deviations from curvilinear regression (Residual S.S.) from the deviations from linear regression, and tested significance of the improvement by an F-test. We noted that the process could be extended to a third-degree polynomial $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{1i}^2 + \beta_3 x_{1i}^3$ to test for a more extreme curvature. We used multiple regression to fit the models, letting $x_{2i} = x_{1i}^2$, (and $x_{3i} = x_{1i}^3$ if the model were extended to test the further improvement of adding a "cubic" term). This kind of procedure is employed in stepwise regression, but is not, of course, restricted to polynomials. Any series of variables can be tested successively for the improvement of fit produced as each new variable is introduced. Computer programs are available that will test all combinations of a set of candidate variables but the results are practically guaranteed to be misleading, as enough manipulation will almost always produce a "good fit". One should use stepwise regression only when there is a logical sequence of models to test, and even then it is likely that the final model will be "over-fit" (i.e., have too many independent variables). One useful approach is to develop a model on half the data and check it on the other half. Usually, ecologists do not have enough data to hold half of it in reserve while studying a model. An alternative is known as "cross-validation". In it a series of fits are used and each observation is left out in turn, and used to check the error variance estimate from the fitted model. Such a test is "computer-intensive", i.e., depends on the ability of the modern computer to

conduct many calculations in a short time. Anyone planning to use stepwise regression should consult references like Draper and Smith (1998) first.

1.13 Logistic regression

This is a form of regression analysis developed for data of the binomial form, i.e., in which the variable of interest is either 1 or 0 (or "yes" or "no", "present or absent", etc., which can be coded as 1 or 0). Usually we express results as a proportion, e.g, the proportion surviving after some time interval or some treatment. Logistic regression originated in the field of bioassay, in which the response to a given dose of some substance is studied quantitatively. If one plots the response (proportion surviving or otherwise responding to some treatment) against the dose (often quantity of some substance given an individual) the resulting curve is usually sigmoid (s-shaped). The cumulative normal curve provides a convenient s-shaped model, and is used in bioassay in "probit" analysis. Details of methods used for bioassay are given by D.J. Finney (Statistical Method in Biological Assay, 3rd Ed. 1978, Charles Griffin and co., Ltd. London).

Joseph Berkson proposed using the logistic function as a bioassay model in 1944. The basic model is:

$$P = \frac{1}{1 + e^{-(a+bx)}} \quad (1.30)$$

where P denotes the dependent variable and x is the independent variable ("dose" in bioassay). Because P is a proportion,

$$Q = 1 - P = \frac{e^{-(a+bx)}}{1 + e^{-(a+bx)}}$$

and we can now consider the ratio of P and Q:

$$\frac{P}{Q} = e^{\alpha + \beta x} \quad (1.32)$$

The ratio of P to Q is sometimes called the "odds ratio", no doubt because it expresses the odds for a particular outcome.

Now the natural logarithm of this "odds ratio" {eq.(1.32)} is a linear function,

$$\ln\left(\frac{P}{Q}\right) = \alpha + \beta x \quad (1.33)$$

This is called the "logit" transformation.

There is an interesting sidelight to the logit transform. Consider a table of proportions (e.g., several species of plants classified by whether they have flowers, fruits or neither). One can then calculate the natural logarithm of the "odds ratio" and analyze the linear model of eq. (1.33). This is termed log-linear regression by some authors and can be extended to behave like the analysis of variance. It has been used largely in the social sciences, but could well be of interest in ecological circumstances where one must analyze tables of proportions (or tables in general, for that matter). It should be noted that we will also use the term "loglinear regression" to refer to the log transform of eq.(1.22).

Example 1.3 An example of logistic regression

In aerial counts of wildlife populations, the number of individuals in a group has a marked effect on visibility. This has been studied by using animals with attached radiotransmitters and recording the frequency of observation of groups containing these individuals. Such a study of elk has been used to correct for visibility (M. D. Samuel et al. 1987. Visibility bias during aerial surveys of elk in northcentral Idaho. *Journal of Wildlife Management* 51:622-630). The following table shows the data (only small samples were available so that larger groups had to be combined).

Table 1.9 Sighting data from an aerial survey of radio-marked elk.

Group size	Missed	Seen	Logit Proportion transformation	
			$\frac{\text{seen}}{\text{seen}}$	$\log_e(P/O)$
1	18	5	0.217	-1.281
2	7	6	0.462	-0.154
3	5	5	0.500	0.000
4	4	6	0.600	0.405
5	4	9	0.692	0.811
6	6	4	0.400	-0.405
11	3	14	0.824	1.540
23	0	10	1.000	

The simplest way to fit this data is to use eq. (1.33), i.e., regress the logit values (right-hand column) against x . In this case, the investigators used the logarithm of group size in their analysis, so we use $\ln(\text{group size})$ for x in Fig. 1.14, which shows the regression fit.

Due to the fact that the independent variable is from a binomial distribution the linear model implied by eq. (1.33) does not give the best fit to the data. Instead, the technique of maximum likelihood estimation is recommended. If we assume a particular frequency distribution (probability distribution function in Section 1.2) underlies a set of observations, then it may be possible to find expressions that often minimize the variance of an estimated quantity. Methods of mathematical statistics are required to derive such estimators, but many of the commonly-used estimates are also maximum likelihood estimates. In the present case, there is no simple expression for estimating the parameters of eq. (1.32) so that an iterative method is required to solve the maximum likelihood equations. The method used here is due to J. Berkson (Tables for the maximum likelihood estimate of the logistic function. *Biometrics* 13:28-34, 1957). Maximum likelihood estimates for logistic regression can also be obtained in some of the available statistics programs (e.g., SYSTAT).

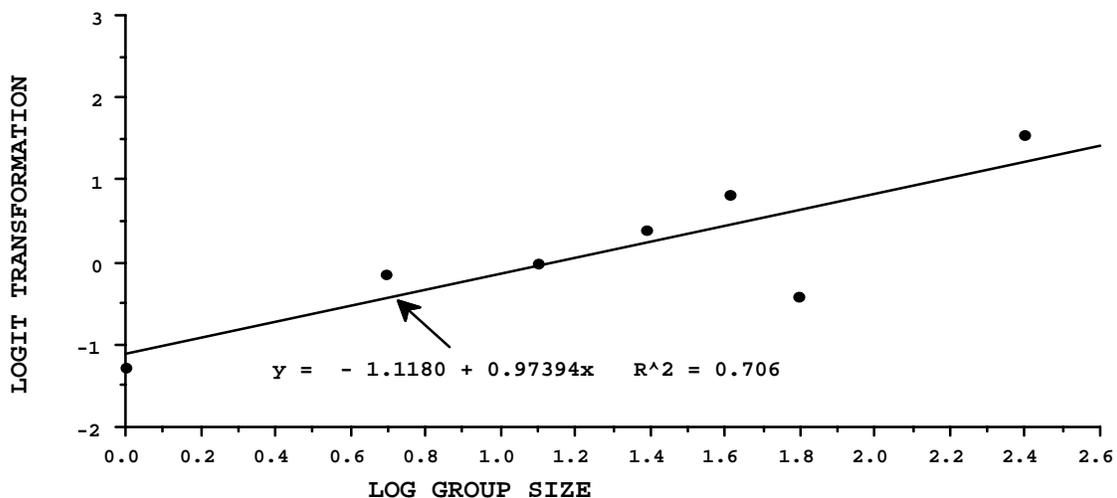


Fig. 1.14. Regression of logit values on logarithm of group size from aerial survey of elk.

The parameters obtained from the regression analysis (Fig. 1.14) are $\alpha = -1.118$ and $\beta = 0.974$, while those obtained from the maximum likelihood fit are somewhat different, being $\alpha = -1.305$ and $\beta = 1.155$. Fits to eq. (1.30) are not substantially different (Fig. 1.15).

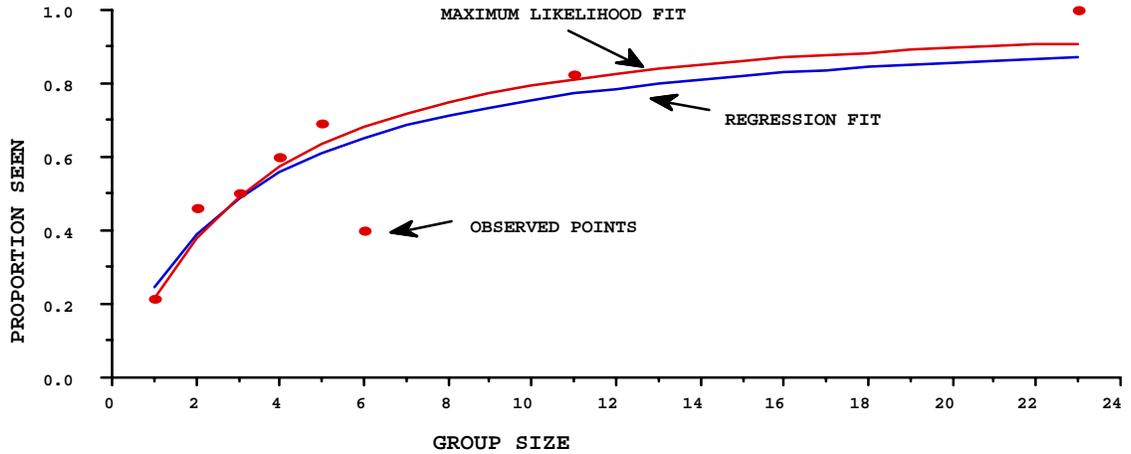


Fig. 1.15 Fits of eq. (1.30) to observed data on elk sightability using regression (eq.(1.33) and maximum likelihood methods.
Example 1.4

Two further examples (Fig. 1.16) are based on reproductive rates in Hawaiian monk seals at two sites. The curves were fitted as above, using regression and maximum likelihood estimates.

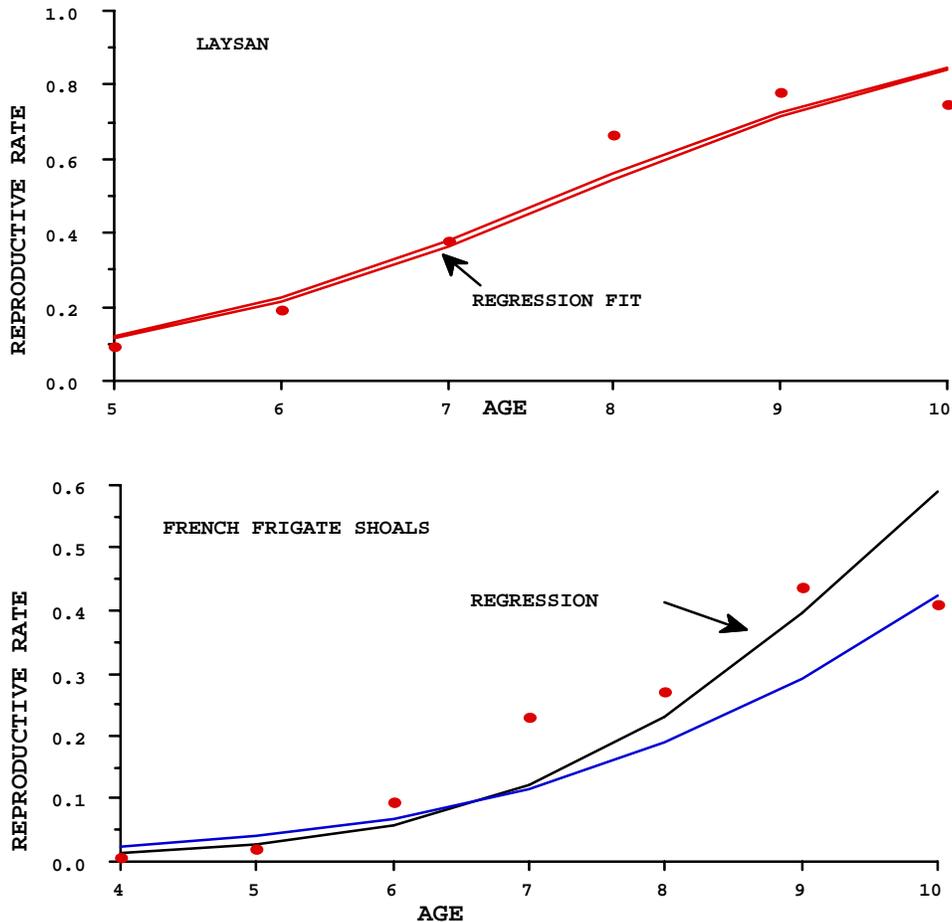


Fig. 1.16 Logistic fits of reproductive rates against age of the female for Hawaiian monk seals at two sites.

In the upper curve, it appears that the regression and maximum likelihood methods give about the same results, while neither provides much of a fit in the lower curve. Deteriorating conditions (poor food supplies and survival) at the site may be changing the curve, so that it does not represent a stable situation. Circumstances at the site shown in the upper curve have been reasonably good, but there is no particular reason to suppose that reproductive rates should follow a logistic curve.

For comparison, some data on judging sound intensity were fitted by the two methods (Fig. 1.17). These data appear to fit the logistic very well, and the two methods of estimation give virtually indistinguishable results.

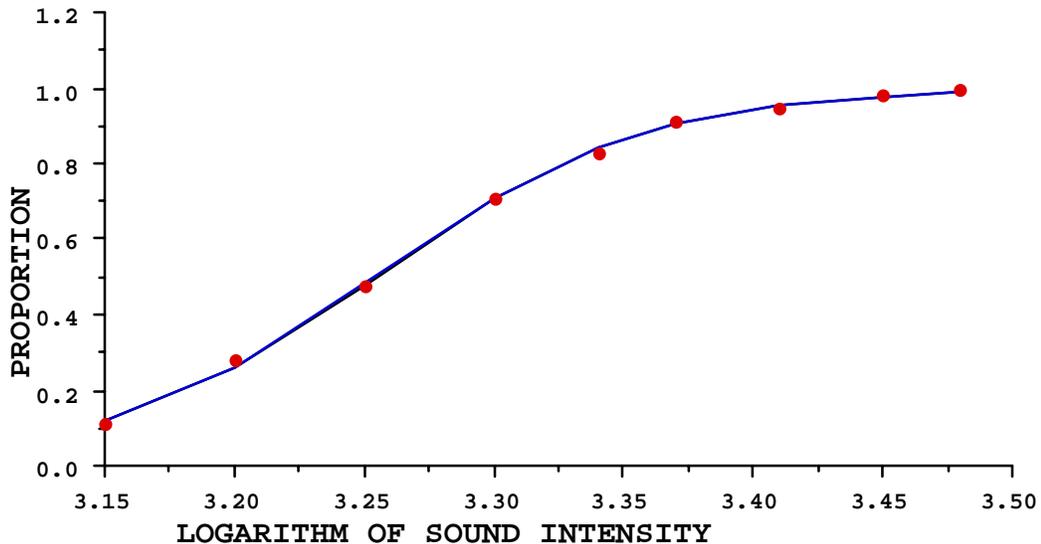


Fig. 1.17. Logistic curve fitted to data on judging sound intensity.

1.14 Locally weighted regression

When there is no suitable model for a curve, locally weighted regression provides a way to fit a smoothed line. The method is variously called "loess" or "lowess". Some authors use "loess", but ecologists will no doubt be confused by the implication of wind-deposited soil!. Weighted linear regressions are fit at each point on the graph (e.g., if the data span 30 years, then such regressions are fit at each of the 30 years) by selecting data points in the immediate neighborhood of each point on the x-abcissa. The number of points in each such neighborhood might be taken to be, say, about 30% of the total number of observations. However, this can be varied in the fitting program, and depends on the purpose at hand. If one wants a thorough smoothing, then 50% or more of the points might be used in each regression. If the smoothed curve is to follow the data point closely, then a small fraction, perhaps as little as 10%, of the points should be used in each fitting. Experimentation with the fitting program will help in developing an approach for a particular data set. Weights diminish by a cubic function, so points very near to the selected point get by far the most weight. The fitted regression line determines only the y-value for the selected abscissal value. In effect, the technique behaves much like a moving average, but has various advantages. Programs to produce lowess fits are available. SYSTAT has a routine for lowess fitting in the plotting routine (after loading the data in a file, bring up "plot", and select the "smooth" function. It will then be necessary to indicate the fraction of the data points to use in each neighborhood). The lowess method was developed by W. S. Cleveland (Journal Amer. Statistical Assoc. 74:829-836).

The smoothed line in Fig. 1.18 illustrate the technique. This approach to smoothing is preferable to the usual moving-average smoothing because it does not leave blanks at the end of the series, and uses what seems to be a better averaging approach. The lowess technique can be illustrated by smoothing French Frigate Shoals monk seal beach count data. At each point along the line (here, each year) the nearest n points are used to form a weighted linear regression (9 points were used in producing Fig. 1.18). The regression line is used only to determine the smoothed value for the given point. Inasmuch as the weights and the regression line must be computed for every point used along the x-axis (the years 1957 to 1993 in the present example), enough calculations are involved to make use of a computer desirable.

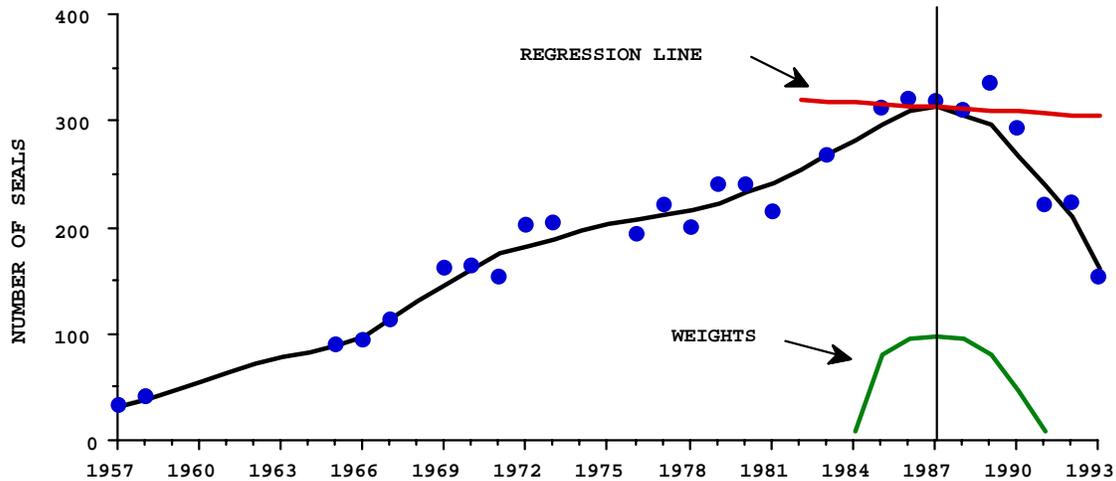


Fig. 1.18. Locally-weighted regression line ("lowess" smoothing) for the French Frigate Shoals monk seal beach counts. For each year on the graph, a weighted linear regression is computed from the n nearest points, with the contribution of each point weighted by a cubic function of distance of the data point from the base point. The regression line for 1987 is shown on the graph, along with the weights assigned to the 9 nearest points. The regression line determines only one point on the smoothed line.

1.15 Non-linear least-squares

The method of least-squares was discussed in Sec. 1.4, and eq.(1.6) was used to develop least-squares estimates for linear regression. The same approach can be used to fit non-linear functions, starting with the same equation for sum of squares:

$$S = \sum [y_i - f(x)]^2$$

where $f(x)$ is now some non-linear function, such as the logistic function of eq. (1.30). One could find a minimum for the sum of squares, S , by a direct search routine. This is labor-intensive, and there are various computer programs that do the job very quickly and efficiently. Some of these call for partial derivatives of the function (used to "linearize" the function so that the approach to a minimum can be done in successive iterations). Others use numerical approximations to the partial derivatives, or direct search routines. SYSTAT contains two such routines under the "nonlin" function. It requires that a model be furnished, but this can be written in the notation used in EXCEL (really statements in BASIC language, which underlies EXCEL). Thus eq. (1.30) is entered as:

$$P = 1/(1 + \text{EXP}(-(A+B*X)))$$

The data need to be entered by using the Editor function (or can be read in from an EXCEL file, or copied to the Editor via a clipboard). Names used for variables (P, X) above are used as column headers in the data file, and the SYSTAT fitting routine recognizes the other labels (except built-in functions like EXP) as variables to fit (such as A and B above). Trial values can be furnished (i.e., rough estimates of A and B) and the number of iterations can be set (these have built-in "default" values). It may be necessary to use trial values if the program doesn't converge in, say 20 iterations (the default value), but further iterations can be

tried, first. Since the program is iterative, it may get stuck in various ways, and it is then desirable to quit, and start over with different guesses at starting parameters.

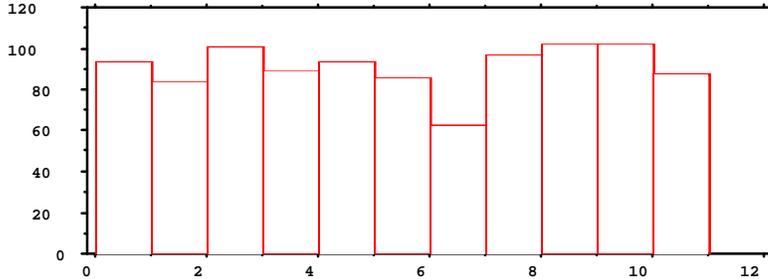
1.16 Exercises

1.16.1 Coin-tossing. Students should try a coin-tossing experiment like the one reported in example 1.1. Put 10 coins in a jar and make 100 tosses, recording the number of heads in blocks of 10. Make a frequency distribution and compare it with Fig. 1.1. Try another set of 100 and compare the two frequency distributions. Compute the sample means and variances, and compare them with the theoretical values.

1.16.2 Simulating the binomial on a computer. Coin-tossing gets tiresome after awhile, and it is important to look at a different probability model. In order to get large samples without the tedium of mechanical approaches, we can resort to the computer. Students familiar with a programming language will likely prefer to write a simple program. However, useful results can be obtained in EXCEL and are readily in reach of those without programming experience. Those with only a passing experience with EXCEL may have to resort to the HELP function (or a colleague with experience) but it is important to carry out the following exercise because it should provide a capsule view of “monte carlo” simulations. Also, the next two chapters on bootstrapping depend on use of EXCEL. Insert the statement “=RANDBETWEEN(0,1)” in a cell in an EXCEL spreadsheet, and copy down to fill 10 cells in a column. This generates a series of 0’s and 1’s with probability 1/2 of getting either. Now copy the row to the right for 100 columns (it is convenient to use the automatic numbering system in a column above the 10 entries to keep track—a handy little number pops up beside to indicate how many numbers you have entered). Now sum the columns (use the summation function in the legend at the top of the sheet). This row of numbers (the sums) is now equivalent to the table of data in Example 1.1. Now use the histogram procedure (in the Tools menu) to construct a histogram of frequencies of results. These should approximate the bars in Fig. 1.1. Note that every time you make a change in the worksheet it recalculates the table of random values (this function can be turned off). It is worthwhile to calculate several histograms just to get a notion of how variable the outcomes are. Next calculate the expected values from eq.(1.1). Find the factorial function (“FACT” in Math and Trig functions). Actually, all you need to know is that FACT(5) gives the value of 5! Use this function to calculate the factorial part of eq.(1.1) next to a column numbered 0 to 10. Then enter the rest of the equation in the next column (because $p=0.5=1-p$ these entries will all be the same, but we’ll use the approach for a case where p is not 1/2 below). The product of the two columns gives the proportions of eq.(1.1) which add to unity. Now multiply 100 times the proportions, and you have the expected values, which should approximate what you have in the histograms. The Chart Wizard in EXCEL will plot expected and observed values (you need to look under “Custom Types” to find one that plots a line and bars). One last chore is to recalculate the expected values using a value of $p=0.9$ which gives a distinctly asymmetric graph. It is always useful to put the numerical value of p above the calculations and use the “\$” (e.g.,\$A\$30) notation to denote p in calculations for the equation. This lets one experiment with different values of p . Students should save a worksheet with the above calculations in order to have it for further reference when we consider other frequency distributions.

1.16.3 Random sampling There will be a great deal of emphasis on random sampling in this course. A relatively new topic in statistical methodology called bootstrapping will be used extensively. It depends on random sampling with replacement. Courses and books on sampling methodology usually depend on sampling without replacement. Consider using a number of sample plots to make counts of plants in order to estimate overall density of some species of plant. Such plots should be located at random in order to assure an unbiased estimate of density, and secure a reliable estimate of variance. Ordinarily, an investigator would find some way to assign a number to all possible plots in the area to be studied, and locate the sample plots by consulting a table of random numbers. If the same plot is drawn twice, it would not be counted twice, as this usually makes no sense. Hence we describe this as sampling without replacement. Textbooks on sampling show that it usually doesn’t make much difference whether we do in fact sample with replacement, inasmuch as the sample usually is a small fraction of the total population. Bootstrapping, however, depends on sampling with replacement as a way to reflect the underlying frequency distribution. Consequently, most of our samples will be with replacement. We will be taking repeated random samples with replacement of a data set. The individual entries in the data set will be in a computer file, and we will randomly select individual entries from this file. It is convenient to number the data items from 1 to n , and we then need to generate random numbers. To illustrate the approach, enter “RANDBETWEEN(1,10)” in a

cell in EXCEL and copy down the column for 100 entries. Make a histogram of the data, as in Example 1.16.2. This is a sample from a uniform distribution, i.e., a frequency distribution where the probabilities are all equal. It is the distribution underlying random sampling. It is easy to extend the process to, say, 1,000 draws as in the frequency distribution plotted below. Note that it is still quite variable, even with 1,000 draws. Make a graph of your data like the following using the Chart Wizard and post it on a spreadsheet with the calculations.

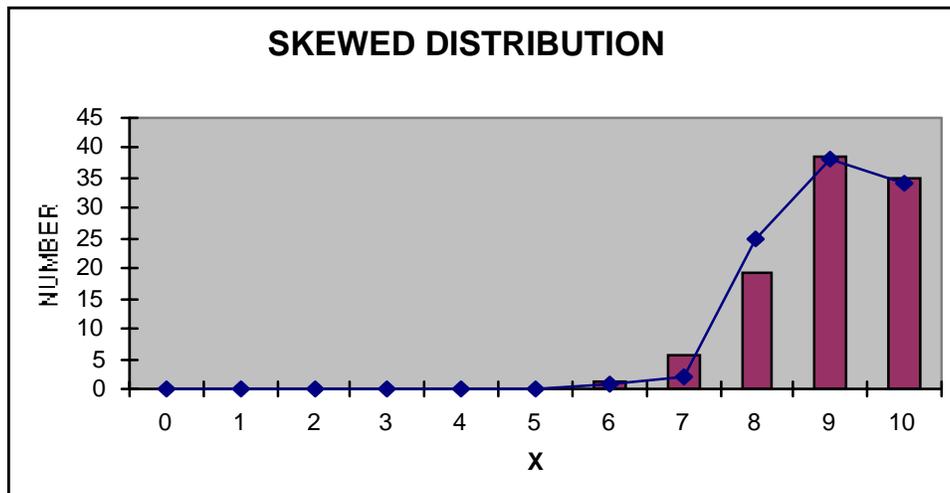


1.16.4 Simulating a discrete skewed distribution. In Exercise 1.16.2 students were asked to calculate expected values for a binomial frequency distribution [eq.(1.1)] with $p=0.9$. A skewed frequency distribution is not hard to simulate, requiring two changes to the methods used in Exercise 1.16.2. Instead of `RANDBETWEEN(0,1)` we use `=RAND()` (don't put anything in the parentheses) which provides random numbers between 0 and 1. We also need an "IF" function which is the basis for a lot of computer work. It evaluates an expression and chooses between two output values, depending on whether the expression is true or false (there are a number of different expressions working along these lines, but we use the simplest here). Set up a spreadsheet with a column of 10 values of `=RAND()`, and copy it to the right 100 times. We again need a numerical value of p above this table for reference, which may be say 0.9. If the first entry in the first column is in position, say, D9, then in the column just below this first column the first entry should be `=If(D9>=A3,1,0)` where the value of p is in A3$. Copy down 10 and across for 100 columns and sum these entries. The IF function checks to see if the entry in D9 exceeds p and enters 1 if true and 0 if false. The sums then provide the basis for a histogram of a skewed discrete distribution. Make histograms with $p=0.1$, and $p=0.5$. Compare the histogram with $p=0.5$ with the one you made in Exercise 1.16.2. Make a new calculation of eq.(1.1) with $p=0.9$ and compare it with the histogram with $p=0.1$ (actually you should have made one in Exercise 16.1.2 and need only copy it over to this worksheet for comparison. Plot the data in Chart Wizard (expected and observed values). It should look like the following graph:

1.16.5. Do the algebra to calculate the expected value of eq.(1.1) as given in the right side of eq.(1.2).

1.16.6 Simulating a continuous skewed distribution. A continuous random variable is one that has values in the real domain. For our purposes, this means values like those generated by `RAND()` -- any number within the range considered (i.e., from 0 to 1). We will consider one way to generate random variables from an exponential distribution here. Consider the function:

$$F(x) = 1 - e^{-\beta x} \tag{1.34}$$



This is an example of a cumulative distribution or cumulative distribution function. It takes values from 0 to 1, and has one parameter, β , which controls the rate at which the function approaches unity. The graph below shows the function for $\beta = 0.1$ and $\beta = 0.5$.

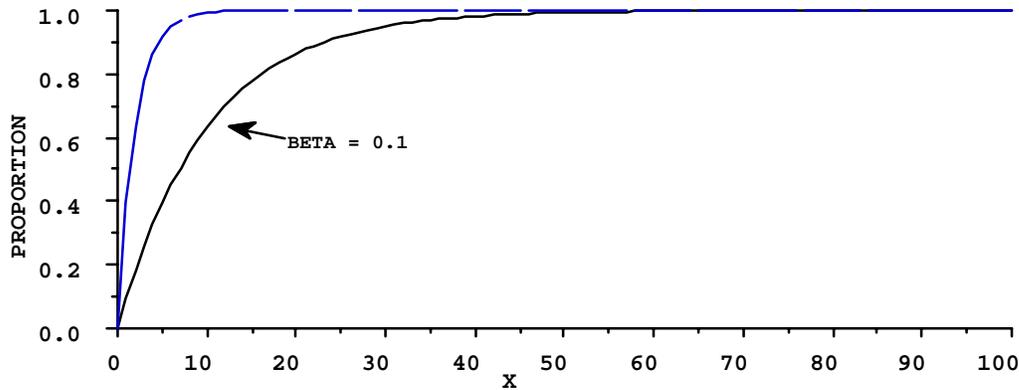


Fig. 1.19 Plot of cumulative distribution function for the exponential distribution for values of $\beta = 0.1$ (solid line) and $\beta = 0.5$ (broken line).

We use the cumulative distribution function here because it takes values from 0 to 1, and we can take a random sample from that range (using `RAND()`) and translate that to find the corresponding x value, by rearranging eq. (1.34) as

$$x = \frac{1}{\beta} \log_e(1 - F(x)) \quad (1.35)$$

Thus the procedure is to draw a random sample of values from `RAND()` and look up the corresponding values of x . Eq. (1.34) is the integral of an exponential distribution over the range 0 to x , hence the name "cumulative". To compare the outcomes of a simulation with the equation for the frequency distribution, one runs a simulation as described in Exercise 1.16.4, and plots the results. Differentiating the cumulative yields the frequency distribution:

$$f(x) = \frac{dF(x)}{dx} = \frac{d[1 - e^{-\beta x}]}{dx} = \beta e^{-\beta x} \quad (1.36)$$

Students whose calculus is a little rusty may want to look up the formula for finding a differential of an exponential; others may want to accept the statement without derivation. We need the right side of eq.(1.36) only to be able to compare simulation outcomes with the theoretical model, given in the figure below. Produce a column of 1000 random variables [$F(x)$] from `RAND()` and convert them with eq.(1.35), make a histogram of the results (using 30 "bins") and then calculate the expected values by multiplying eq.(1.36) times 1000. Plot these as before and see how your result compares with the graph below.

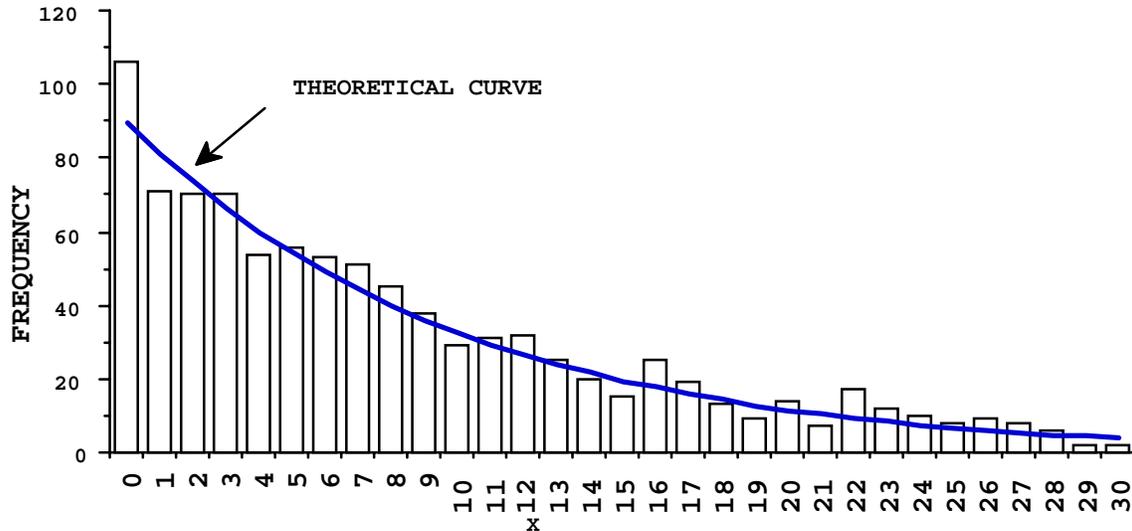


Fig. 1.20. Simulated exponential data compared to theoretical curve.

1.16.7 Simple linear regression. Data on counts of deer on a study area are given below. Fit the linear regression of Fig. 1.3 by using eqs. (1.9). This is readily done in EXCEL (in fact, EXCEL has a regression fitting routine which we will use for additional exercises, but students should do the calculations directly from the definitions in order to see how they “work” and then check by using the built-in fitting routine). Some graphics programs will also do the fitting automatically.

Year	Number of deer
x_i	y_i
1	10
2	21
3	52
4	71
5	97
6	146
7	212

1.16.8 Check the fact that a and b give minimum values of eq.(1.6), the sum of squares, for the deer data of Exercise 1.16.7. Copy the results of Exercise 1.16.7 into a new worksheet and compute eq. (1.6) for a and b , setting up the worksheet so that a and b are listed as separate entries on the worksheet as shown below. Then vary a and b by small amounts and write down the resulting sums of squares in the table. That is, make a table like the following and fill in the entries. It is easiest to first make your entries in pencil as transferring them individually to a summary table in EXCEL calls for a lot of tedious use of “Paste Special” in the menu, and/or provides opportunities to forget which cell you were working with. You should find a minimum in this table. If you want to try to get closer to the values of a and b found in Exercise 1.16.6, make a new table with fractional values in the row and column headings (e.g., 31.1, 31.2, etc.) and fill in the new table. This approach provides a device that is sometimes useful to solve a pair of more complex equations without needing to use a non-linear least-squares fitting routine. It is tedious unless you can guess reliably in advance just which part of the “Sums-of-Squares” space the answer lies. But the purpose here is just to show how things work.

Sums-of-Squares table(eq.1.6)

		b				
		30	31	32	33	34
a	-39		2046			
	-40					
	-41				2520	
	-42			1987		
	-43					
	-44					

1.16.9 Use EXCEL and eq. (1.11) to calculate ANOVA for a regression equation for the data of Exercise 1.16.7 and compare your results with those given in Table 1.1. Now use the EXCEL regression program (found in the same group of analysis tools as are the ANOVA programs) to see how it works, and add the results to your direct computations.

1.16.10 Compute the correlation coefficient for the deer data from eq. (1.12). It can also be directly computed using a function CORREL found in the functions menu.

1.16.11 Compute s_b^2 from eq. (1.15) for the deer data. Now compute it assuming that you have 3 observations (9,10,11) from year 1 and 4 observations (207,210,212, 219) at year 7 (and no observations for years 2,3,4,5 and 6). You will need to recalculate everything for the new data. What do you conclude about the effect of this arrangement of the data on s_b^2 ? Would you recommend this approach? Why?

1.16.12 Compute confidence limits for b from eq.(1.16) using the following set of data. Show details of your computation (i.e., the components of the calculation on a spreadsheet).

1	2.86
2	0.90
3	1.56
4	3.85
5	1.62
6	4.39
7	3.66
8	3.95
9	4.45
10	4.50

Note that the α in eq.(1.16) is not the same as α in the regression model. It is standard notation for the probability level. Use $\alpha=0.05$ here. You can obtain the needed t-value from the functions in EXCEL (f_x on the Toolbar) which is $TINV(\alpha,d.f.)$ where α is the desired probability for a 2-tailed t-test. You can run the regression analysis in EXCEL to confirm your results.

1.16.13 Multiple regression. Calculate a multiple regression equation on the following data, using eqs.(1.19) and check your results in EXCEL. The data were used in an early effort to construct an index of abundance for grizzly bears in Yellowstone National Park. Use the logarithm of the count as y and "Yr." As x1 and "Freq. Sight" as x2. It is important not to use the actual 4 digit year as x1 because it can cause a loss of accuracy when larger data sets are involved.

Year	Count	ln count	Yr.	Freq. sight.
1976	17	2.8332	1	1.64
1977	13	2.5649	2	1.50
1978	9	2.1972	3	1.28
1979	13	2.5649	4	1.08
1980	12	2.4849	5	1.40
1981	14	2.6391	6	1.58
1982	11	2.3979	7	1.62
1983	13	2.5649	8	1.20
1984	17	2.8332	9	2.29
1985	9	2.1972	10	2.00
1986	25	3.2189	11	3.12
1987	13	2.5649	12	1.64
1988	19	2.9444	13	2.12
1989	16	2.7726	14	1.86
1990	25	3.2189	15	1.95
1991	24	3.1781	16	2.65
1992	23	3.1355	17	1.65
1993	20	2.9957	18	1.67
1994	20	2.9957	19	1.47

1.16.14 Perform the test for curvilinearity described in the text (Sec. 1.7) and illustrated in Table 1.5 on the following sets of data. Make a spreadsheet containing the ANOVA tables (as in Table 1.5); note that the deer data are also included here so you have an example of the expected results at hand) and discuss the results as they apply to the notion that one should test for the assumptions before doing an analysis. Do the tests of the ANOVA tables provide convincing evidence of nonlinearity in the data?

Year	Horses	Year	Deer	Year	Elk
1	340	1	10	1	3172
2	423	2	21	2	4305
3	482	3	52	3	5543
4	611	4	71	4	7281
5	762	5	97	5	8215
6	879	6	146	6	9981
		7	212	7	10529
				8	12607

Year	Gray seals	Year	Muskox
1	751	1	49
2	854	2	57
3	869	3	65
4	898	4	61
5	1019	5	76

1.16.15 The following data are replicate monk seal beach counts from French Frigate Shoals. Conduct a test for significant deviations from regression using the “pure error” model of Section 1.6. There may be an advantage in using logarithms of the counts (to approximately “normalize” the data), as was done in Section 1.6, but try the analysis without the log transform. Report your results in an analysis of variance on a spreadsheet, as in Table 1.3.

1985	298	1990	264	1994	193
	250		271		183
	301		262		219
	403		300		190
1986	401		299		196
	285		300		198
	278	1991	176		202
1987	351		191		232
	285		216		222
	316		217		249
	301		197	1995	141
	320		185		124
	350		281		168
	333		273		132
	252	1992	204		140
	362		202		144
1988	292		226		174
	303		227		156
	288		234		164
	286		271		
	315		231		
	327	1993	156		

	327	195
	354	186
1989	331	182
	337	189
	322	221
	313	161
	279	184
	292	187
	319	208
	354	194
	375	219
	363	

1.16.16 The following data are from three years of a survey of harbor porpoises in which there were replicate transects flown and the transect lengths were recorded.

Year	Km.	Count	Km.	Count	Km.	Count		
1986	552	48	1987	326.6	1	1988	199.1	1
	318	31		117.5	12		66.5	12
	445	9		752	30		374.7	5
	399	59		384.4	24		685.7	71
	195	1		58.5	0		333.7	0
	150	10		223.2	6		311.9	18

1.16.17 Perform an analysis of covariance (Section 1.10) on the data of Exercise 1.16.15, and report your results on a spreadsheet. Is this a legitimate analysis in view of the results of Exercise 1.16.15? Explain.

2.0 INTRODUCTION TO BOOTSTRAPPING

2.1 Introduction

There are a lot of valuable statistical methods that are practically guaranteed to work well if the data are approximately normally distributed and we are mainly concerned with linear functions of random variables. As was remarked in Chapter 1, the mean or average of a data set is a linear combination of random variables, and the central limit theorem says that we can expect means to converge on normality as the sample size increases. However, ecologists often are forced to use small samples. Very often we want to consider ratios of random variables, which are definitely nonlinear combinations, and difficult to deal with in any consistent manner. Many models of importance in ecological studies contain products, ratios and exponents, and are simply not susceptible to a standard statistical analysis in terms of the available theory.

A relatively new development in statistical methodology offers a way out of this dilemma. The technique is called "bootstrapping", which, according to Efron and Tibishirani (1993) was named from the phrase "to pull oneself up by one's bootstraps", i.e., to accomplish a physical impossibility. Efron and Tibishirani (1993:56) note that the bootstrap was introduced by Efron in 1979, making it quite a recent development in contrast to many other statistical techniques. It was preceded by "jackknifing" which was originated by Quenouille (1956) as a way to study bias in estimators, but named by John Tukey (1958) due to its all-purpose applicability, like one's handy jackknife. A related topic is the use of the "delta method" to estimate variances for estimates based on complicated models. We will touch on these latter two methods later, but will mainly depend on bootstrapping as the principal tool for handling difficult problems.

One of the nice things about bootstrapping is that it is simple to apply, so long as one has access to a computer. Detailed application requires access to a desk computer and some knowledge of a programming language. However, bootstrapping can be done in EXCEL, as used here. There are several programming languages that can be used for bootstrapping. The BASIC language is simple and easy to use, particularly because it is "interpreted", i.e., one can issue the RUN command at almost any stage and find out whether the code written up to that point works as expected. QUICK BASIC contains a built-in compiler, so a large number of runs can be made quite rapidly. QUICK BASIC runs using almost identical statements on both MS/DOS and Macintosh operating systems, and has been issued with the Microsoft operating systems, but unfortunately in the MS/DOS version which makes it awkward to use. Most of the examples given here were also done in EXCEL, which has a random number generator in the statement RANDBETWEEN(N1,N2) where N1 and N2 represent the range of the random numbers to be generated. Be sure the ANALYSIS TOOLPAK is loaded before attempting the EXCEL versions of bootstrapping. Pull down the Tools menu and use the add-ins element to find the Analysis Toolpak. Details of use vary with the version, so you may need to use the "help" function on occasion.

2.2 The mechanics of bootstrapping

Bootstrapping is easy to apply. The process for approximating the standard error of a mean is illustrated in Fig. 2.1. An original data set containing n items (here $n = 10$) is randomly sampled with replacement B times. Four of these B samples are shown above the original data set in Fig. 2.1. Note that an individual value from the original data set, such as 106 may appear repeatedly in a bootstrap sample. Each of the B bootstrap samples is averaged, as shown above the individual samples. This is the bootstrap replication. We then use these B replicate values to compute the standard error of the mean. The equation is

exactly the same as that for calculating a variance, namely $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$. However, Efron and Tibishirani

use a different notation to distinguish bootstrap variables from the original data, using $x^{*1}, x^{*2}, \dots, x^{*B}$ to denote the vectors containing the bootstrap samples of n observations (i.e., the four sets of bootstrap samples of 10 items each shown in Fig. 2.1). Thus the first bootstrap sample is

$$x^{*1} = (203, 203, 106, 106, 106, 160, 106, 8, 301, 160).$$

The original data is represented by the vector $x = (x_1, x_2, \dots, x_n)$.

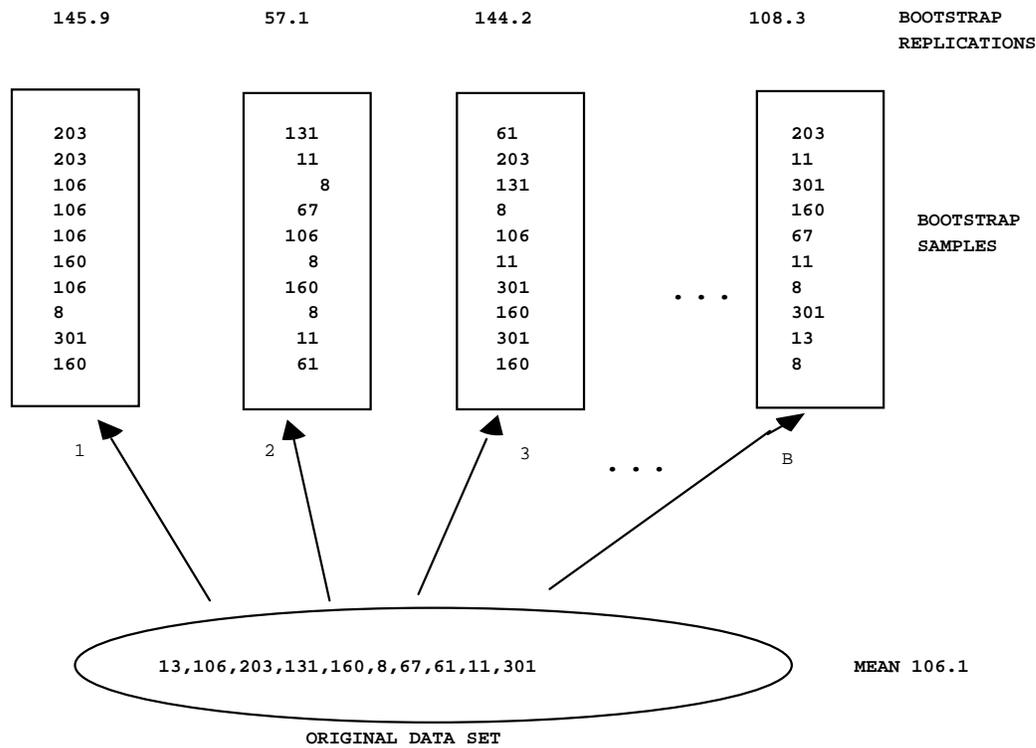


Fig. 2.1 The bootstrapping scheme for estimating a standard error. An original data set containing n items is randomly sampled B times with replacement using samples of size n . Each such bootstrap sample is averaged, and these means are used to estimate the standard error of the mean of the original data set.

The quantity $s(x^{*1})$ denotes a statistic computed from the corresponding bootstrap sample. In this case $s(x^{*1})$ is the average of the first bootstrap sample, 145.9. Using the bootstrap notation the standard error of the mean estimated by bootstrap sampling is written as:

$$\hat{s}e_{boot} = \left\{ \frac{\sum [s(x^{*b}) - s(\cdot)]^2}{B-1} \right\}^{1/2} \quad (2.1)$$

where the summation runs from $b = 1$ to B , and $s(\cdot)$ represents the mean of the bootstrap sample, i.e., $\sum x^{*b}/B$, where again the summation runs from $b = 1$ to B . The important thing to remember here is that $s(x^{*b})$ represents the mean of the b^{th} bootstrap sample, so that $s(\cdot)$ is the average of B such averages. Note, too, that the standard error of a set of random variables is computed as $s/n^{1/2}$, but here we are computing the standard deviation of a set of means and this is the standard error of the mean (i.e., don't make the mistake of dividing by the square root of B).

The first few columns of an EXCEL worksheet used to bootstrap the data of Fig 2.1 follow. The first row shows the assignment of a serial number to the original data items, while the original data appear in the second row. The next 10 rows list random numbers from 1 to 10 obtained from the statement `RANDBETWEEN(1,10)`. The next set of numbers are random samples, with replacement, from the original data set.

ITEM NUMBER	1	2	3
DATA	13	106	203
	1	6	8
	2	9	7
			2
			6

RANDOM NUMBERS	3	1	3	10	
	4	8	9	9	
	5	2	2	1	
	6	5	6	1	
	7	3	8	9	
	8	1	2	6	
	9	7	2	2	
	10	3	2	8	
	BOOTSTRAP SAMPLES	1	8	61	106
		2	11	67	8
3		13	203	301	
4		61	11	11	
5		106	106	13	
6		160	8	13	
7		203	61	11	
8		13	106	8	
9		67	106	106	
10		203	106	61	
SUM		845	835	638	
MEAN		84.5	83.5	63.8	

These are obtained by using a "table lookup" function in EXCEL. It can be explained by referring to the following formulas for the first column.

ITEM NUMBER	1		
DATA	13		
	1	=RANDBETWEEN(1,10)	
	2	=RANDBETWEEN(1,10)	
RANDOM NUMBERS	3	=RANDBETWEEN(1,10)	
	4	=RANDBETWEEN(1,10)	
	5	=RANDBETWEEN(1,10)	
	6	=RANDBETWEEN(1,10)	
	7	=RANDBETWEEN(1,10)	
	8	=RANDBETWEEN(1,10)	
	9	=RANDBETWEEN(1,10)	
	10	=RANDBETWEEN(1,10)	
		1	=HLOOKUP(C3,\$C\$1:\$L\$2,2,FALSE)
	BOOTSTRAP SAMPLES	2	=HLOOKUP(C4,\$C\$1:\$L\$2,2,FALSE)
3		=HLOOKUP(C5,\$C\$1:\$L\$2,2,FALSE)	
4		=HLOOKUP(C6,\$C\$1:\$L\$2,2,FALSE)	
5		=HLOOKUP(C7,\$C\$1:\$L\$2,2,FALSE)	
6		=HLOOKUP(C8,\$C\$1:\$L\$2,2,FALSE)	
7		=HLOOKUP(C9,\$C\$1:\$L\$2,2,FALSE)	
8		=HLOOKUP(C10,\$C\$1:\$L\$2,2,FALSE)	
9		=HLOOKUP(C11,\$C\$1:\$L\$2,2,FALSE)	
10		=HLOOKUP(C12,\$C\$1:\$L\$2,2,FALSE)	

SUM =SUM(C14:C23)
 MEAN =C25/10

The statement HLOOKUP(C3,\$C\$1:\$L\$,2,FALSE) specifies a horizontal lookup table (VLOOKUP permits a vertical lookup table). The first entry is the column entry for the value to be looked up in the table, i.e., C3 denotes a random number entry, for which we need to find the corresponding entry in the original data row. The lookup table is specified by the array, \$C\$1:\$L\$1 in which the first row is the index value corresponding to a data entry in the next row. The subsequent value in HLOOKUP is the row containing the data to be returned by the HLOOKUP function, and the final entry ("FALSE") insures that the function returns the exact value required (using "TRUE" would permit returning the value nearest in numerical magnitude to a lookup entry). As with any of the more complex functions in EXCEL, a little practice will make the role of the individual entries clear. An important proviso with the HLOOKUP and VLOOKUP functions is that the lookup table must be in the first rows (or first columns for VLOOKUP) of the spreadsheet. The last entries above give the sums and means of the bootstrap samples. The means are used in eq.(2.1) to calculate the bootstrap standard error. Readers should understand that the example used here is mainly intended to demonstrate the approach. The best estimate of a standard error of a set of numbers is that calculated by the usual formula, i.e., from

$$\text{S.E.}^2 = \frac{\sum(x_i - \bar{x})^2}{n(n-1)} .$$

Bootstrapping is used to calculate standard errors for more complex functions, for which a direct estimate of a variance is not available from statistical theory.

Example 2.1. The original data of Fig. 2.1: 13,106,203,131,160,8,67,61,11,301 represent data on survival times in days. They can be considered to come from an experiment on the effect of some treatment on survival of experimental animals (whereupon there should be a corresponding set of data from a control group) or the survival times of a set of radio-tagged wild animals. It is a small sample, but this is common, inasmuch as there is increasing public pressure to reduce experimental use of live animals, and collecting data from wild animals is expensive and can be quite difficult. We would thus like to extract as much information as possible from the data. The data in Table 2.1 are from an EXCEL worksheet that computes the bootstrap standard error. It shows the first 10 columns of a total of 50, which is likely the minimum size that should be used to demonstrate behavior of bootstrapping. In preparing such spreadsheets, one should change the calculation mode from automatic to manual (in the TOOLS menu, under OPTIONS or PREFERENCES depending on the version of EXCEL) while building the worksheet. The calculate command can then be used to see how the result varies from run to run.

2.3 Empirical probability distributions

The probability distribution of a random variable, X, is any complete description of the probabilistic behavior of x. In coin-tossing with a "fair" coin, there are two possibilities, each occurring with probability 1/2. In rolling a die, there are 6 outcomes, each having $\text{Pr}\{x=k\} = 1/6$ for $k = 1,2,3,4,5, \text{ or } 6$. It is convenient to define the sample space, S_X , as a list of possible outcomes. Thus for a fair die, $S_X = \{1,2,3,4,5,6\}$ and we assign probability 1/6 to each event in the sample space. Consider the binomial distribution which assigns a probability to each sample point in the sample space $\{0,1,2,3, \dots, k, \dots, n\}$ but these probabilities depend on the parameter, p, of the distribution. The binomial distribution is:

$$\text{Prob}\{x_i = k\} = f_k = \binom{n}{k} p^k (1-p)^{n-k} \quad (2.2)$$

where $\binom{n}{k}$ is evaluated as $\frac{n!}{(n-k)!k!}$, in which, for example, 5! (read as "five factorial") is calculated as $5 \times 4 \times 3 \times 2 \times 1 = 120$.

Table 2.1 Sample from bootstrapping the data of Fig. 2.1.

ITEM NUMBER		1	2	3	4	5	6	7	8	9	10
DATA		13	106	203	131	160	8	67	61	11	301
	1	7	8	1	6	5	4	6	2	8	6
	2	10	10	3	8	10	10	2	4	3	2
RANDOM	3	2	7	4	10	7	4	3	5	10	10
NUMBERS	4	4	1	4	4	5	6	8	10	3	7
	5	8	6	1	6	7	8	5	3	1	4
	6	9	7	9	4	4	9	1	7	6	6
	7	1	5	6	9	10	3	5	3	9	7
	8	1	8	3	2	10	10	10	1	1	7
	9	5	2	6	6	9	10	1	10	10	6
	10	2	6	3	8	7	3	2	2	2	6
	1	67	61	13	8	160	131	8	106	61	8
BOOTSTRAP	2	301	301	203	61	301	301	106	131	203	106
SAMPLES	3	106	67	131	301	67	131	203	160	301	301
	4	131	13	131	131	160	8	61	301	203	67
	5	61	8	13	8	67	61	160	203	13	131
	6	11	67	11	131	131	11	13	67	8	8
	7	13	160	8	11	301	203	160	203	11	67
	8	13	61	203	106	301	301	301	13	13	67
	9	160	106	8	8	11	301	13	301	301	8
	10	106	8	203	61	67	203	106	106	106	8
SUM		969	852	924	826	1566	1651	1131	1591	1220	771
MEAN		96.9	85.2	92.4	82.6	156.6	165.1	113.1	159.1	122	77.1

For convenience in discussing bootstrapping, we can describe a probability distribution as $F\{f_1, f_2, f_3, \dots, f_k, \dots, f_N\}$ where f_i is the limiting frequency of the i^{th} event. For a single die, we infer that $f_i = 1/6$, and would expect to eventually come very close to that value, given enough rolls of the die. If we determine f_i from observations, then it can be considered to be an empirical probability distribution. Instead of rolling dice, we can set up a spreadsheet using `RANDBETWEEN(1,6)`, copy this down through, say, 1,000 cells, and tabulate the outcomes by using the histogram function in the data analysis menu under `TOOLS`. This gave the following results:

<i>Bin</i>	<i>Frequency</i>	<i>Proportion</i>	<i>F(x)</i>
1	181	0.1810	0.1810
2	179	0.1790	0.3600
3	162	0.1620	0.5220
4	172	0.1720	0.6940
5	152	0.1520	0.8460
6	154	0.1540	1.0000

TOTAL	1000
-------	------

The column under proportion gives the empirical frequency distribution, with the corresponding cumulative frequency distribution being shown under $F(x)$.

It is necessary to note that, in mathematical statistics, $F(x)$ represents the cumulative probability distribution function, $F(x_0) = \Pr\{x \leq x_0\}$. The table above provides an example. We will use the description $\hat{F}\{f_1, f_2, f_3, \dots, f_k \dots, f_N\}$ for a finite number of events as a handy way to represent an empirical probability distribution in discussing bootstrapping. The "hat", $\hat{}$, over a symbol means that the quantity is an estimate of the true, but unknown, value, F . The cumulative, $F(x_0) = \Pr\{x \leq x_0\}$, will mainly be used here in simulations. It is important to remember that the sum of the frequencies in $F\{f_1, f_2, f_3, \dots, f_k \dots, f_N\}$ is always unity, described as $\Pr\{x \in S\} = \sum f_k = 1$, where " \in " means "contained in".

2.4 Sample sizes for bootstrapping

Bootstrapping is a resampling procedure, that is, we take repeated samples of the original data set, calculate values of some statistic $s(x^*)$ and use these to infer something about the true, but unknown, value of some parameter. In the example used thus far, the statistic was the standard error of the mean. How many bootstrap replications are needed? Efron and Tibishirani (1993: Eq.(6.9)) give a formula for examining the effect of varying sample size, but also indicate that, in their experience, $B = 200$ is usually adequate for estimating the standard error, while $B = 50$ may provide useful information. In the problems I have dealt with thus far by using bootstrapping, I tend to use $B = 100$ for exploring data and debugging programs, and $B = 1,000$ or $2,000$ for the published result. With desktop computing so cheap, one might as well resort to "overkill" unless the statistic being bootstrapped is very complicated and requires a lot of computing time. However, this choice of $B \geq 1,000$ is also largely driven by the fact that larger samples are needed to compute confidence limits on a mean by bootstrapping, as we'll see in the next section. When making calculations using EXCEL one can only get about 250 bootstraps in the horizontal plane, so it is necessary to use VLOOKUP and set up the table in the vertical plane, whereupon it is possible to get 2,000 bootstraps for confidence limits. If more bootstraps are needed, one can copy off data to another sheet and recalculate, copy off those results, and recalculate again. Large samples can thus be obtained. However, as noted above, 2,000 is usually adequate for confidence limits.

2.5 Percentile confidence limits

Calculating confidence limits by bootstrapping can be extremely simple, if the percentile method is used. Follow the same process demonstrated in Fig. 2.1, generating at least 1,000 bootstraps (I tend to use 2,000 if computing doesn't take too long), store the data in a file, arrange it in numerical order, and count in $\alpha B/2$ observations from both ends, where α is the chosen "significance level". These are the percentile confidence limits. Although there is nothing in the underlying theory that dictates a choice, most biologists tend to use $\alpha = 0.05$, for 95% confidence limits.

To accomplish this in EXCEL, use VLOOKUP and put the data in the first 2 columns, the same number of random numbers in the third row as there are data points, and the corresponding lookup values in the fourth row. Calculate the function being bootstrapped in subsequent rows. You can then order the data and count in from both ends for confidence limits. In this case, the function being bootstrapped is simply the mean. Confidence limits are obtained by ordering this column (use SORT in the DATA menu of EXCEL) and counting in from the ends of the ordered data. To use $\alpha = 0.05$ on 1,000 bootstraps, one would count in 25 observations from each end. In this example, the approximate 95% confidence limits were 55.4 to 169 around the mean of 106.1 of the original data, based on 1,000 bootstraps. Using a BASIC program to do the bootstrapping is faster and requires less effort once the programming is done. Results of 2,000 bootstraps from a BASIC program (Fig. 2.2) gave confidence limits of about 55-164.

As noted previously, EXCEL will accommodate at least 2,000 bootstraps in the vertical arrangement. However, if an older (and thus slower) computer is used, it may be best to do only 200 bootstraps at a time. That is, set up the operation as shown on the attached sheet, run 200 bootstraps, and copy the results to a second worksheet. Do this 5 times (or 10 if you want 2,000 bootstraps) and then order the data on the second worksheet to locate the confidence limits.

Students should review normal theory confidence limits in the statistics text of their choice at this point. Under normal theory, we would calculate a standard error of the mean of the original data of Fig. 2.1 (mean = 106.1), getting $s = 95.33$, and $S.E. = 95.33/(10^{1/2}) = 30.14$, and calculate 95% confidence limits of ± 1.96 SE. I tend to use 2 rather than 1.96 for convenience, and a little extra margin. Using 2 S.E. gives approximate 95% limits of 46 to 166. Survival data generally follow a highly skewed distribution, and the sample variance tends to vary appreciably. In this case, the limits are so wide that the data don't give us a very good notion of average survival time.

Statistics books recommend transforming skewed data in order to approximate normality. One then produces normal-theory confidence limits as above, and transforms back to the original scale. It can, however, be a considerable chore to find a normalizing transformation suitable for the data at hand. Further, the small sample of Example 2.1 simply does not supply enough information to evaluate possible transformations. It is thus reassuring that Efron and Tibishirani (1993: Chap. 13) indicate that the percentile method automatically supplies limits that would be obtained under normal theory if we knew the proper transformation to normalize the data.

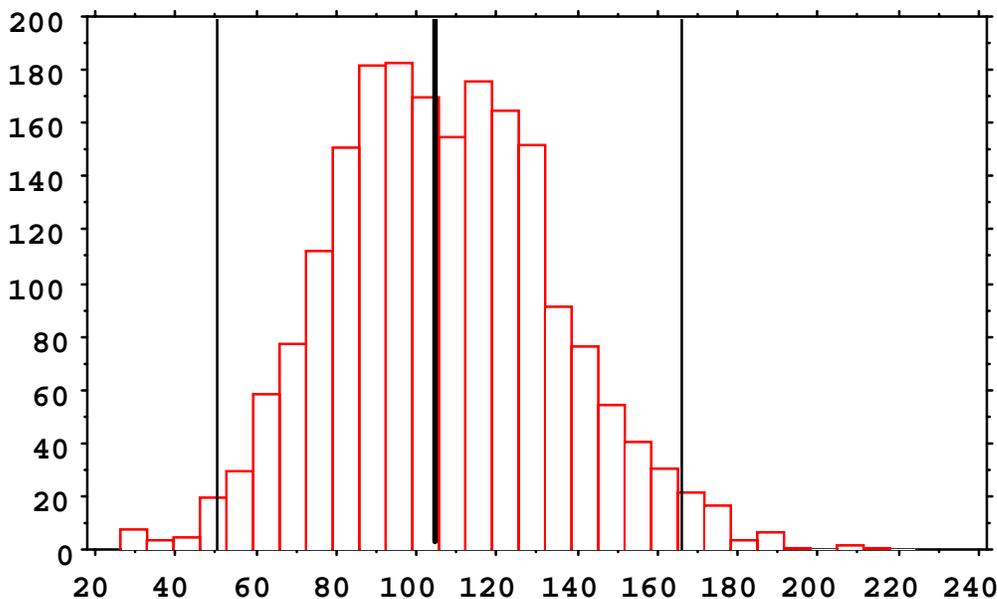


Fig. 2.2 Frequency plot of 2,000 bootstraps of the original data of Fig. 2.1, showing mean of the original data (heavy central vertical line), and 95% confidence limits (lighter lines to right and left) from the bootstrap percentile method for calculating confidence limits. These limits are at about 55 and 164 for the data shown

2.6 Regression models and parametric bootstrapping

Regression models provide extremely valuable tools in ecological studies. Many investigators use regressions without giving much thought to the matter, and may thus report some erroneous results without realizing that this is possible. Regression models are classified as linear and nonlinear. Linear models are most commonly used, with the main example being $y = \alpha + \beta x$, where α denotes the "intercept" and β the "slope". Ecologists also use multiple regression models with two or more x -values, e.g., $y = \alpha + \beta_1 x_1 + \beta_2 x_2$, and may also use multiple regression models like $y = \alpha + \beta_1 x_1 + \beta_2 x_2^2$. These are both linear

models, being "linear in the coefficients", but a version like $y = \alpha + \beta_1 x_1 + \beta_2 x_2^\gamma$ is nonlinear. A frequently encountered nonlinear model is $y = \alpha e^{-\beta x}$. This model can, however, be transformed into a linear model by taking logarithms (usually to base e) giving $\log_e y = \log \alpha - \beta x$. The model $y = \alpha + \beta_1 x_1 + \beta_2 x_2^\gamma$ is said to be intrinsically nonlinear, inasmuch as a simple transformation will not convert it to a linear version (unless, of course, we know or assume we know γ). Dealing with intrinsically nonlinear models can be difficult, and they are most often fitted with nonlinear least-squares. Programs are available for fitting by nonlinear least-squares.

Regression models may be bootstrapped in exactly the same way as shown in Fig. 2.1, except that now the original data will consist of x, y pairs, and the statistic computed from bootstrap replications consists of paired estimates of α and β , rather than the mean as used in the example of Fig. 2.1. How these sets of paired estimates are treated depends on the purpose of the study. Often the main interest is in estimates of β , but we may also want to set confidence limits on an estimate of some value of x computed from the estimates of α and β . Texts on regression analysis are available; one of the more widely used is that of Draper and Smith (1998), and most basic statistics texts give a good deal of attention to regression models. To set confidence limits on some regression estimate by bootstrapping, one simply needs to follow the procedure presented above, with the "statistic" being the estimate of interest in the study at hand.

The chief problem for ecologists in this approach is the usual one -- small sample sizes. With smallish samples, bootstrapping pairs may give some strange and variable results. We will thus need to consider parametric bootstrapping. The procedure again is simple. One fits a regression model to the original data, calculates residuals about the fitted line, and bootstraps the residuals. Consider the result of fitting a simple linear regression to n original pairs of x, y observations. The outcome is a fitted regression line, denoted as $\hat{y}_i = a + bx_i$, where a and b represent the estimates of α and β , and there are n pairs of original data. The residuals about regression are calculated as:

$$e_i = \hat{y}_i - y_i \quad (i = 1, 2, 3, \dots, n) \quad (2.2)$$

where \hat{y}_i is calculated from the fitted line, $\hat{y}_i = a + bx_i$. We now bootstrap the residuals, taking repeated random samples with replacement of n observations from the residuals, add these residuals to the fitted regression line to get a new set of n values of y_i . Combined with the original set of x -values (unchanged throughout) these new pairs constitute the bootstrap samples of Fig. 2.1. We then calculate the bootstrap replication by fitting a new regression line to the bootstrap sample. Of course, if we are only interested in, say, the slopes, b , then only that calculation needs to be carried out. The only tricky part is to remember that the new values of y_i are computed from the i^{th} value of x_i , so that the same residual (e_i) may be associated with several values of x_i , depending on the random selection. That is, the new set of y_i values is computed from:

$$y_i = a + bx_i + e_i \quad (i = 1, 2, 3, \dots, n) \quad (2.3)$$

with a and b coming from the regression line fitted to the original data and the values of e_i come from a random sample with replacement of the n data points generated by eq.(2.2). Students should repeat the calculations shown below to fix the scheme in mind.

Example 2.2 Parametric regression bootstrapping.

For simplicity, we will suppose that we want 95% confidence limits on the slope, β , of a regression line. The slope estimate is calculated as:

$$b = \hat{\beta} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \quad (i = 1, 2, 3, \dots, n) \quad (2.4)$$

and the intercept estimate is $a = \hat{\alpha} = \bar{y} - b\bar{x}$.

STEP 1 Compute a regression line from the original data:

$$b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{325.181239}{630.5} = 0.5157$$

$$a = \bar{y} - b\bar{x} = 16.725 - (0.5157)(20.5) = 6.1525$$

$$\hat{y} = 6.1525 + 0.5157x \quad \text{Regression line from original data}$$

STEP 2 Calculate the deviations $e_i = \hat{y}_i - y_i$

Original data					
i	x_i	y_i	\hat{y}_i	$e_i = \hat{y}_i - y_i$	
1	10	12.672	11.31	1.362	
2	12	8.9391	12.3415	-3.402	
3	14	13.934	13.373	0.561	
4	15	16.377	13.8888	2.488	
5	17	13.252	14.9203	-1.668	
6	21	19.121	16.9833	2.137	
7	23	17.821	18.0148	-0.194	
8	28	18.879	20.5936	-1.715	
9	30	21.047	21.6251	-0.578	
10	35	25.213	24.2038	1.009	

STEP 3 Draw random samples of 10 with replacement from the e_i :

i	e_i	Random samples with replacement from the e_i						
1	1.362	-0.1936	0.5607	-1.6681	-0.5776	-3.4024	1.0091	
2	-3.402	-0.5776	2.4879	-0.1936	-0.1936	1.3618	-3.4024	
3	0.561	-1.7150	-1.7150	1.0091	-3.4024	1.0091	0.5607	
4	2.488	0.5607	1.3618	2.4879	-0.1936	-3.4024	1.3618	
5	-1.668	1.3618	-0.1936	-1.7150	-0.1936	1.0091	2.1373	
6	2.137	-1.6681	-1.6681	1.3618	-0.5776	1.0091	1.3618	
7	-0.194	2.1373	-0.1936	0.5607	-1.6681	-3.4024	-3.4024	
8	-1.715	-0.1936	1.0091	2.4879	-1.7150	-3.4024	-3.4024	
9	-0.578	-1.7150	-0.5776	1.3618	1.3618	2.4879	0.5607	
10	1.009	1.3618	-1.6681	1.3618	1.0091	-1.7150	2.1373	

STEP 4 Add the random samples of e_i to the predicted regression line to obtain new sets of y_i :

i	x_i	\hat{y}_i	$\hat{y}_i +$ random samples with replacement from the e_i						
1	10	11.31	11.1164	11.8708	9.6419	10.7325	7.9076	12.3192	
2	12	12.342	11.7640	14.8294	12.1479	12.1479	13.7033	8.9391	
3	14	13.373	11.6580	11.6580	14.3822	9.9706	14.3822	13.9338	
4	15	13.889	14.4495	15.2506	16.3766	13.6952	10.4864	15.2506	
5	17	14.92	16.2821	14.7267	13.2053	14.7267	15.9294	17.0576	
6	21	16.983	15.3152	15.3152	18.3451	16.4057	17.9924	18.3451	
7	23	18.015	20.1521	17.8212	18.5755	16.3467	14.6124	14.6124	
8	28	20.594	20.3999	21.6027	23.0814	18.8786	17.1911	17.1911	
9	30	21.625	19.9101	21.0475	22.9868	22.9868	24.1129	22.1858	
10	35	24.204	25.5656	22.5357	25.5656	25.2130	22.4888	26.3411	

STEP 5 Calculate regression slopes for each of these new sets of "data", using the same set of x_i . This gives the values, 0.542, 0.443, 0.608, 0.582, 0.513, and 0.520. In practice, of

course, one would calculate a large set of such estimates, 1,000 or more. The frequency distribution of these values then provides the basis for confidence limits, as calculated previously for means.

A generalized summary of the steps in parametric bootstrapping is as follows:

- 1) Compute estimates of the parameters of the model from the original data. In this case, the regression coefficients, a and b .
- 2) Calculate deviations, $e_i = \hat{y}_i - y_i$, between the observed data (y_i) and the fitted model (\hat{y}_i).
- 3) Draw B (at least 1,000 for confidence limits) random samples of n with replacement from this set of deviations.
- 4) Add these deviations to the \hat{y}_i to create the bootstrap replications.
- 5) Compute parameter estimates from each of these B sets of data.
- 6) Arrange these B estimates in a frequency distribution and count in $\alpha B/2$ observations from each end to obtain $(1-\alpha)\%$ confidence limits.

Calculations can be carried out in EXCEL by using the same arrangement as used in Sec. 2.5 to get confidence limits on a mean. The data to be bootstrapped are now the deviations from regression, and the bootstrap operation proceeds in exactly the same manner. However, another stage has to be incorporated in which the bootstrapped deviations are added to the predicted regression. These new regression values are then used to estimate the parameters of the regression equation. In the present example, only the slope is calculated. This can be done by using the SLOPE function, which returns the slope of two arrays. The x -values are the original values, while the y -values are those in the body of the table. The 1,000 slope values were then ordered, and approximate 95% confidence limits obtained by counting up and down 25 entries. The limits obtained from the EXCEL calculation ($B = 1,000$) were 0.375 and 0.648. A calculation using a program written in BASIC were 0.377 to 0.652. A plot of the results of 2,000 bootstraps computed by the BASIC program appears in Fig.2.3.

Students should review normal theory regression calculations in Chap. 1.0 or in a statistics textbook. The variance about regression is calculated as follows:

$$s^2 = \frac{\sum(y_i - \hat{y}_i)^2}{n - 2} = \frac{\sum(y_i - (a + bx_i))^2}{n - 2} \quad (2.5)$$

An estimate of the variance of the regression coefficient is given by:

$$s_b^2 = \frac{s^2}{\sum(x_i - \bar{x})^2} \quad (2.6)$$

and this variance has the t -distribution with $n-2$ degrees of freedom under normal theory. For 95% confidence limits in the present example, we look up the 0.025 ($\alpha/2$) value of t with 8 degrees of freedom, finding it to be 2.306, and calculate:

$$\text{Upper 95\% confidence limit} = b + t(s_b) = 0.516 + 2.306(3.952/630.5)^{1/2} = 0.698,$$

and the analogous lower limit is 0.333. Note that these limits are somewhat wider than the 95% limits obtained by bootstrapping. A small sample of quite variable data is involved. It is always important to look at a plot of the data.

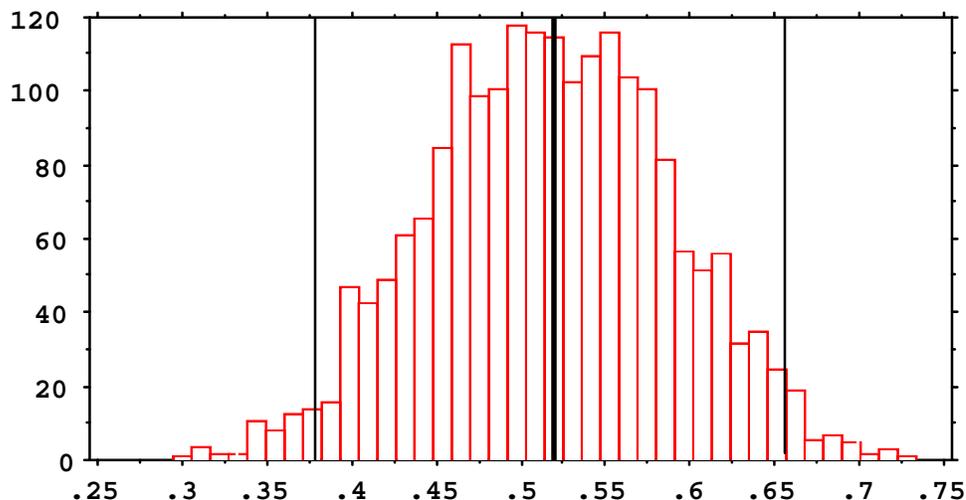


Fig. 2.3. Frequency plot of 2,000 bootstraps created by a BASIC program. The heavy central line shows the position of the regression slope calculated from the original data, while the lighter solid lines show the 95% bootstrap confidence limits. Broken lines show normal 95% confidence limits calculated from the observed data.

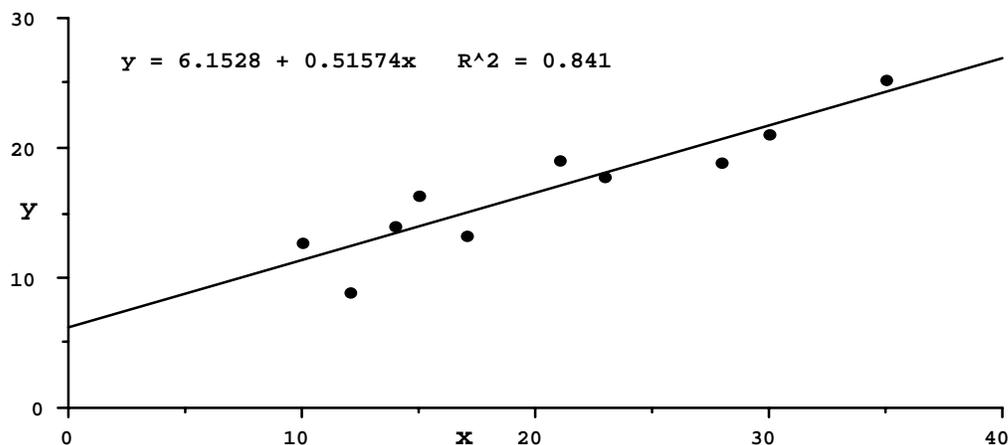


Fig. 2.4. Regression plot of the original data used in Example 2.2.

Most "canned" statistical programs also give the correlation coefficient, which is defined as:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{[\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2]^{1/2}} \quad (2.7)$$

In the present example, $r = 0.917$, and is reported by the graphics program that produced Fig. 2.3 as $r^2 = 0.841$. A very serious problem for ecologists is that much of the data they encounter is not normally distributed, and routine use of statistical packages without examining the assumptions or studying the data can lead to important errors in interpreting the data. Bootstrapping provides a way to examine the data without the normal theory assumptions, and thus helps to avoid blunders. The above set of regression data does conform to the normal theory model, so it is worthwhile to look at another example from a different source for contrast. The basis for claiming conformity to normal theory is that the data were constructed using normally distributed errors.

Example 2.3. A regression estimate of survival rate. As a further example, we consider a common use of regression methods. Many investigators are interested in estimating survival rates. Suppose we observe 100 marked animals over 10 years, and tally the number of survivors at the end of each year. If the probability of survival holds constant from year to year and animal to animal, then we can consider that the expected number surviving x years is just $E(n) = Np^x$, where N is the number originally marked and p is the probability of surviving a year. We might then use a model, $y_i = Ns^x$, where y_i is the number observed at the end of the x^{th} year and s is the survival rate. Taking logarithms gives:

$$\log y_i = \log N + x \log s \quad (2.8)$$

and an easy approach is just to fit a simple linear regression equation, $y = a + bx$, where $b = \log s$, and use e^b to estimate s . An example of such a data set follows:

Year	Survivors	Log survivors
1	89	4.48864
2	83	4.41884
3	74	4.30407
4	68	4.21951
5	65	4.17439
6	60	4.09435
7	55	4.00733
8	51	3.93183
9	48	3.87120
10	38	3.63759

Plotting log survivors against year gives the following graph:

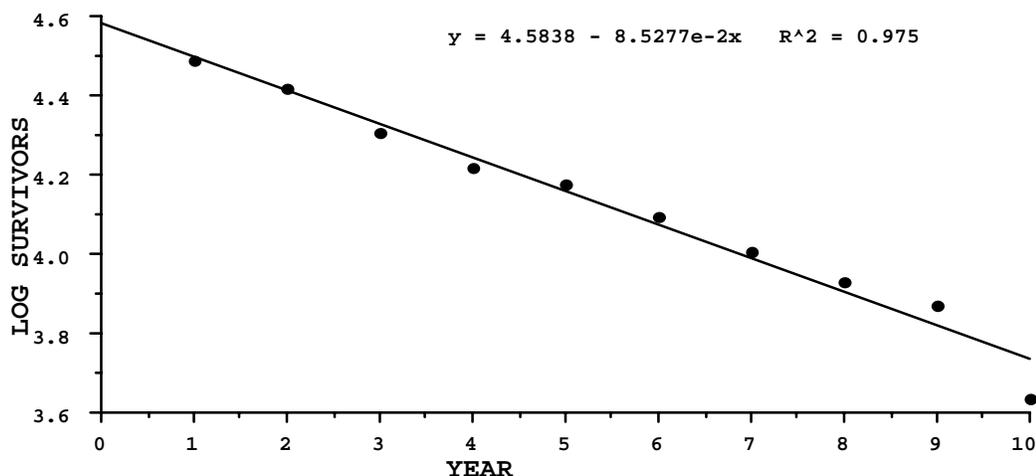


Fig. 2.5. Logarithms of the number of animals surviving at the end of each year regressed against time in years.

Using normal linear regression theory as in the previous example gives a slope of -0.08528 with 95% confidence limits of -0.0742 to -0.0963 , and translating these back to a survival rate and confidence limits gives

$e^{-0.08528} = 0.918$ with approximate 95% confidence limits of 0.908 to 0.928 . A run of 2,000 bootstraps gave the following frequency distribution, and 95% confidence limits of -0.09381 and -0.07717 , which translate to an annual survival range of 0.910 to 0.926 . The

bootstrapping was done with a BASIC program, but could have been conducted in EXCEL, just as in the previous example.

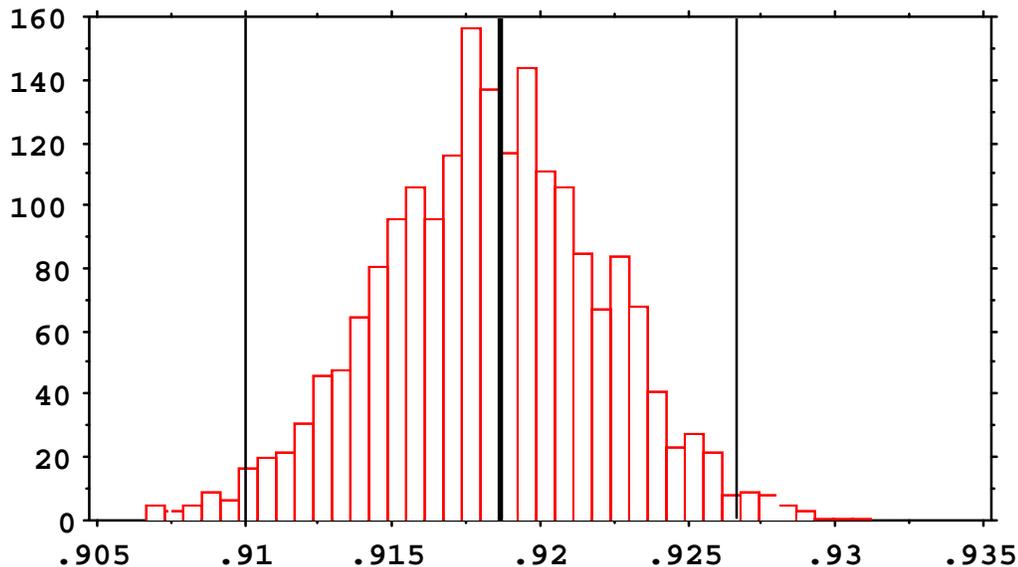


Fig. 2.6. Results of 2,000 parametric regression bootstraps for survival data. Bold central line shows the regression survival estimate (0.928) and solid lines the bootstrap 95% confidence limits. Broken lines are the 95% confidence limits obtained from normal regression theory.

The bootstrap limits appear to be a little "tighter" than the normal theory limits. How can we determine which method is right? One approach is to use "Monte Carlo" methods, which in this case amount to running many stochastic simulations of survival data, and determining which of the two choices for calculating confidence limits gives the best "coverage", i.e., do the calculated confidence limits include the true survival rate in 95% of simulated cases?

Example 2.4. The correlation coefficient, r . The correlation coefficient (r) calculated as in eq.(2.7) is widely used, along with the assumption that a transformation to:

$$z = 0.5 \log_e \frac{1+r}{1-r} \quad (2.9)$$

is normally distributed with expected value

$$\mu = 0.5 \log_e \frac{1+\rho}{1-\rho} \text{ and variance } \frac{1}{n-3}.$$

Approximate 95% confidence limits are obtained from $z \pm 2\left\{\frac{1}{n-3}\right\}^{1/2}$. Thus in Example 2.2, we had $r = 0.917$ which is transformed to:

$z = 0.5 \log_e \left\{ \frac{1.917}{.083} \right\} \pm 2\left[\frac{1}{7} \right]^{1/2}$ or $z_1 = 2.326$ and $z_2 = 0.814$. These confidence limits for the transformed variable are usually transformed back by iterative solutions of eq. (2.9), i.e., we find r_1 and r_2 from:

$$2.326 = 0.5 \log_e \frac{1+r}{1-r} \text{ and } 0.814 = 0.5 \log_e \frac{1+r}{1-r},$$

which gives 95% confidence limits on r as 0.67 to 0.98. If we resort to bootstrapping, then the 2000 bootstraps used to produce Fig. 2.3 (values of r were computed at the same time that values of b were calculated) gave approximate 95% confidence limits of 0.86 to 0.98, essentially the same upper limit, but an appreciably higher lower limit. A graph of the results follows:

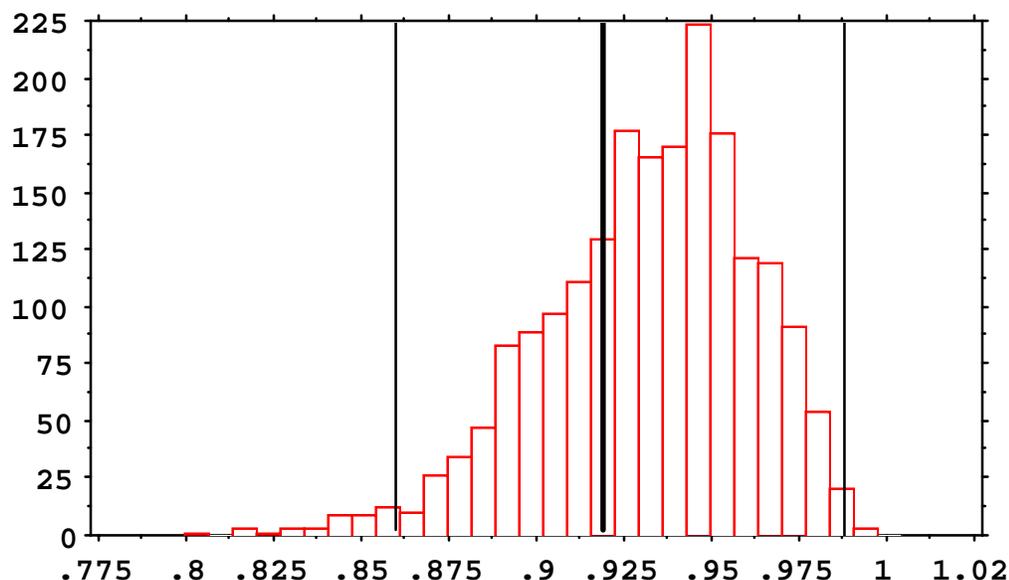


Fig. 2.7. Results of 2,000 bootstraps for the correlation coefficient of the regression data of Example 2.2. Heavy line shows correlation coefficient calculated from the original data, while lighter lines are approximate 95% confidence limits.

Clearly, the bootstrapped values of the correlation coefficient are quite skewed, but this is the situation with respect to normal regression theory also; the correlation of two jointly normal distributions has a skewed distribution (unless $\rho = 0$).

2.10 EXERCISES

Where bootstraps of 1,000 or more are involved, students should do the work in individual spreadsheets unless their computer has a large memory. Otherwise you are likely to get an “out of memory” notice when you try to copy from one workbook to another, etc. Make two worksheets for each such exercise, one summarizing results and the second containing the calculations. The practical approach is to save only the first 20 lines or so, when you have finished an exercise. You then can likely consolidate all results in one workbook to hand in. It is important to have your exercises in a workbook, as that makes it possible to try to find out where you went wrong if necessary. **IF YOU WANT TO LEARN TO BOOTSTRAP IN EXCEL, IT IS ESSENTIAL TO DO THE EXERCISES!** The exercises are more or less interlocking so you will need to do most of them. If you do, you should have a pretty fair notion of how to bootstrap. The bootstrapping technique will be used for examples and exercises in the rest of the book, so you need to know how to do it. If you know a programming language, you can certainly do the exercises that way, and provide summary tables and graphs to turn in. If you have not used the “graph wizard” function before, you may have trouble getting appropriate x-values on the graph. The trick is to first make an “xy (scatterplot)” graph, finish it and then open the CHART menu and select the bar chart. This changes the xy plot to a bar chart with the appropriate x-axis labels.

2.10.1 Set up an EXCEL worksheet to carry out bootstrap calculations on the data of Fig. 2.1, following the approach outlined in Table 2.1. Use 200 bootstraps. Set up the worksheet to use manual calculation as indicated in Example 2.1 and make 30 runs, recording the mean of the 200 bootstrap means in a separate column (you need to either type in the observed values as you repeatedly run the bootstrapping or use the “PASTE SPECIAL” command). Also calculate the variance of each group of 10 bootstrap samples, listing it at the bottom of the set along with the sum and mean of each set of 10. Record your results on a spreadsheet and save it for the next exercise. Calculate s.e.(boot) of eq. (2.1).

The explanation of using HLOOKUP in EXCEL manuals may not be very helpful. The sample below might help. This is part of a worksheet set up as indicated above and the HLOOKUP string displayed in the header

of the worksheet is as follows, referring to the bootstrap sample in the box in the body of the table. It commands EXCEL to lookup the random number in cell D3 (which is 3) in the table of the first two rows (designated in the command as the array \$D1:\$M\$2 and shown in boldface type below) and find the corresponding item in the second row of the array table (which is 203).

=HLOOKUP(D3,\$D\$1:\$M\$2,2,FALSE)

A	B	C	D	E	E	G	H	I	J	K	L	M				
EXERCISE				1	2	3	4	5	6	7	8	9	10			
2.10.1	ORIGINAL DATA			13	106	203	131	160	8	67	61	11	301			
			1	3	9	8	3	7	4	9	1	1	5	10	5	5
			2	2	4	5	2	6	3	9	1	8	7	6	6	9
			3	7	4	1	6	3	2	3	1	3	6	6	1	1
			4	5	8	3	5	9	8	10	7	4	7	10	10	8
	RANDOM NOS.		5	5	2	3	10	10	4	3	8	9	4	5	9	5
			6	8	3	9	4	5	10	4	3	10	6	9	3	8
			7	8	2	8	8	7	9	1	10	6	10	9	8	3
			8	2	5	1	10	5	5	9	2	1	4	7	1	3
			9	7	8	10	3	6	5	2	3	5	10	4	8	5
			10	6	9	2	1	5	6	6	7	6	4	1	8	8
			1	203	11	61	203	67	131	11	13	13	160	301	160	160
			2	106	131	160	106	8	203	11	13	61	67	8	8	11
			3	67	131	13	8	203	106	203	13	203	8	8	13	13
	BOOTSTRAP		4	160	61	203	160	11	61	301	67	131	67	301	301	61
	SAMPLES		5	160	106	203	301	301	131	203	61	11	131	160	11	160
			6	61	203	11	131	160	301	131	203	301	8	11	203	61
			7	61	106	61	61	67	11	13	301	8	301	11	61	203
			8	106	160	13	301	160	160	11	106	13	131	67	13	203
			9	67	61	301	203	8	160	106	203	160	301	131	61	160
			10	8	11	106	13	160	8	8	67	8	131	13	61	61
	BOOTSTRAP			1	2	3	4	5	6	7	8	9	10	11	12	13
			SUM	999	981	1132	1487	1145	1272	998	1047	909	1305	1011	892	1093
			MEA	99.9	98.1	113	149	115	127	99.8	105	90.9	131	101	89.2	109
			N													

The second item in the column of bootstrap samples (just below the item in a box) has the following command:

=HLOOKUP(D4,\$D\$1:\$M\$2,2,FALSE)

which instructs EXCEL to find the random number in the position D4 (remember that the first row of the table above, with letters A, B, C etc. is NOT part of a worksheet but merely gives locations in that worksheet). This random number is 2 and thus EXCEL picks out the second item in the array which is 106. The third command is as follows, and EXCEL uses the random number in D5 to pick out the 7th item in the array which is 67.

=HLOOKUP(D5,\$D\$1:\$M\$2,2,FALSE)

You may need to exert considerable patience and some trial and error efforts to get EXCEL to do the job if you have not worked much with it before, but once you have the hang of it, things should go along o.k.

2.10.2 Copy the original data and the run of 30 means obtained above to another spreadsheet, and compute means and variances for the two sets of data (the original data, 10 observations and the 30 means) using the built-in functions, i.e., AVERAGE() and VAR(). Make histograms of these means and variances from a run of the spreadsheet made in Exercise 2.10.1 (that is, make histograms of the 200 bootstrap means and

variances listed at the bottom of that sheet). Show how se_{boot} of eq.(2.1) compares with the variance of the original 10 observations and the variance of the 30 means. Is there any advantage to using bootstrapping in this example?

2.10.3 Repeat the die-tossing example of Section 2.3. Calculate the expected values. Explain the difference between a p.d.f. and an empirical probability distribution. Which is which in this example? What is the difference between a cumulative distribution function and a distribution function? State the distribution function for this example (as an equation).

2.10.4 Carry out the parametric bootstrap calculations for Example 2.3 using an EXCEL spreadsheet (the approach is given in Example 2.2). Use 2,000 bootstraps, and use VLOOKUP(). It is easier to use for larger numbers of bootstraps. Calculate confidence limits and prepare a graph of the frequency distribution to compare with Fig. 2.7. Use HISTOGRAM to obtain the frequency distribution. Be sure to “freeze” the appropriate cell references so that the x-values remain the same in calculating the bootstraps. You can obtain confidence limits simply by ordering the slopes using the SORT function.

2.10.5 Bootstrap the data of Example 2.2, using 1,000 bootstraps and computing the correlation coefficients rather than the slopes. You can start by copying the bootstrap calculation of Exercise 2.10.4 to a new spreadsheet and inserting the x and y values in this sheet. One can often convert a bootstrap operation to a new data set this way, so it is wise to keep examples on hand. Make a frequency distribution of z (eq.(2.9)). Does this look like a normal distribution as assumed in calculating confidence limits under the usual theory?

2.10.6 How would you obtain bootstrap confidence limits on α in Example 2.3? Calculate the 95% bootstrap confidence limits using 1000 bootstraps. Run the EXCEL regression on the data and compare the confidence limits on α with those you obtained from bootstrapping.

2.10.7 The regression bootstrap of Example 2.3 used parametric bootstrapping in which deviations from a model fitted to the original data are bootstrapped. The first example of bootstrapping given (Example 2.1) might thus be called “nonparametric” bootstrapping. Try this approach on the data of Example 2.3. Remember that you need to bootstrap pairs of observations. This may require setting up the slope calculations in blocks of 10, but careful use of the \$function will facilitate copying down in blocks of 10 without too much trouble. Use $B = 500$, and make a frequency plot of the calculated slopes and compare it with the frequency diagram of Exercise 2.10.4. . This should illustrate why parametric bootstrapping is preferred for small samples in regression studies.

2.10.8 Referring to the data of Example 2.3, calculate bootstrap 95% confidence limits on the variance about regression as shown in eq. (2.5). Compare your results with the value you get from a regression calculation on the original data of Example 2.3. Report your results on a worksheet (along the lines of those used thus far in the exercises above). You can use the results of Exercise 2.10.6 as a starting point adding on columns containing the sum-of-squares calculation and adding these up to get a variance estimate.

3.0 FURTHER TOPICS IN BOOTSTRAPPING

3.1 Introduction

There is much more to bootstrapping than can be covered here. The book by Efron and Tibishirani (1993) provides a detailed reference. Some useful further features are given here, and additional examples appear in subsequent chapters.

3.2 Estimates of bias

Bootstrapping provides a handy way to check for bias in an estimate. Suppose we calculate some statistic, such as the variance, from a set of data. We would like that statistic to provide an unbiased estimate of the true parameter, σ^2 . If we represent the statistic that we intend to calculate from the data as $\hat{\theta} = s(\mathbf{x})$, where \mathbf{x} is the vector of observed data, x_1, x_2, \dots, x_n and $s(\mathbf{x})$ is some function of that data, then we would like to have the expected value of that function to be $E(\hat{\theta}) = E[s(\mathbf{x})] = \theta = \sigma^2$. When the theoretical distribution of the estimator is known or assumed to be known, then an unbiased estimator can often be found by the methods of mathematical statistics, i.e., we can find the expected value of a trial statistic directly. Thus we know that the sample variance:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} \quad \text{has expected value } E(s^2) = \sigma^2 \text{ for the normal distribution of eq. (1.3). In}$$

this case, $\hat{\theta} = s(\mathbf{x}) = s^2$ (the notation can be a little confusing, as we use the notation $s(\mathbf{x})$ to represent any statistic calculated from a data set, \mathbf{x} , whereas s is also commonly used to represent a specific quantity, the sample estimate of the standard deviation).

Very often we are not sure what theoretical distribution may be appropriate for an observed sample, and it is frequently true that there may not be any such distribution. Statisticians thus spend a lot of time trying to choose the "right" distribution or manipulating (transforming) the data to approximate some known distribution. Bootstrapping can avoid a lot of that trouble and uncertainty. In this section we consider how bootstrapping can be used to check for bias in an estimator. We define bias as:

$$\text{bias}_F = E_F[s(\mathbf{x})] - t(F) \quad (3.1)$$

where the subscript F serves as a reminder that the bias and expectation are taken with respect to some probability distribution function F (quite likely an unknown distribution), and $\theta = t(F)$ denotes our statistic as calculated from the true probability distribution. The bootstrap estimate of bias is calculated as:

$$\hat{\text{bias}}_B = \hat{\theta}^*(\cdot) - t(\hat{F}) \quad (3.2)$$

where $\hat{\theta}^*(\cdot)$ is the mean of our estimator calculated from a large number of bootstrap samples, and $t(\hat{F})$ is the same estimator calculated from the original data.

Example 3.1. A numerical example from a normal distribution may help fix ideas at this point. The following 10 "observations" are from a normal distribution with mean zero:

1.6718, -3.061, 0.9338, 2.8766, -1.248, 2.6206, 0.3212, -1.121, 0.0475, 1.7129

Consider estimating the variance of these observations from:

$$s^2 = \sum \frac{(x_i - \bar{x})^2}{n} \quad (3.3)$$

This is the equation for variance often used by engineers, and produced by some of the earlier pocket calculators. If we apply the formula to the 10 original observations it gives $t(\hat{F}) = 3.177$. Now calculate the average of 2,000 bootstraps using this estimator. This gives $\hat{\theta}^*(\cdot) = 2.799$, and we get:

$$\hat{\text{bias}}_B = \hat{\theta}^*(\cdot) - t(\hat{F}) = 2.799 - 3.177 = -0.378,$$

a negative bias, suggesting that we have an underestimate. From theory, we know that an unbiased estimator comes from

$$s^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$$

which gives us $s^2 = 3.530$ from the above data set. Our bias estimate is negative, meaning that we underestimate the true quantity. We could thus add this quantity, 0.378, to our underestimate from the original data, $t(\hat{F}) = 3.177$, getting an improved estimate (3.555). It is perhaps better to define a bias-corrected estimator :

$$\hat{\theta} = \hat{\theta} + [\hat{\theta} - \hat{\theta}^*(\cdot)] = 2\hat{\theta} - \hat{\theta}^*(\cdot) = 2(3.177) - 2.799 = 3.555 \quad (3.4)$$

which gives the same result. This is close to the result (3.530) one

would get by using the proper equation $s^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$ in the first

place. The point here is that we often don't know what the proper equation is, and the bootstrap provides a way to check for bias in whatever equation we do have available to estimate some quantity. Eq. (3.4) came very close to the correct answer in this example, but in practice, if we have indications of an important bias, simply correcting by eq. (3.4) may not necessarily improve the situation. The estimator may be subject to a great deal of variability, so that the adjustment may not help. The essential conclusion here, is that if the bootstrap indicates small bias and small standard error, then we can be very comfortable indeed with our estimator, even if we don't have a theoretical model. Note that this result came out very close to the expected answer just by chance; repeating it gives a smaller bias, as will likely be evident in Exercise 3.11.1.

3.3 An improved bias estimate

Efron and Tibishirani (1993:Ch. 10) recommend an improved bias estimate that converges on the asymptotic estimate with a smaller bootstrap sample, B . They define a resampling vector for each bootstrap sample that contains the proportions of that bootstrap sample calculated from the frequencies with which the individual entries are observed. Thus in Fig. 2.1, the original data set was:

1	2	3	4	5	6	7	8	9	10
13	106	203	131	160	8	67	61	11	301

and the resampling vector for the first bootstrap sample of Fig. 2.1 would be:

0	0.4	0.2	0	0.2	0.1	0	0	0	0.1
---	-----	-----	---	-----	-----	---	---	---	-----

that is, 13 doesn't appear at all, but 106 appears 4 times, and so on. They then average these B resampling vectors, obtaining a final vector with the proportions averaged over all B vectors, denoted as \bar{P}^* and use this instead of $t(\hat{F})$ in eq.(3.2), obtaining,

$$\overline{\text{bias}}_B = \hat{\theta}^*(\cdot) - T(\bar{P}^*) \quad (3.5)$$

Example 3.2 We can illustrate the improved bias correction by using the original data of Example 3.1, but using $B = 500$. The proportions of \bar{P}^* will add to unity and $T(\bar{P}^*)$ is then calculated as a weighted variance, using the proportions of \bar{P}^* as weights (w_i). The weighted mean is

$\bar{x}_w = \sum w_i x_i$, and the variance is calculated as $s^2 = \frac{\sum w_i (x_i - \bar{x}_w)^2}{n}$. A run with $B =$

500 gave $\overline{\text{bias}}_B = \hat{\theta}^*(\cdot) - T(\bar{P}^*) = 2.910 - 3.223 = -0.313$ and bias-corrected estimate of s^2 is then 3.536, which is very close to the result (3.530) obtained by dividing by $n-1$, as should be done in practice. I would be inclined to use this approach on complicated problems, where bootstrapping uses a fair bit of computer time. Otherwise, one can simply use a sizable number (say 2,000) of bootstraps as in Example 3.1, inasmuch as it is likely that percentile confidence limits will be also be calculated in a practical example--here we know the "right" answer (i.e., divide by $n-1$) from theory.

3.4 Cross-validation

Models applied to ecological data may serve various purposes, but the more important uses may be to see how well we understand the data, and to make predictions. One of the earliest approaches to evaluating predictions from a model is very simple. One develops and fits the model on half of the available data, and then tests its predictions on the other half. Using all of the data for development and testing invariably results in underestimating the prediction errors. With the increased computing power now available, models can be fit to various subsets of the data and tested on the remainder. The logical outcome appears to be fitting the model to all but one of the observations, making a prediction for the remaining observation and repeating the process for all n observations, getting n predictions and deviations from predicted value. The variance is then calculated as:

$$\text{CVE} = \frac{1}{n} \sum (y_i - \hat{y}_{-i})^2 \quad (3.6)$$

where \hat{y}_{-i} denotes a predicted value based on all of the observations except the i^{th} value, and the summation runs from 1 to n . One thus deletes one observation, fits the model, and makes a prediction for the missing value, doing this n times to calculate the cross-validation error, CVE.

Example 3.3 Cross-validation error. To demonstrate the cross-validation idea, we use a larger set of regression data ($n = 30$). The data are as follows:

No.	x	y	No.	x	y
1	14.48	18.30	16	10.00	17.60
2	22.49	19.98	17	20.06	19.33
3	19.71	18.51	18	9.70	17.89
4	29.89	21.00	19	26.86	20.79
5	30.01	21.00	20	34.23	20.36
6	21.61	19.77	21	27.53	20.65
7	16.71	18.77	22	19.88	18.82
8	26.78	20.26	23	21.99	19.26
9	17.85	18.70	24	21.09	20.02
10	33.04	20.25	25	28.68	20.38
11	18.92	18.66	26	21.91	19.61
12	20.23	18.45	27	28.50	20.45
13	28.24	20.60	28	20.01	18.93
14	29.77	21.10	29	17.62	19.09
15	22.92	19.24	30	22.90	19.87

Applying cross-validation is simple in this case, with the only difficulty being one of arranging to drop each observation in turn. We then can compute CVE from eq.(3.6), which turns out to be 0.201. Note that the values of \hat{y}_{-i} in eq. (3.6) are computed from individual regressions dropping the i^{th} point, and y_i is the y -value of the i^{th} observation. For comparison the value of the variance about regression (eq.(2.5)) is 0.182. This is somewhat smaller, as might be expected because the deviations from regression are from a normal distribution in this example, and thus the normal-theory model gives the best estimate. Cross-validation would be used only in the absence of suitable theoretical estimators for the model parameters.

3.5. Bootstrapping for predictions. According to Efron and Tibishirani (1993:Chap. 17) bootstrapping offers an alternative to cross-validation. They focus on estimating the variance (in this example, variance about regression as given by eq. (2.5)). There are two stages in the bootstrapping approach. The first is to obtain bootstrap samples from the data set (given in Example 3.3), calculate regression lines for each bootstrap sample, and calculate a variance about each such regression (eq.(2.5)) using the original data set as x_i and y_i values. The second stage is to calculate the variance about regression for the bootstrap sample now using only the bootstrap sample (the y_i^* and x_i^* values). Thus, two variances about regression are calculated from the same regression equation, using the

following variances (the regression coefficients are those calculated on the bootstrap data in both cases):

$$s^2 = \frac{\sum(y_i - (a + bx_i))^2}{n - 2} \quad \text{and} \quad s^2 = \frac{\sum(y_i^* - (a + bx_i^*))^2}{n - 2}$$

Often, the variance about regression obtained from the bootstrap sample will be appreciably smaller than that obtained from the original data; and it is the mean difference of these two variances that is sought here. A few values from bootstrapping follow:

Bootstrap no.	Variance about regression using orig. data for y and x	Variance about regression using bootstrap values of y and x	Difference
1	0.1786	0.1905	-0.0119
2	0.1804	0.1863	-0.0059
3	0.1820	0.1519	0.0301
4	0.1724	0.1514	0.0210
5	0.1814	0.1759	0.0055
6	0.1948	0.1301	0.0647
7	0.1863	0.1277	0.0586
8	0.2023	0.1449	0.0574
9	0.1731	0.1189	0.0542

The "inflation factor" (mean difference) is added to the variance of the original data set to give an improved estimate. In the present example, the mean difference in the two variances about regression is small (0.028) so adding it to the variance about regression calculated from the original data (0.184) makes only a minor change. However, adding the correction to 0.184 gives a value (0.212) closer to that obtained in example 3.3. Nonetheless the best estimate is that of the original regression calculation because the data of Example 3.3 were generated from a bivariate normal distribution. Note that the bootstrap operation is a sampling procedure so that there will be small differences in the mean differences in repeat runs. Two further runs with B=2000 gave mean differences of 0.027 and 0.029.

It turns out that the correction indicated above is really a correction for bias. Inasmuch as the data we used in this example were normally distributed, the variance estimate should be unbiased, and the bootstrap analysis consequently comes up with a minor change, as would be expected with an unbiased estimator. The data used were drawn from a bivariate normal distribution, which is the basis for the normal theory confidence limits on a correlation coefficient, so it is worthwhile to compare (Fig. 3.2) the confidence limits on r based on bootstrapping x,y pairs with the normal theory limits in this example. The two different calculations of confidence limits are now appreciably closer than they were in Example 2.4. However, regression data approximating the bivariate normal distribution are not often encountered in practice, because one usually somehow selects the x-values used, rather than obtaining them at random, as the bivariate normal regression theory assumes.

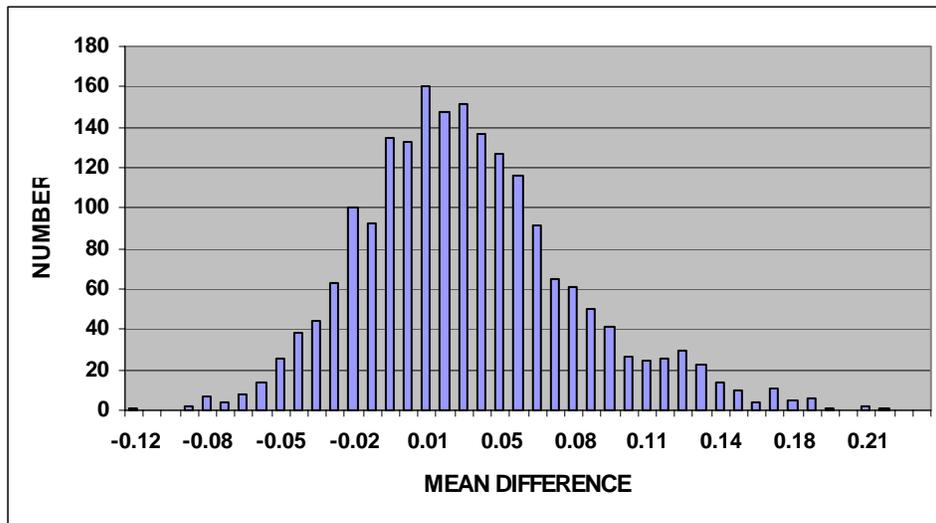


Fig. 3.1. Mean differences between variances about regression calculated from 2000 bootstraps. The differences are between variances about regression calculated from a regression line based on an individual bootstrap sample, with the first variance calculated using the original data, and the second calculated from the data of the particular bootstrap sample.

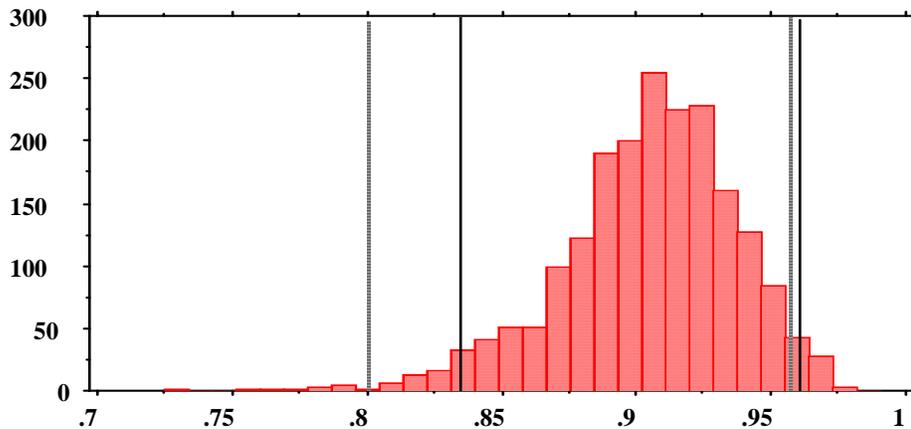


Fig. 3.2. Distribution of 2,000 bootstrap samples for correlation coefficient obtained from the data of Example 3.3. Upper 95% confidence limits are essentially the same, while the lower limit from normal theory is substantially lower than that from bootstrapping.

3.6 Improved confidence intervals

Efron and Tibishirani (1993:Chap. 14) recommend an improved bootstrap confidence interval which they call the BCa method ("bias-corrected and accelerated"). These intervals require more calculations than the simple percentile bootstrap confidence intervals. The data from Example 3.3 will be used to illustrate. Two proportions, α_1 and α_2 are calculated, and the total number of bootstraps, B , multiplied by these values. One

then orders the B bootstrap values and takes the $\alpha_1 B$ and $\alpha_2 B$ values as confidence limits. Thus if $B = 2,000$ and α_2 turns out to be 0.0166 as in the example here, then we take the bootstrap value $0.0166(2000) = 33$ as the ordered value of the bootstrap replications that gives the lower confidence limit.

Two initial values are calculated. The first is a bias-correction:

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^*(b) < \hat{\theta}\}}{B}\right) \quad (3.7)$$

here, $\Phi^{-1}(\cdot)$ indicates the inverse of the cumulative normal distribution, which can be looked up in tables and is available in various computer programs (in Microsoft Excel it is "NORMSINV"). The quantity in parentheses has, as numerator, the number of bootstrap samples that are less than the parameter estimate. In the example given below, we consider the correlation coefficients of Section 3.5, for which the correlation coefficient calculated from the original data was $r = 0.9043$, so we tally the number of bootstrap results that were less than this value and divide by $B = 2,000$, and look up the inverse cumulative normal value. In this case (Fig. 3.2) there were 942 values less than $r = 0.9043$, so we look up $942/2000 = 0.471$ in the inverse normal tables, getting -0.0728 for \hat{z}_0 .

The second value is the "acceleration", \hat{a} . This is calculated by jackknifing, (Section 3.7) using much the same procedure as in the cross-validation example above. We delete each observation in the original data set in turn, and calculate the correlation coefficient from the remaining observations. From the 30 observations tabulated in Example 3.3, we thus get 30 correlation coefficients, which are here designated as $\hat{\theta}_{(i)}$, where the subscript (i) indicates that the parameter estimate $\hat{\theta}$ (here, r) has been calculated from the original data set with each observation deleted in turn, and $\hat{\theta}_{(\cdot)}$ is used to indicate the average of these 30 values. The estimate of "acceleration" is then calculated as:

$$\hat{a} = \frac{\sum(\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^3}{6\{\sum(\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2\}^{3/2}} \quad (3.8)$$

where the summations are from 1 to $n = 30$ in this example. These two parameters (\hat{z}_0 , \hat{a}) are then used to calculate α_1 and α_2 , but require two more values for the calculation. These are designated $z^{(\alpha)}$ and $z^{(1-\alpha)}$, and are the values that cut off a proportion, α , from each tail of the unit normal distribution.. For 95% confidence limits, we look up $z^{(\alpha)}$ as 0.975 in the inverse cumulative normal table, getting 1.95996, and use 0.025 for $z^{(1-\alpha)}$, giving -1.95996. The calculations then are:

$$\alpha_1 = \Phi\left[\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha)})}\right] \quad (3.9)$$

and

$$\alpha_2 = \Phi \left[\hat{z}_o + \frac{\hat{z}_o + z^{(1-\alpha)}}{1 - \hat{\alpha}(z_o + z^{(1-\alpha)})} \right]$$

where values of $\Phi(\cdot)$ are to be looked up in tables of the cumulative normal distribution (in Excel, these are available as NORMSDIST). As noted in the introduction to this section, the lower limit is the 33d ordered value of the bootstrapped correlation coefficients (0.825), and the upper limit is $\alpha_1 B = 0.964(2000)$ giving the 1927th ordered value (0.957).

An example for the data of Example 3.3 is given below. The calculations are time-consuming and are best done with a computer program (see Section 3.10).

The improvement in bootstrap confidence limits in this example is not large, but suggests that the calculations do result in better bootstrap confidence limits. Data for Example 3.3 came from a bivariate normal distribution in which $\rho = 0.90$. From normal theory, 95% confidence limits were 0.803 to 0.954, while the percentile bootstrap limits in one run with $B = 2,000$ were 0.830 to 0.959. The improved confidence limits were 0.825 to 0.957, giving a lower limit closer to the normal theory result. Percentile limits vary a little in successive runs, giving 0.837-0.958 and 0.833-0.960 in two additional runs with $B = 2,000$.

	Correlations	CUBE TERM	SQ TERM
1	0.8978444	2.67703E-07	4.1537E-05
2	0.9075823	-3.57065E-08	1.08434E-05
3	0.9083542	-6.71633E-08	1.6523E-05
4	0.8996521	9.97194E-08	2.1504E-05
5	0.8993546	1.2017E-07	2.43518E-05
6	0.9066940	-1.39052E-08	5.78254E-06
7	0.9018049	1.53344E-08	6.17225E-06
8	0.9027515	3.63677E-09	2.36489E-06
9	0.9018945	1.37354E-08	5.73536E-06
10	0.9181345	-2.65399E-06	0.00019169
11	0.9038609	7.86151E-11	1.83515E-07
12	0.9118957	-4.40089E-07	5.78576E-05
13	0.9013883	2.4416E-08	8.41621E-06
14	0.9008332	4.12818E-08	1.19446E-05
15	0.9065879	-1.2145E-08	5.28363E-06
16	0.8874195	4.80101E-06	0.000284591
17	0.9044161	-2.03529E-12	1.60602E-08
18	0.8932690	1.3384E-06	0.000121448
19	0.9068001	-1.58286E-08	6.30419E-06
20	0.9205568	-4.30489E-06	0.000264632
21	0.9029564	2.36797E-09	1.77658E-06
22	0.9045651	-2.09626E-11	7.60263E-08
23	0.9048018	-1.34558E-10	2.62587E-07
24	0.9130072	-6.62566E-07	7.6001E-05
25	0.9019668	1.2528E-08	5.39412E-06
26	0.9047293	-8.51567E-11	1.93559E-07
27	0.9014553	2.27629E-08	8.03189E-06

	28	0.9040066	2.26062E-11	7.995E-08	
	29	0.9046723	-5.61539E-11	1.4664E-07	
	30	0.9054246	-1.46324E-09	1.28887E-06	
	AVERAGE	0.9042893	-1.44489E-06	0.00118043	SUMS
	Z(0)-HAT		-0.07276		
	ACCELERATION		-0.0059377		
	Z(0)-HAT+Z-ALPHA		1.88716	1.95996	
1-ACCEL	Z(0)-HAT+Z-ALPHA		1.0112055		
	RATIO		1.8662478		
	ALPHA1		0.9635494	1927.0988	
	Z(0)-HAT+Z(1-ALPHA)		-2.03276	-1.95996	
1-ACCEL	Z(0)-HAT+Z(1-ALPHA)		0.987930		
	RATIO		-2.057595		
	ALPHA2		0.016569	33.139	

3.7 The jackknife

The jackknife technique, as noted in the introduction to Chapter 2, pre-dates bootstrapping, and was originally derived (Quenouille(1956)) to evaluate biases in an estimator. The technique is very simple and easy to apply. Given an original data set, one simply leaves out each observation in turn and calculates the statistic of interest on the remaining observations, as was done in the calculations for improved confidence limits above, getting $\hat{\theta}_{(i)} = s(\mathbf{x}_{(i)})$, where $\mathbf{x}_{(i)}$ is the vector of observations with the i^{th} observation removed, and $s(\cdot)$ denotes some statistic calculated from these observations. The bias estimate is calculated as:

$$\hat{\text{bias}}_{\text{jack}} = (n-1)(\hat{\theta}_{(\cdot)} - \hat{\theta}) \quad (3.10)$$

where $\hat{\theta}_{(\cdot)}$ denotes the mean of the $\hat{\theta}_{(i)}$ and $\hat{\theta}$ is the statistic estimated from the original data. The jackknife estimate of standard error is:

$$\hat{\text{se}}_{\text{jack}} = \left[\frac{n-1}{n} \sum (\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)})^2 \right]^{1/2}. \quad (3.11)$$

We can illustrate the calculations with the data of Example 2.1. The following table shows the 10 original observations and the 10 jackknife samples created by dropping each observation in turn. If we consider the mean as the statistic to be jackknifed, then the bias estimate from eq.(3.10) turns out to be zero, inasmuch as the mean of the original observations necessarily equals the grand mean of the jackknife samples.

Orig. data	Jackknife samples									
	1	2	3	4	5	6	7	8	9	10
13		13	13	13	13	13	13	13	13	13
106	106		106	106	106	106	106	106	106	106
203	203	203		203	203	203	203	203	203	203
131	131	131	131		131	131	131	131	131	131
160	160	160	160	160		160	160	160	160	160
8	8	8	8	8	8		8	8	8	8
67	67	67	67	67	67	67		67	67	67
61	61	61	61	61	61	61	61		61	61
11	11	11	11	11	11	11	11	11		11
301	301	301	301	301	301	301	301	301	301	
Ave.	116.4	106.1	95.33	103.3	100.1	117	110.4	111.1	116.7	84.44

Eq. (3.11) gives $\hat{se}_{\text{jack}} = 30.15$, while a bootstrap estimate of standard error [eq.(2.1)] is 28.70, and the standard error of the original data is also 30.15, as it should be in this case, because the jackknife standard error formula gives the same result for the standard error of a mean.

Example 3.4 Jackknifing a regression equation

Grizzly bears are very difficult to census. Also, they range very widely, are difficult (and somewhat dangerous) to trap and are not numerous. Adult females with cubs-of-the-year may tend to spend more time in the open than other bears, and such family groups can be approximately identified by group size, age of cubs, location, etc. Consequently the only long-term index of abundance for bears in Yellowstone has been an annual "count" of such family groups. The index is quite variable, so it is essential to learn as much about the effect of variability as possible, and to look for ways to improve the index. For further study here, logarithms of the index count are used because a linear relationship would likely result if the counts are directly proportional to population abundance. A plot of the data (Fig. 3.3) shows the substantial variability.

The jackknife, the bootstrap, and cross-validation can be used to study the index. To use the jackknife approach, one proceeds as in the example shown in Section 3.7. There are 19 annual values of the index, so the original data are copied 19 times, and each of the paired items (year and ln count) is removed in turn and placed at the top of the table. The gaps in the main body of data are then closed. For each of the paired columns of data, one then estimates a slope (using the SLOPE function in EXCEL) and calculates the intercept from y and x means. This thus gives the basis for a regression line at the bottom of each set of data. This regression line is then used to compute an estimate for the missing point (using the x-value at the top of the table) and that prediction is placed below the value left out located at the top of the table. The resulting 19 data pairs then provide data for calculation of CVE by eq. (3.6). The first two columns of a calculation appear in a table below.

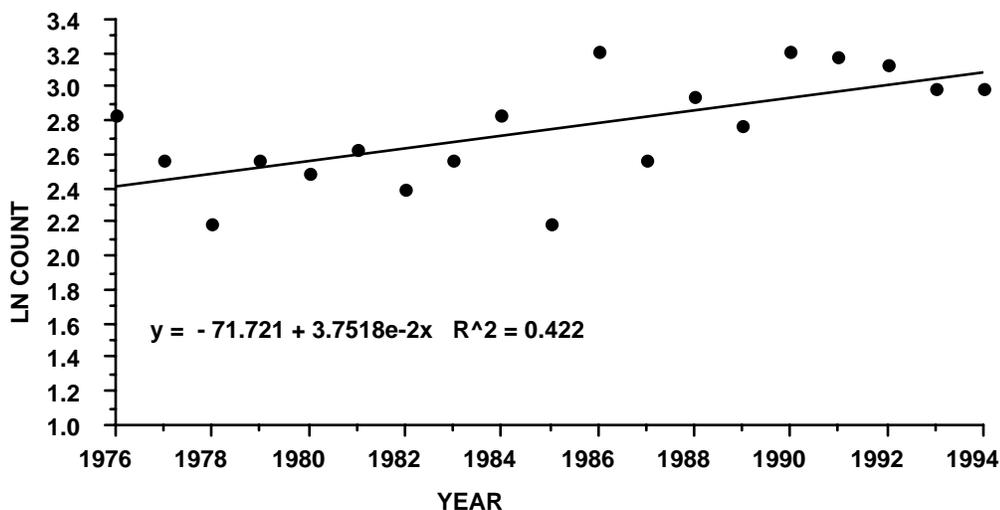


Fig. 3.3 Index of grizzly bear abundance in Yellowstone National Park.

The slope estimates at the bottom of the table are used to jackknife the data for calculating a standard error from eq. (3.11) which turns out to be 0.0104. This is perhaps most useful if divided by the jackknife mean, giving a coefficient of variation of 0.28, indicating the considerable variability in the data. The data can also be used to calculate the jackknife bias estimate of eq.(3.10), which appears to be very small. The operation can be described in steps as follows:

- (1) Duplicate the two columns of data (x and y variables) n times, where n is the number of observations available.
- (2) Remove each value in turn and put it above the table of values, leaving space for a predicted value.
- (3) Move up the data to close the gaps.
- (4) Calculate slopes and x and y means for each column.
- (5) Use this regression data to calculate a predicted value for the x-value of the item removed from that column and place the predicted value below the removed value. The squared difference is then summed and divided by n to calculate CVE.
- (6) Use the calculated slope values to produce jackknife estimates of standard error and bias.

All of the above provides some information on how an index behaves. It is, however, more useful in the situation where we have several possible candidates for an index, as the estimates of bias and CVE (and possibly other statistics calculated for the data) can be used to decide which of the candidates might give the best notion of trends in the bear population, which is of major importance in managing an important species. An improvement in the index is available by way auxiliary variables that provide a correction for the variation in visibility of bears, which presumably is at least partially responsible for fluctuations in the number seen from year to year. This improved index was described by Eberhardt et al.(1999).

Table of first few columns of data from Jackknifing bear index.

LEDT OUT				2.8332	1	2.5649	2
PREDICT				2.3140		2.4304	
DEV SQ				0.2696		0.0181	
ORIGINAL DATA							
		LN COUNT					
		Y	X	Y	X	Y	X
1976	17	2.8332	1	2.5649	2	2.8332	1
1977	13	2.5649	2	2.1972	3	2.1972	3
1978	9	2.1972	3	2.5649	4	2.5649	4
1979	13	2.5649	4	2.4849	5	2.4849	5
1980	12	2.4849	5	2.6391	6	2.6391	6
1981	14	2.6391	6	2.3979	7	2.3979	7
1982	11	2.3979	7	2.5649	8	2.5649	8
1983	13	2.5649	8	2.8332	9	2.8332	9
1984	17	2.8332	9	2.1972	10	2.1972	10
1985	9	2.1972	10	3.2189	11	3.2189	11
1986	25	3.2189	11	2.5649	12	2.5649	12
1987	13	2.5649	12	2.9444	13	2.9444	13
1988	19	2.9444	13	2.7726	14	2.7726	14
1989	16	2.7726	14	3.2189	15	3.2189	15
1990	25	3.2189	15	3.1781	16	3.1781	16
1991	24	3.1781	16	3.1355	17	3.1355	17
1992	23	3.1355	17	2.9957	18	2.9957	18
1993	20	2.9957	18	2.9957	19	2.9957	19
1994	20	2.9957	19				
MEANS		2.75	10.0	2.75	10.5	2.76	10.4
SLOPES		0.0375		0.0457		0.0394	
INTERCEPTS		2.3776		2.2683		2.3516	
<u>S.S. of</u>		<u>SLOPES</u>		0.0001		0.0000	

3.8 The Monte Carlo method

In many situations, it is desirable to seek a way to check on the validity of possible estimators. If the stochastic process leading to the data under study can be modelled in a realistic manner, then it is usually possible to test estimation and analysis methods by "Monte Carlo" simulations. Many detailed papers and a sizable number of books deal with such approaches, and all that will be attempted here is to provide a sketch of the method, and a simple example. Exercise 1.14.6 discusses simulation of a continuous frequency distribution, the exponential distribution. The underlying model for survival times is, in fact, the exponential, although survival may also need to be described by more complex models. Given a way to generate a sample from a plausible distribution, one can then use such data to test estimation or analysis schemes.

For a concrete example, we consider the percentile confidence limits discussed in Chapter 2, and demonstrated in Fig. 2.2, and ask whether these limits are valid. This question is usually discussed in terms of coverage. For convenience, consider 95% confidence limits. These are described as limits that should include the true but unknown mean in 95% of a very large series of repetitions of the same process from which a given observed sample is generated. Note that nothing is said about a particular case -- it is only the long-run average that we can depend on. If confidence limits are properly constructed, then they should "cover" the (unknown) true mean 95% of the time. If we assume that observed survival time data come from an exponential distribution, then we can generate a very large number of samples of $n = 20$ "observations", calculate bootstrap

confidence limits from these samples, and see how well they "cover" the true mean. In this case, we can know the true mean, inasmuch as it can be calculated for the exponential distribution, $E(x) = 1/\beta$. Using $\beta = 0.01$ results in an expected ("true") mean of $1/\beta = 100$.

A BASIC program was used to study the confidence limits. It turns out that 1,000 simulation runs with 1,000 bootstraps for each sample of $n = 20$ yields 906 cases where the calculated percentile limits included the true mean of 100, whereas one would expect 950 cases inside the limits for a true 95% level of significance. Note that this result (906 of 1,000) is subject to sampling error; a binomial calculation gives $v(p) = p(1-p)/1000$ where $p = 0.95$, so that two standard errors on p will be about 0.013. Consequently, it would appear that the bootstrap "coverage" is significantly short of the expected 95%. Nonetheless a nominal 91% isn't too bad for confidence limits. Fig. 3.4 provides an example of coverage from this study.

The exponential distribution is sharply skewed and the standard deviation equals the mean so that the survival time thus generated is highly variable. For an alternative, we can run the Monte Carlo study using normally distributed variables with the same mean (100) and a smaller standard deviation (10). This can be done by using the Box-Muller approximation (Bratley et al. 1983) to generate unit normal random variables, replacing the exponential in a BASIC program. This generates two approximately normal random variables with zero mean and variance of unity from two uniform random variables, and these are then transformed to have standard deviation of 10 and mean of 100. Running 1,000 simulations each using 1,000 bootstraps on samples of $n = 20$ from the normal distribution gives coverage of 931, appreciably closer to the expected 95%.



Fig. 3.4. An example of "coverage" for the simulation to test bootstrap percentile limits on simulated data from an exponential distribution. The figure shows confidence limits for a sample of 20 observations out of the 1,000 simulations used to test confidence limits calculations (with 1,000 bootstraps per sample of 20 observations). The true mean is at 100, while squares represent upper 95% confidence limits and circles the lower limits.

Lines connect the limits for the three cases where the limits did not include the true mean. In this small sample, coverage was $17/20 = 0.85$.

Inasmuch as we are dealing with means, the usual approach to confidence limits would be to calculate a variance from the original data, and obtain confidence limits with a multiplier from the t-distribution. Such results can easily be simulated, using the same methods for generating exponential and normal random variables. A BASIC program was used to simulate samples and confidence limits from the normal distribution, and the same program was used with the exponential generator. Results using sample sizes of $n = 10, 20,$ and 30 appear in the following table. These results suggest that constructing confidence limits in the usual manner from exponential data does a little better job than bootstrapping, and for data from a normal distribution the limits are within sampling error of the expected 95%, while bootstrapping falls a little short. One would not, of course, use bootstrapping to obtain confidence limits on means. It is best reserved for situations where there is no convenient theoretical approach.

Sample size	Exponential simulation		Normal simulation	
	Bootstrapping	Usual limits	Bootstrapping	Usual limits
10	837	900	904	944
20	906	923	931	941
30	927	921	933	953

3.9 The delta method

The delta method is a useful adjunct to bootstrapping. It has been used for many years to approximate the variance of complex functions of random variables. It is obtained from a Taylor expansion of the function in which the second degree terms are retained, and rearranged to represent variances of the random variables. The expression is as follows:

$$V[g(x)] = \sum_{i=1}^n V(x_i) \left(\frac{\partial g}{\partial x_i} \right)^2 + 2 \sum_{i < j} \text{cov}(x_i, x_j) \left(\frac{\partial g}{\partial x_i} \right) \left(\frac{\partial g}{\partial x_j} \right) \quad (3.12)$$

where $V[g(x)]$ represents the variance of some function, $g(\mathbf{x})$, where \mathbf{x} is a vector of random variables, x_1, x_2, \dots, x_n . $V(x_i)$ denotes the variance of the variable x_i , which is multiplied by the square of the partial derivative of $g(x_i)$. Covariance terms are calculated for those cases where $i < j$. In many cases, it may be that the random variables are independent, so that the covariance terms can be assumed to be zero, and the right-hand portion of eq.(3.12) can then be dropped.

Bootstrapping can be used to calculate a variance for $g(\mathbf{x})$ without any need to calculate variances and covariance of the individual random variables or to obtain partial derivatives. The delta method becomes a valuable adjunct, however, when it is possible to design the study in order to minimize $V[g(\mathbf{x})]$. In practice, $V[g(\mathbf{x})]$ may be appreciably larger than is desirable, and we may wish to design a new study with larger samples (or to supplement the existing samples). In this case, it is essential to be able to determine the

effect of increasing the sample sizes for individual random variables. The delta method provides a way to calculate the effects of changing sample sizes on the overall variance.

Example 3.5 Application of bootstrapping to a complex function. Obtaining a variance for the Lotka-Leslie model provides a good example of the utility of bootstrapping. The underlying equation for this model is:

$$1 = \sum \lambda^{-x} l_x m_x \quad (3.13)$$

Here, λ represents the rate of change of an age-structured population having age-specific survivorship rates l_x and age-specific reproductive rates m_x . The general model for the Lotka-Leslie function does not have a "closed-form" solution. That is, there is no way to write eq.(3.13) in a linear form, that is to provide an expression stated as $\lambda = g(x)$. It is thus necessary to solve eq.(3.13) for λ by an iterative procedure, i.e., by varying values of λ until one satisfies the equation. Because there is no linear expression for a solution for λ , developing an expression for the variance becomes very difficult. Bootstrapping then provides a convenient approach. One only needs to set up the data on l_x and m_x in tables, sample these tables of data with replacement, and calculate values of λ from the samples. The percentile method then provides convenient confidence limits.

In many instances, the samples available for calculations are too small to make calculations from eq.(3.13) feasible. An alternative may then be needed. A useful approximation (Eberhardt 1985) is:

$$\lambda^a - s\lambda^{a-1} - l_a m [1 - (\frac{s}{\lambda})^{w-a-1}] = 0 \quad (3.14)$$

Here, a is the age at which full reproductive rate is achieved, l_a is survival from birth to age a , s is survival beyond that age, and w is an age at which calculations are truncated in order to compensate for the effects of senility. This equation again must be solved by iteration, and can readily be bootstrapped. The delta method can be used to study the components of variance and thus to determine the effect of increasing sample sizes for the several components on the variance estimate for λ . In several examples, the delta method gives very much the same variance estimate as bootstrapping. Another benefit of the delta method calculations is that the partial derivatives serve to indicate the relative importance of the several components, indicating, for example, that small changes in adult survival have the maximum effect on λ . Because there is no linear solution for λ , the delta method has to be applied by using implicit differentiation. Solutions appropriate for eq.(3.14) appear in the following references, which also give details and result of the application of bootstrapping to this complex function. Calculations for grizzly bears appear in Eberhardt et al. (1994), for sea otters in Eberhardt (1995), for monk seals in Gilmartin and Eberhardt (1995), and for manatees in Eberhardt and O'Shea (1995). Selected examples appear in Chapter 11. A program (APPLMB) to calculate confidence limits on data used in eq.(3.14) is available in Appendix A.

3.10 Programs for data analysis

Some extensive calculations are involved in the preceding sections. Computations for cross-validation and jackknifing are not too burdensome and can be done from the

same setup as illustrated in Example 3.4. Programs that avoid most of the hand-calculations are available in the Appendices. Appendix A does not require any knowledge of computer programming, while Appendix B requires some effort to gain minimal proficiency in the R-language. However, once one gains some experience with it, results are easily and quickly obtained. The following notes give cross-checks on several of the Sections of this Chapter.

Cross-validation

As remarked above, cross-validation is not too difficult to do directly on a spreadsheet, and Example 3.3 was worked that way. If one loads the “boot” package in R (Appendix B), and brings in vectors x and y with the data of Example 3.3, then the following code produces data for the cross-validation error (CVE). Note that it may be necessary to use statements, “ $x=as.matrix(x)$ ” and “ $y=as.matrix(y)$ ” to be sure that the x and y data are not treated as “lists” (this can be checked by statements like “length (x)”, which should give the number of items in the vector (30). If it comes up with the length as “1”, then make the indicated change to matrix form).

```
{theta.fit=function(x,y){lsfit(x,y)}
theta.predict=function(fit,x){cbind(1,x)%*%fit$coef}
results=crossval(x,y,theta.fit,theta.predict)}
```

Typing “results” gives:

```
$cv.fit
[1] 18.39368 19.52994 19.16999 20.58385 20.60205 19.40486 18.70319 20.16172
[9] 18.88301 21.19587 19.04597 19.24901 20.36361 20.55723 19.61994 17.76607
[17] 19.18723 17.64574 20.14681 21.40597 20.25370 19.18175 19.48001 19.31688
[25] 20.44596 19.45570 20.41327 19.19649 18.82414 19.59521
$ngroup
[1] 30
$leave.out
[1] 1
$groups
NULL
$call
crossval(x = x, y = y, theta.fit = theta.fit, theta.predict = theta.predict)
```

The values under “\$cv.fit” are the predicted values from regressions calculated with one observation left out. Set $z=$cv.fit$ and calculate CVE from “ $sum((y-z)^2)/30$ ” which gives 0.2009693, which is CVE, as obtained from the more tedious calculation with EXCEL.

Bootstrapping for predictions

Section 3.5 points out that bootstrapping can provide an alternative to cross-validation, and describes the methodology. One can get results for the data of Example

3.3 from “BOOT2” in Appendix A. Load in the 30 x,y pairs and compute slope and intercept below columns 4 and 5 (which contain the bootstrap samples) as well as below the first 2 columns (which contain the original data). On the right side of the sheet (away from the bootstrap calculations) set up two columns, one to calculate the regression sum of squares (eq.(2.5)) from the original x,y data but using the parameters (slope and intercept) from the bootstrap data and a second column to calculate regression sum of squares from the bootstrap samples also using parameters from the bootstrap “replication”. Sum these columns and divide by $n-2 = 28$ to get two estimates of regression sum of squares. Load the difference (calculation from original data minus calculation from bootstrap sample) into the “parameter input” box and run, say 2,000 bootstraps. The mean of these gives the “inflation factor” to add to the variance of the original data set (about 0.028) to give the improved estimate of variance, as described in Sec. 3.5. The R-language program(Appendix B) is:

```
{theta.fit=function(x,y){lsfit(x,y)}
theta.predict=function(fit,x){cbind(1,x)%*%fit$coef}
sq.err=function(y,yhat){(y-yhat)^2}
results=bootpred(x,y,2000,theta.fit,theta.predict,err.meas=sq.err)}
```

It gives:

```
> results
$app.err
[1] 0.1702293
$optim
[1] 0.0283673
$serr.632
[1] 0.1984177
$call
bootpred(x = x, y = y, nboot = 2000, theta.fit = theta.fit, theta.predict = theta.predict,
  err.meas = sq.err)
```

Where “\$optim” gives the result obtained above (0.028). Add this to “\$app.err” to get 0.1985966 which compares to 0.2009693 obtained from cross-validation of the data in the section above. We thus have two methods for improved estimates of variance about regression. One should not, of course, report all of the values beyond the decimal point given by this and similar programs!

Improved confidence limits

The BCa method of Section 3.6 is time-consuming to calculate directly. Two lines in R give the results immediately (again using data of Example 3.3). One first combines the x and y vectors as “xdata” with the statement “xdata=cbind(x,y)”, and then uses:

```
{theta=function(x,xdata){cor(xdata[x,1],xdata[x,2])}
results=bcanon(1:30,2000,theta,xdata)}
```

This gives:

```
> results
```

```

$confpoints
  alpha bca point
[1,] 0.025 0.8286273
[2,] 0.050 0.8384993
[3,] 0.100 0.8542274
[4,] 0.160 0.8662237
[5,] 0.840 0.9275239
[6,] 0.900 0.9355572
[7,] 0.950 0.9448605
[8,] 0.975 0.9524536
$z0
[1] -0.1218725
$acc
[1] -0.005937748
$u
[1] 0.8978444 0.9075823 0.9083542 0.8996521 0.8993546 0.9066940 0.9018049
[8] 0.9027515 0.9018945 0.9181345 0.9038609 0.9118957 0.9013883 0.9008332
[15] 0.9065879 0.8874195 0.9044161 0.8932690 0.9068001 0.9205568 0.9029564
[22] 0.9045651 0.9048018 0.9130072 0.9019668 0.9047293 0.9014553 0.9040066
[29] 0.9046723 0.9054246
$call
bcanon(x = 1:30, nboot = 2000, theta = theta, xdata)

```

One can simply pick off the desired confidence limits (0.8286 to 0.9524 for 95% confidence interval), but the program also provides the key items from the calculations of Section 3.6, namely z_0 and the “acceleration”, along with the 30 values of the correlations given in the first column of the table of Section 3.6. Equation (3.7) gives the calculation of the bias correction:

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^*(b) < \hat{\theta}\}}{B}\right)$$

which involves counting the number of bootstrap values below the correlation from the original data ($r = 0.9043$) and transforming by the inverse of the cumulative normal distribution. This can be somewhat variable, so the value from the R-program is a little larger than that calculated in Section 3.6 (another bootstrap run of 2,000 gave z_0 of -0.127). Note that the “acceleration” agrees closely with that obtained in Section 3.6.

3.11 Exercises

3.11.1 Inasmuch as bootstrapping is a sampling procedure, additional runs of B bootstraps will give slightly different results, even if B is large. Conduct a bootstrapping check on the data of Example 3.1 to see how your bias adjustment compares with the results given there. Use $B=2000$. Do 10 trials and record results on a summary sheet (don’t forget to use PASTE SPECIAL and VALUES or you may get a statement like “Circular References” or “Link to another spreadsheet”). This should show that the bias is consistent, and that the corrected value is a much better estimate of the true value. However, when there is an unbiased estimate based on theory (as in this case), one

obviously should use that value. The bias estimate is important only when you don't have an estimate that is known to be unbiased (which is often the case with ecological data, even though it might not be a widely recognized fact). Also number the observations serially (1-10) and calculate the correlation coefficients.

3.11.2 Bias corrections.

Use the data of Example 2.3 (calculations in Exercise 2.10.4) to further explore bias corrections. In that example we used regressions of the natural logarithm of number of survivors (Fig. 2.5) on year to estimate a survival rate (slope of the regression line) and then transformed it back to an annual rate by calculating $y = \exp(b)$. Bootstrap confidence limits were obtained and also transformed back to annual rates. Use eq.(3.2) to examine the bias in transforming back. When there is an evident bias, one should examine the confidence interval on the estimate to see if the bias is large relative to the confidence interval.

$$\hat{\text{bias}}_B = \hat{\theta} * (\cdot) - t(\hat{F}) \quad (3.2)$$

3.11.3 Make a frequency distribution of z (eq.(2.9)) using the correlation coefficients computed in exercise 3.11.1. Does this look like a normal distribution as assumed in calculating confidence limits under the usual theory? Compare your results with Exercise 2.10.5.

3.11.4 The regression bootstrap of Example 2.2 used bootstrapping residuals in which deviations from a model fitted to the original data are bootstrapped. In exercise 2.10.7 we tried bootstrapping the x, y pairs directly, and got some strange-looking results. However, larger samples (more x and y values) appear to give comparable results with both methods. Efron and Tibishirani warn that the approach using residuals is "model-dependent", i.e., if the model is wrong, the results may be doubtfully useful. Hence, its worthwhile to repeat the exercise using the data of Example 3.3. Doing this directly is cumbersome, so it is best to use the program furnished in Appendix A (BOOT2). Compare your results with the slope and confidence intervals given by the regression program in EXCEL. This exercise is worthwhile in that ecologists use regressions with smallish samples and the independent variables are not always known with certainty. There don't seem to be any guidelines as to sample sizes in such cases, so its wise to use both approaches and to check for bias (Eq. (3.2)) if you want to be comfortable with your results. The frequency diagrams of Exercise 2.10.7 were distinctly bimodal, making it clear that the nonparametric approach is not advisable with only 10 pairs of observations.

3.11.5 Example 3.4 gives the approach to jackknifing a regression line in which logarithms of data on an index of bear abundance are fitted by linear regression (Fig. 3.3) and the fit examined by cross-validation, with a check on bias from eq.3.10. Complete the analysis just as in Example 3.4. Compute the cross-validation error (CVE), jackknife standard error of the slopes, and Biasjack of the slopes. Compare the jackknife standard error with that of the slope computed with the usual EXCEL regression analysis. Also compare CVE with the residual mean square of the regression analysis.

3.11.6 Jackknifing was used in Example 3.4 because it is fairly easy to apply and we could compute CVE of eq. (3.6) in the same operation. However, bootstrapping has some advantages, and likely should be used to estimate bias and confidence limits whenever it is feasible. Use the data of Example 3.4 to conduct bootstrapping of the deviations to compute the bootstrap bias estimate of Eq. (3.2) and 95% confidence limits on the slope. Use 1,000 bootstraps (for convenience in calculations using EXCEL – with 19 observations, I would be inclined to try both approaches). The regression bootstrap using x,y pairs can readily and quickly be computed from the program in Appendix A (BOOT2).

3.11.7 The approach of Section 3.5 is most readily calculated by using a programming language. However, it is feasible to do the calculations in EXCEL if one is willing to devote enough time to the job. A program in Appendix A (BOOT2) will do the job in short order and should be used to repeat the results of Section 3.5.

4.0 SAMPLING METHODS

4.1 Introduction

One of the weaknesses in ecology today is that too many investigators fail to realize the importance of sampling. A logical reason for this difficulty is that studies are often centered on one or two study areas, so that the investigator tends to forget that he is in fact studying a sample from some larger population. This may not be seen as a handicap until it becomes necessary to attempt to extend the results of the study to the larger area. The perceptive observer may then suddenly realize that there really isn't much basis for such an extrapolation, unless he does in fact have data from a number of subareas (i.e., a representative sample) on which to base the extrapolation and to provide a basis for assessing its validity.

The intent here is to provide a brief overview of sampling methodology. Most of the material follows the lines of survey sampling methods, as given in much more detail by Cochran (1977). Thompson (1992) includes methods of particular interest in ecology. The very basic features are those of an elementary statistics course. Most students will prefer to refer to familiar textbooks for these aspects. The essential material has to do with some elementary statistical concepts and a few standard distributions, mainly the binomial, hypergeometric, Poisson, and normal. Students not familiar with these distributions and the basic rules of probability should look them up in one of the elementary references. A brief sketch of the statistical background appears also in Chapter 1.

4.2 Simple random sampling

The main complication in defining simple random sampling is one of defining the meaning of the word "random". Our approach is that of probability theory, in which it is assumed that every sampling unit (some sort of explicitly defined entity) has the same probability (chance) of being drawn into the sample. The mechanics of drawing a random sample then depend on giving each unit the same chance of inclusion in a sample while keeping the choices independent of one another. The standard procedure is to assign a number to each unit in the population, and to refer to a table of random numbers as a device for selecting the sample.

Once the sample has been drawn and measurements made on the sample units, various problems of analysis of the data must be dealt with. However, procedures for analysis of the data need to be considered well in advance of the sampling to be sure that the right kinds of data are collected. That is, the investigator must first prepare a sampling plan, which designates exactly how the sample will be obtained. Secondly, there should be a definite plan for the analysis of the resulting data, specifying what statistical analyses will be carried out, and what will be done if a particular kind of result is obtained in the analysis. Many of the problems in field research are caused by the lack of such a study design. It may be objected that one cannot produce such a plan if it is not known in advance how the study will turn out. There are several answers to this objection. One is that few studies are conducted in completely new situations. Usually there are previous investigations that can be used in

the design stages, and data that can be used to test both the sampling plan and the analytical procedures.

4.3 The finite population correction

The estimated variance of a mean for sampling without replacement from a finite population is:

$$s^2(\bar{x}) = \frac{s^2}{n} \left[\frac{N - n}{N} \right] = \frac{s^2}{n} \left[1 - \frac{n}{N} \right] \quad (4.1)$$

and the finite population correction is just the quantity in brackets, or one minus the fraction of the population actually sampled (frequently designated as f). Thus when nearly all of the population is taken into the sample, the variance of the estimated mean becomes very small, as it logically should.

When the fraction of the population sampled is small, this equation implies that size of the population has little effect on the standard error and thus on confidence limits. This is a result that comes as a surprise to many people, who intuitively suppose that bigger samples are required for very large populations. This is, however, simply not true. A large population may, of course, offer more logistic problems in sampling and thus be more expensive to sample.

As a general rule of thumb, when the sampling fraction is less than about 5%, it is customary to neglect the finite population correction factor, and treat the sample as though it had been obtained by sampling with replacement. Sampling with replacement refers to circumstances where objects can be drawn from the population one at a time and replaced before the next object is drawn. With such a process, probabilities remain unchanged as the drawing proceeds, making calculations much simpler than if replacement does not occur, when removal of one individual changes the odds on selecting others in the next draw.

Students whose statistical training has come from courses in which hypothesis testing was mainly emphasized may not have encountered the notion of a finite population correction. This is because most tests of hypotheses are formulated on the basis of sampling from an infinitely large population, or on the basis of sampling with replacement.

4.4 Confidence intervals

Ideally, one would like to be able to know how far "off" a particular estimate is from the true parameter value. Statistical methods offer no such utopian result, and the best that we can do is to make probability statements that apply to the long-run of future trials, or to some hypothetical population just like the one currently under study. These take the form of confidence limits, which are a statement of the following kind:

$$P_{\Gamma}\{X_L < \mu \leq X_U\} = 1 - \alpha \quad (4.2)$$

where X_L denotes the lower confidence limit, X_U the upper, and the probability that the true, but unknown, value (μ) of the random variable of

interest will fall between these limits is $1 - \alpha$. The proper interpretation of this statement is that a very large number of repetitions of the "experiment" at hand would yield confidence limits that include the true, but unknown, μ in a fraction $1 - \alpha$ of the trials. It must be emphasized that the statement cannot be interpreted as pertaining to a single set of sampling results that are in hand. Such a statement would be ridiculous because the confidence interval then either includes the true value or it does not, and no probability is involved--it's just that we have no way of knowing where the true value lies. Hence we have to adopt some sort of long-run view of the "odds on being right."

One of the most common mistakes in reporting the results of a statistical analysis is to assert that "the probability is $1 - \alpha$ that the hypothesis is false". Just as with the confidence limits statement above, a testable hypothesis is either true or false, but there is no need for statistical analysis if one knows the answer. If the answer is not known, then the statistical approach attempts to supply some quantitative assessment of the "odds" for and against the hypothesis. The problem that many people have with this is that they have been admonished from childhood to "make up your mind". Such decisions should be stated as a belief based on the evidence, but announced separately from the probability statements used to assess the evidence.

Most investigators tend to use confidence limits that are symmetric about the estimate. No doubt this is a consequence of the common use of the symmetric normal distribution, which leads one to tend to cut off about $\alpha/2$ of the probability distribution on each side of the mean, and thus get symmetric limits. In point of fact there is nothing in theory or practice that says that the limits should be symmetric--all that is required is that there be $1 - \alpha$ of the distribution within the limits. Also, setting limits for a distribution like the Poisson is likely to result in asymmetric limits. One reason is the difficulty of cutting off an exact fraction (α) of the distribution when one must set the limits in terms of integer values. This difficulty can quickly be appreciated by trying some examples with tables of the Poisson distribution.

For ease in understanding and remembering the procedure for obtaining confidence limits, we will use the standardized or unit normal curve, and reverse the usual process of going from some other normal distribution to the standardized--that is we now look up a value (z) in tables of the unit normal that cuts off the desired proportion, α , of the distribution. If α is to be 0.05, then we find $z = 1.96$, and

$$\frac{X_c - \mu}{\frac{\sigma}{\sqrt{n}}} = \pm 1.96 \quad (4.3)$$

where X_c represents upper or lower confidence limit respectively corresponding with the plus and minus signs on the right hand side of the equation. Thus we have $X_c = \mu \pm 1.96 \sigma/\sqrt{n}$ and the probability statement previously given is satisfied by the corresponding choices of X_c (which are X_L and X_U). In practice, it is necessary to substitute \bar{x} for μ .

The above results assume that one knows σ , which is almost always not the case. If the sample size is about 30 or more, it really doesn't matter much, in that s/\sqrt{n} then provides an adequate estimate of the standard error (σ/\sqrt{n}). If sample size is small, then it is preferable to use tables of the "t" distribution (instead of the unit normal distribution), which make allowance for uncertainty about the variance. Values of the t-distribution are available in EXCEL.

One also needs to bear in mind that the estimate of σ (usually denoted by s) that is appropriate depends on what quantity one is setting limits on. For a mean, we proceed as above, but if limits on a single observation are to be secured, then we naturally use the standard deviation, s , in place of s/\sqrt{n} .

Example 4.1 Calculating confidence limits

For "95 percent confidence limits" ($\alpha = 0.05$), many people round 1.96 to 2.0 so that the limits can be calculated from $x_c = \bar{x} \pm 2s/\sqrt{n}$. Suppose $s = 9$, $n = 16$ and $\bar{x} = 10$. Then we have $x_c = 10 \pm 2(9)/4$, which can be expressed as $5.5 \leq \mu \leq 14.5$.

4.5 Determining sample size

Determining the sample size required to provide confidence limits of preassigned width on a mean is again a matter of using the z values. From the results above we have:

$$\pm \frac{zs}{\sqrt{n}} = X_c - \mu \quad \text{and we let } D = |\mu - X_c|$$

where the vertical lines denote "absolute value of...", so that D amounts to the half-width of the desired (symmetric) interval. Hence:

$$n = (zs/D)^2 \tag{4.4}$$

Another way to approach sample size estimation is to express D relative to the mean:

$$D/\bar{x} = \frac{zs}{\sqrt{n} \bar{x}}$$

so that

$$n = [z(\text{c.v.})/D(\%)]^2 \tag{4.5}$$

where $\text{c.v.} = s/\bar{x}$ is the coefficient of variation, and $D(\%) = D/\bar{x}$ expresses D as a proportion (note that $D(\%)$ is used as a proportion, NOT as a percentage). The utility of this approach is that we often have an idea of the coefficient of variation, but may not know what the mean is likely to be, so it is possible to set proportionate limits this way. The advance specification can then be, for example, that "I want 95% confidence limits no wider than ± 20 percent of the mean".

When sampling without replacement is important (i.e., when the sample is likely to be a significant proportion of the population, say, 5% or more) the above relationships serve to get an approximate value of the sample size needed, which can be labelled n_0 . The final estimate of n is then obtained from:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (4.6)$$

This estimate results from using a variance estimate that includes the finite population coefficient as previously described (Equation 4.1).

Example 4.2 Determining sample size

Suppose we want smaller (narrower) confidence limits, say $8 \leq \mu \leq 12$, in Example 4.1. Using Eq.(4.4) $D = 2$, and $n = [2(9)/2]^2 = 81$. If an appreciable fraction of the population is to be sampled, then the above result needs to be corrected by using Eq. (4.6). Assume $N = 100$, then we have $n = 81/(1 + 81/100) = 44.8$, which may be rounded to $n = 45$. Suppose it is required that $D(\%) = 0.1$ for an approximate 95% confidence interval of $\pm 20\%$. Then $D(\%) = 0.1 = 2(9)/10\sqrt{n}$, and $n = 324$, which exceeds the supposed population size of $N=100$. But Eq.(4.6) gives $n=76$.

4.6 Stratified random sampling

Almost invariably, ecologists have some advance information about populations that they wish to sample. This prior knowledge may well be one of the reasons for rejecting random sampling and substituting some sort of purposive selection, wherein one chooses sampling units that "look" to be representative or typical. There are, however, methods that take into account advance knowledge and at the same time provide the protection against bias that is given by random selection. One such method is to classify all of the population units into one of several strata (groups), and to then take random and independent samples in each such stratum.

The name, stratified sampling, comes from the close analogy to the layering effect seen in various circumstances, since we normally attempt to have the strata represent gradations in value of the random variable of interest. If one can do a fairly good job of segregating units by magnitude of the random variable under consideration, then it is apparent that the variability to be encountered in sampling within a particular stratum may be substantially reduced over that without stratification. Hence greater overall precision results for a particular total sample size.

Stratified sampling will require somewhat more advance effort than simple random sampling, but the usual result is that it turns out to be less extra effort than one might suppose. The method provides some side benefits in terms of better understanding of the material being studied, and of the nature of the problem dealt with. In some circumstances it may be that a portion of the sampling units in the population are very difficult to reach, or otherwise expensive to sample. In this case, stratification can specifically take

differences in cost into account, and provide alternative sampling schemes aimed at getting the most information for the effort expended.

The material here is devoted to an exposition of the basic technology in stratified sampling. Many additional features will be found in texts like that of Cochran (1977) and Thompson (1992). Here we will deal with such things as how to go about stratifying a population, the estimation procedures, various equations for obtaining variance estimates, and determining sample size.

4.7 Mechanics of stratification

Undoubtedly the most common ecological application of stratified sampling has to do with locating sampling plots or other measurement schemes on a map of a region of interest. The process of stratification is then intuitively obvious--one finds a way to break the map up into strata, each of which is composed of some large number of individual sampling units, usually plots of square or rectangular shape. It might be noted that the units in one particular stratum do not have to be contiguous--this is in fact one of the primary advantages of stratification. However, it is advantageous to keep units in a stratum more or less contiguous if the survey is designed for analytical purposes (e.g., to make comparisons between strata).

The basic process is just to assign units to strata according to the available prior information, seeking to get the units in any one stratum as much alike as possible in terms of the random variable or variables being studied. The next step is to assign serial numbers to every unit in each stratum. This does not, of course, require that someone write down all the numbers--all that is required is a trustworthy scheme for assigning and finding again any particular number. Often it will turn out to be simplest to delineate stratum boundaries with colored pencil and to note the starting and ending point of each row of units in a given section of one stratum by writing the corresponding numbers on the map. Sometimes large blocks can be counted as the equivalent number of sampling units, that is sampling units might be mil-acre (or meter-square) plots but the strata may be made up of larger units. A little practice soon settles the details for any particular set of conditions.

We use the following notation, which follows that of Cochran (1977) for convenience in referring to the much more complete description available there.

N_h ($h = 1, 2, \dots, L$) The number of units in the stratum. There are L strata in all, and $N = N_1 + N_2 + \dots + N_L$

n_h The number of units in the h^{th} stratum that are selected for enumeration (a random sample of n_h units from the h^{th} stratum). In most circumstances n_h should be at least 4.

y_{hi} ($i = 1, 2, \dots, n_h$) The observed value of the random variable (Y) on the i^{th} unit of the h^{th} stratum. There are n_h such observations taken from the h^{th} stratum.

$W_h = \frac{N_h}{N}$ The proportion of the total units present in the h^{th} stratum (N units in population; $N_1 + N_2 + \dots + N_L = N$).

$\bar{y}_h = \Sigma y_{hi} / n_h$ Average of the sample units from the h^{th} stratum.

$f_h = \frac{n_h}{N_h}$ The sampling fraction in the h^{th} stratum (i.e., the fraction of the units in the h^{th} stratum that are actually examined).

4.8 Estimates from a stratified sample

Since the W_h represent the proportion of the total population in the h^{th} stratum, they are logical weighting factors for estimating the overall population mean. The equation is:

$$\bar{y}_{st} = w_1 \bar{y}_1 + w_2 \bar{y}_2 + \dots + w_L \bar{y}_L \quad (4.7)$$

where \bar{y}_{st} refers to the mean of a stratified sample.

By the rule for variance of a linear combination of independent random variables (independent because of the random sampling in separate strata), we have the variance of the estimate as:

$$V(\bar{y}_{st}) = w_1^2 V(\bar{y}_1) + w_2^2 V(\bar{y}_2) + \dots + w_L^2 V(\bar{y}_L) \quad (4.8)$$

This result assumes that the sampling fraction (f_h) in each stratum is small enough to neglect the finite population correction. If the fpc is included, the equation for variance in the h^{th} stratum is:

$$V(\bar{y}_h) = \frac{S_h^2 (N_h - n_h)}{n_h N_h} = \frac{S_h^2}{n_h} [1 - f_h] \quad (4.9)$$

where S_h^2 is obtained from:

$$S_h^2 = \sum_{i=1}^{N_h} (y_{hi} - \bar{y})^2 / (N_h - 1) \quad (4.10)$$

These two equations can be combined to get the desired result:

$$V(\bar{y}_{st}) = \sum_{h=1}^L (W_h^2 S_h^2 / n_h) - \sum_{h=1}^L W_h S_h^2 / N \quad (4.11)$$

However, S_h^2 is not a quantity that can be determined from sampling, being the variance of the entire population in the h^{th} stratum. A logical procedure is to estimate it from:

$$s_h^2 = \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2 / (n_h - 1) \quad (4.12)$$

which is the usual variance estimate for a simple random sample. The last term in the equation for $V(\bar{y}_{st})$ is the finite population correction so if all of the f_h are small (say, less than 5%) this term can be dropped

4.9 Confidence limits

When sample sizes in the several strata are all substantial, the confidence interval takes on the same form as we have previously encountered it for simple random sampling, and can be written as:

$$\bar{y}_{st} \pm z s(\bar{y}_{st})$$

where z is the value from the unit normal curve corresponding to the confidence level wanted, for example $z = 1.96$ for $\alpha = 0.05$. We now use $s^2(\bar{y}_{st})$ or, in this case its square root, to denote that this is an estimate of the true variance given by Equation (4.11).

When sample sizes for individual strata are small, as is not uncommonly the case, there is a difficulty brought in by the fact that use of z corresponds to virtually knowing the true variance. As was noted earlier, samples of 30 or so give close enough estimates that one does not need be too concerned about an effect on the confidence limits. With only a few observations in one or more strata, however, the estimates of stratum variance (s_h^2) may not be so precise, and this situation would usually be handled by substituting a "t" value for the z value, that is, by making use of the t-distribution, which allows for uncertainty about the variance estimate. The trouble here is that we need to combine the several stratum variances to estimate $V(\bar{y}_{st})$, but it is not proper to average the various "t" values corresponding to stratum sample sizes (we would ordinarily look up a t-value with $n_h - 1$ degrees of freedom for each stratum). Cochran (1977:96) and Thompson (1992:106) give a rather complicated expression for calculating an "effective" number of degrees of freedom for use in this case.

Example 4.3 A caribou census

An example that incorporates nearly all of the basic elements in stratified random sampling is furnished by an aerial census of Alaskan caribou. Details will be found in a paper by Siniff and Skoog (1964). We here extract that part which is appropriate for illustration. Six strata were selected on the basis of preliminary observations of relative caribou abundance, and delineated on detailed maps. Sampling units were four-square-mile blocks, and each such unit within each stratum was assigned a number. There were no advance estimates of within-stratum variances available, so the stratum standard deviations (S_h) were assumed proportional to the preliminary rough estimates of population levels in the stratum. This provided the following data for allocation according to Equation (4.14):

Stratum	N_h	W_h	S_h	$\frac{W_h S_h}{\sum W_h S_h}$	$n_h(\text{opt.})$	$n_h(\text{actual})$
A	400	.572	3000	.428	96	98
B	30	.043	2000	.022	5	10
C	61	.087	9000	.195	44	37
D	18	.026	2000	.013	3	6
E	70	.100	12000	.299	67	39
F	120	.172	1000	.043	10	21
	699	1.000	29000	1.000	225	211

The "optimum" allocation was based on the total number of sample units (225) that the investigators believed could be surveyed in the time available. The actual allocation amounted to "hedging" against uncertainty about the likely values of S_h . Thus there were several strata (B, D and F) where the supposed optimum allocation called for rather small samples, so these were increased at the expense of strata (C and E) where the optimum plan called for censusing a substantial fraction of the units in the stratum. Survey results were as follows (N_h , W_h were as used above, and n_h as given in the last column above):

<u>Stratum</u>	\bar{y}_h	s_h^2	$\frac{W_h^2 s_h^2}{n_h}$	$W_h s_h^2$	$\bar{y}_h W_h$
A	24.1	5,575	18.613	3,189	13.79
B	25.6	4,064	.751	175	1.10
C	267.6	347,556	71.098	30,237	23.28
D	179.0	22,798	2.569	593	4.65
E	293.7	123,578	31.687	12,358	29.37
F	33.2	9,795	13.800	1,685	5.71
	-	-	138.518	48,237	77.90

From Equation (4.7), \bar{y}_{st} is the sum of the last column above, or 77.9 caribou per four-square-mile unit. This is readily converted to a total for the area surveyed by multiplying by the total number of units, giving $(77.9)(699) = 54,450$ caribou. The variance estimate (Equation (4.11) is:

$$v(\bar{y}_{st}) = \sum_{h=1}^L (W_h^2 S_h^2)/n_h - \sum_{h=1}^L W_h S_h^2/N = 138.518 - (48,237/699) = 69.51$$

Confidence limits on the mean may be obtained by assuming z to be equal to 2 (or 1.96 to be exact, for $\alpha = 0.05$), since rather substantial samples are involved here, in most strata, giving:

$\bar{y}_{st} = \pm 2(69.51)^{1/2}$ or 77.9 ± 16.7 caribou per four-square-mile unit. For total caribou on the study area, we estimate the variance as:

$$v(y_{tot}) = N^2 v(\bar{y}_{st}) = (699)^2(69.51) = 33,962,655$$

and limits are:

$$54,540 \pm 2(33,962,655)^{1/2} \quad \text{or} \quad 42,885 \leq x_{tot} \leq 66,195.$$

Notice that s_h^2 increases with increasing y_h in the table above. The investigators plotted $\log_{10}(s_h^2)$ against $\log_{10} \bar{y}_h$ (Siniff and Skoog, 1964:398) and obtained a regression relationship:

$$y = 1.63 + 1.42 x$$

where $y = \log_{10}(s_h^2)$ and $x = \log_{10} \bar{y}_h$. This is equivalent to the relationship:

$$s_h^2 = 42.66(\bar{y}_h)^{1.42}$$

which might be used to estimate variances in planning similar surveys. However, it is important to remember that size of the sampling unit (4 sq. mi. in this case) will affect such a relationship.

Example 4.4 A mortality survey

A Michigan study of over-winter losses of whitetailed deer (Whitlock and Eberhardt, 1956) provides an example where the finite population correction is negligible. In this case, nearly 19,000 square miles (all of the northern lower peninsula of Michigan) were classified into five strata on the basis of estimates made by field biologists. The primary units were half-sections (one-half square mile), but these were subsampled in the actual search by using a strip 88 yards wide laid out as a rectangular course 1/2 mile long and 1/4 mile wide. Width of the strip was based on use of four-man teams, with each individual responsible for searching a 22 yard wide interval. Various complications were involved in the design inasmuch as it was necessary to consider prospects for missing dead deer on the strip, the number of men available in various locations (and transportation), the necessity for one man to act as compass-man, need for a biologist in each crew, and so on.

Advance data from a previous survey of an area of high mortality suggested that the coefficient of variation (s/\bar{x}) might be about 1.30, so estimates of S_h were obtained by multiplying 1.3 times an estimated number of deer to be found on each plot (these guesses were made in the process of setting up strata).

It was decided that about 110 plots could be surveyed with available manpower so the allocation was devised as follows:

<u>Stratum</u>	Expected losses per sq. mile	Expected losses expressed as dead deer/plot	Square miles in stratum	W_h	Estimate of S_h	Preliminary allocation
I	20+	3.75	408	.0220	4.88	23
II	10-20	1.88	1,048	.0566	2.44	29
III	5-10	.94	2,293.5	.1240	1.22	32
IV	1-5	.31	5,567.5	.3010	.40	25
V	0-1	.01	9,181.5	.4964	.01	1
			18,498.5	1.0000		110

The allocation again followed Equation (4.14) but four plots were added to stratum V, giving 114 in all, of which 113 were actually searched (one plot was completely flooded at the time of the survey). Survey results were:

<u>Stratum</u>	n_h	s_h	\bar{y}_h	Contribution to $V(\bar{y}_{st})$	Coefficient of variation	s_h^2 / \bar{y}_h
I	23	2.146	1.826	.000097	1.18	2.52
II	29	1.082	.621	.000129	1.74	1.88
III	31	.724	.484	.000260	1.50	1.08
IV	25	.541	.280	.001062	1.93	1.04
V	5	--	0.00	.00	--	--
	113			.001548		

In this instance, only very small fractions of each stratum were searched so Equation (4.11) reduces (by dropping the right-hand term) to:

$$v(\bar{y}_{st}) = \sum_{h=1}^5 (W_h^2 S_h^2) / n_h$$

and the individual terms are listed under the heading "Contribution to $v(\bar{y}_{st})$ " so that one can see in which stratum most of the variability

turns up. Comparing the expected losses and the \bar{y}_h , it becomes apparent that the over-estimates were largely in strata I and II, which was not especially surprising since the winter turned out to be milder than anticipated when the survey was planned, and starvation losses were correspondingly lower (major starvation areas nearly all were in strata I and II).

The coefficients of variation in the above table show the advance estimate to be somewhat low. The last column of the table gives s_h^2 / \bar{y}_h

which is the "index of dispersion" and is unity (within sampling or "chance" errors) for a Poisson distribution. This suggests that such a distribution (i.e., wholly random dispersal of dead deer) may apply to strata III and IV, in which case the S_h for allocation might simply have been taken as equal to the square roots of the expected numbers of deer per plot.

4.10 Allocating the sample to strata

We have thus far gotten the cart before the horse, having considered how to analyze sample results without considering how the total sample ought to be distributed over the strata. Two kinds of allocation are in common use. The first is perhaps what one would expect to do without any advance information about the variability in various strata, that is, distribute the sample in proportion to the size of the strata ("proportional allocation"). This is also known as a self-weighting sample, since fractions going into each stratum will be equal to W_h , so that a simple mean of all of the sample results will be equal to the weighted mean previously given. In this case we have $n_h/N_h = f = n/N$ so that the sampling fraction is the same in all strata. This leads to a simpler expression for the variance:

$$V(\bar{y}_{st}) = \frac{1-f}{n} \sum_{h=1}^L W_h S_h^2 \quad (4.13)$$

and we again have to substitute sample estimates for S_h^2 .

Proportional allocation is easy to accomplish and to analyze, but often is not a very efficient way to use sampling resources. In most ecological work it turns out that the variance and mean tend to increase together, so that the strata likely will have rather different variances, and proportional allocation will then undersample some strata and oversample others. An allocation which allows for the effect of differences in stratum variances is the scheme called "optimal allocation". This method can be shown to minimize $V(\bar{y}_{st})$ for a fixed n . Optimum allocation is given by the following relationship:

$$n_h = n W_h S_h / \sum_{h=1}^L W_h S_h \quad (4.14)$$

Of course use of the formula demands at least a guess at the S_h . In many studies, there will be some preliminary information about the magnitude of variances to be encountered, quite often in the form of coefficients of variation, which may be applied to the expected stratum means to get an estimate of stratum standard deviations. It also turns out that this kind of allocation is not too sensitive to errors in advance estimates of S_h so one can usually expect to do a better job with this method so long as the stratum variances do differ appreciably and the guessed values of S_h are in the right "ballpark". In many natural populations the stratum with the lowest mean can be expected to have roughly a Poisson distribution of individuals (assuming the purpose of the survey is to estimate total individuals) so the investigator can set that variance equal to the expected mean density, and go on from there on the basis of any information about how variability increases with the

means of the remaining strata.

One further feature of allocation worth considering here is the circumstance where sampling costs differ among strata. Perhaps the simplest assumption is that total cost of the sampling can be written as:

$$\text{cost} = c_0 + \sum_{h=1}^L c_h n_h \quad (4.15)$$

where c_h represents the cost of measuring each sample unit in stratum h , the c_h are not all equal, and c_0 is a fixed or "overhead" cost. In this case, Cochran (1977:97) shows that the allocation should be:

$$n_h = \frac{n W_h S_h}{\sum_{h=1}^L W_h S_h / \sqrt{c_h}} \quad (4.16)$$

so that the number of samples in a stratum depends on the stratum size, its variability, and cost of sampling. One takes more samples in large and variable strata, but also increases sample size if sampling is cheap in the stratum. This kind of allocation can be rather useful in dealing with sampling problems where either access or measurement may be quite difficult for part of the population. It is worth noting that other kinds of cost functions might be obtained from knowledge of the sampling problem, and special allocations then devised. Cochran (1977) discusses "cost functions" for various circumstances.

Example 4.5 A deer population estimate

Counts of "pellet-groups" have been used to estimate deer populations for many years. Daily defecation rates are remarkably constant (about 13 groups per day) and over-winter accumulations of pellets can be identified by the underlying mat of leaves dropped the previous fall. There is thus a straightforward conversion from numbers of pellet-groups to "deer-days" which in turn can be converted to average population levels for the over-winter season. Stratified random sampling has been used to conduct such surveys in northern Michigan for more than 25 years. About 35,000 square miles are surveyed, requiring on the order of 500 man-days of effort. Some nine separate areas (Game Management Districts) are surveyed independently. An example for one such are (District 7 in 1962; Ryel 1971:131) appears in the following table:

<u>Stratum</u>	<u>Area (sq.mi.)</u>	<u>Prop.(W_h)</u>	n_h	\bar{y}_h	$\frac{W_h^2 s_h^2}{n_h}$
I	190	.0541	9	65.22	1.7568
II	425	.1211	12	29.25	1.2649
III	1544	.4399	34	15.35	1.7803
IV	1144	.3259	10	10.70	1.7748
V	207	.0590	1	0.0	--
	3510	1.0000	66	-	6.5768.

The overall weighted mean number of groups per sampling unit was

$\bar{y}_{st} = 17.31$ with two standard errors being 29.6 percent of that estimate. With the very large areas involved, an appreciable amount of time is expended in traveling to the sampling units. Since experience shows that the individual plots should not be very large (to avoid missing groups in the counts), a cluster sample is used at each site, comprised of eight individual plots arranged along a half-mile line. For convenience, square miles (sections) serve as the primary sampling unit, with a random starting place and distance from the boundary used to locate each systematically arranged cluster of plots.

To reduce the effort required to plan and execute these large-scale surveys (150 to 200 people may be involved annually), the same plots have been used for a number of years in succession. This makes it possible for the field men to plan their work efficiently, since they know the plot location well in advance and can anticipate just when the plots will be accessible (and free of snow). Ryel (1971:222) calculated the optimum allocation for a number of years. Results for the District used as an illustration above are:

<u>Stratum</u>	<u>Actual allocation</u>	<u>Calculated optimum allocation</u>					
		<u>1959</u>	<u>1960</u>	<u>1961</u>	<u>1962</u>	<u>1963</u>	<u>1964</u>
I	9	18	14	6	13	4	6
II	12	11	14	21	13	19	18
III	34	27	29	21	26	32	32
IV	10	9	9	18	14	11	10

It can thus be seen that the original allocation was, on the average, quite satisfactory. Two kinds of factors may affect these results. One is that the distribution of deer may change somewhat from year to year, in consequence of winter weather conditions. Another is that variances for each stratum are estimates, and thus will vary somewhat due to chance alone.

4.11 Further remarks on stratified sampling

Cochran (1977), Thompson (1992), and other texts on survey sampling supply a good deal of auxiliary information on methods and techniques for various special cases. A few points that are examined in more detail in those references are summarized here:

(1) Gains in stratified sampling for the estimation of a proportion are usually not sizable unless the proportion (P) varies sharply from stratum to stratum, and in most cases, proportional allocation is preferable.

(2) Many surveys are designed to measure more than one random variable, whereupon the question of allocation gets complicated. An initial approach is to calculate allocation for the variables of main interest separately and determine whether the several allocations differ appreciably. If so, then it may be possible to devise some sort of cost function to help in a decision. If

there are two variables of main interest, then there are some handy schemes for handling the two in a single allocation.

(3) We have not considered how to determine either the number of strata nor stratum boundaries. Most studies of natural populations will involve the location of strata on maps, and the quantity of advance information on the population will usually be such that the number of strata probably will not be less than three nor more than six. A few trial efforts at laying out strata will usually resolve most questions of boundaries. Probably the major difficulty comes up if large areas are to be covered, so that there are a number of people involved, each having rather good local knowledge. Then the principal job turns out to be in getting individuals to agree on what constitute definite strata, and how they should match at the junction of two districts where different people are locally "expert" on the subject matter. Some one person usually has to umpire the decisions, and this can perhaps be done after individuals have made up maps reflecting their knowledge.

(4) Sometimes it is possible to make use of stratification after a simple random sample has been taken. It must be emphasized that stratification cannot legitimately be undertaken on the basis of examining the sample results, but it may turn out that it is not possible to assign individual units to strata until after they have been surveyed, that is, the total number of units in each stratum may be known in advance, but the stratum to which a particular sample unit belongs cannot be determined until the measurement is made.

(5) Most experience with natural populations shows that variability increases with the mean. This is fairly sound grounds for recommending that "optimum" allocation always be carefully considered before selecting one of the other possibilities.

4.12 Ratio Estimation

The main results for ratio estimation require that the population total of an auxiliary variate (X) be known, and the correlation between X and the variable of main interest (Y) is used along with the known total to obtain an estimate of either the mean of Y or its total with greater precision than may be obtained from simple random sampling of Y alone. So far, ratio and regression methods have been little used in ecology and resource management surveys, partially perhaps because of a lack of suitable correlated variables with known totals, but also because many investigators are not familiar with these methods.

In the usual notation, X is used to represent the known population total. Since we have been using X to represent a random variable, X_T will denote the population total here. The ratio estimate is:

$$\hat{Y}_R = \frac{\sum y_i}{\sum x_i} X_T = \frac{\bar{y}}{\bar{x}} X_T$$

as an estimator of the population total for Y . The mean value of Y is estimated by replacing the population total (X_T) by the mean above.

The population ratio is estimated by:

$$\hat{R} = \frac{\sum y_i}{\sum x_i}$$

If interest is principally in the population ratio, then it is not necessary to know X_T .

An important application of ratio methods is worth mentioning here in order to provide an illustration of the nature of the above relationships. This is the use of "strip transects" (discussed in more detail in Chap.5) on irregularly-shaped areas. A strip transect is just a long, narrow plot extending completely across a study area. For present purposes, we assume that all of the objects of interest are counted on each of a number of strips. Each such transect then constitutes a sample plot. If the region under study is rectangular in shape, then each sample plot will have the same area, and no adjustment for transect length is needed. However, in most practical situations, study areas will be irregular in shape. Strip transects across such a site will thus have different individual areas, presenting a problem in the analysis of the data, since plot size is now also a random variable.

It is true that a simple random sample of strips will provide an unbiased estimate of the total number of objects on the study area. The appropriate random variable is the total number on each strip, and the calculations proceed as previously described for simple random sampling of a finite population (the total number of possible sample strips). However, such a theoretically correct result is of almost no practical interest in dealing with natural populations, just because such populations exhibit high variability even with efficient methods of sampling. We thus cannot afford to bring in any further variability. Ratio methods can conveniently be used to resolve the problem simply by regarding the area of each sample strip as X_i , so that X_T is the total area of the study region, and letting Y_i represent the total number of objects on each sample plot. We then have that \hat{R} estimates the average density (number per unit area) observed in the sample, and $\hat{Y}R$ is an estimate of the total number of objects on the entire study area.

The ratio estimate is biased, but the bias is considered unimportant for large samples. In this case, a rule of thumb is n of at least 30, and the coefficients of variation of the means of X and Y should both be less than 0.10 (Cochran, 1977:153). Stratification and ratio estimation may serve roughly the same purposes, and it is likely that an effective stratification could be obtained through the use of the auxiliary variable X . Thus in the example given above, one could stratify the study area into blocks such that the length of potential sample transect strips is about constant in each stratum. However, the ratio method provides a "natural" approach in this instance, and is thus the appropriate choice.

The ratio estimate effectively assumes the relationship between Y and X to be $Y = RX + e$, where e represents an "error" component and R is an unknown constant. In some instances it may not be reasonable to assume that the relationship goes through the origin, so that a regression estimate is appropriate. This method is also biased, so that large samples are generally recommended. Details are available in Cochran (1977), Thompson (1992), and many other texts on sampling. Before undertaking to use the ratio or

regression techniques, an investigator should have some preliminary observations (or good general knowledge) that indicate a relationship between the variable of primary interest (y) and an auxiliary variable (x). A first step is to plot the data and to note whether the regression line clearly does not go through the origin. If this is the case, then it is advisable to look into regression estimation, rather than using ratio methods. Occasionally there may be more than one useful auxiliary variable and it is then possible to use multiple regression.

4.13 Variance of ratio estimates

An estimate of the variance of a ratio is given by:

$$V(\hat{R}) \doteq \frac{1-f}{n\bar{X}^2} \left[\sum_{i=1}^N (Y_i - \hat{R} X_i)^2 / (N-1) \right] \quad (4.17)$$

Here f again represents the finite population correction, and may be neglected if n/N is less than about 5 to 10 percent. N is the total number of units in the population, n , the number in the sample, and \bar{X} the population mean of the auxiliary variate. Note that the summation runs over the entire population, so that this is an approximation to the "true" variance, and it will in turn have to be estimated by a quantity that can be calculated from a sample; that is, we replace the quantity in the right-hand brackets by sample data, getting:

$$s^2(\hat{R}) \doteq \frac{1-f}{n\bar{X}^2} \left[\sum_{i=1}^n (y_i - \hat{R} x_i)^2 / (n-1) \right] \quad (4.18)$$

When interest is in the mean or total of Y , the estimates are as given before; and variances can be calculated from Equation (4.17) by recalling the rule that $V(aR) = a^2V(R)$, where the constant a is now either \bar{X} or $N\bar{X}$, since both of these quantities are assumed known, and thus play the part of constants. Calculation of an estimate of $V(Y_T)$ is easier in the following equivalent form:

$$s^2(\hat{Y}_R) \doteq \frac{N(N-n)}{n(n-1)} \left[\sum y_i^2 + \hat{R} \sum x_i^2 - 2\hat{R} \sum y_i x_i \right] \quad (4.19)$$

Note that this is the variance for estimating a total.

Since it is advisable to check that the coefficients of variation of the means of Y and X are less than 0.10, another form for calculation of variability is:

$$[c.v.(\hat{Y}_R)]^2 = \frac{s^2(\hat{Y}_R)}{\hat{Y}_R^2} = \frac{1-f}{n} [c_{yy} + c_{xx} - 2c_{yx}] \quad (4.20)$$

where c_{xx} , c_{yy} , and c_{yx} are the coefficients of variation of y , x and the analogously defined cross-product term:

$$c_{yx} = \frac{\Sigma yx - n\bar{y}\bar{x}}{(n-1)\bar{y}\bar{x}}$$

Readers who refer to Cochran (1977) should note that he used the coefficient of variation of the mean, e.g., c_{yy}/n .

The squared coefficient of variation of Y_R is often termed the relative variance, and can be used to calculate variances of any of the three estimates of interest (the coefficient of variation, being a relative quantity, has the same value for \hat{Y}_R , \hat{R} , or \hat{R}_T).

Confidence limits can be obtained as before:

$$\hat{Y}_T \pm zs(\hat{Y}_T) \quad \text{or} \quad \hat{R} \pm zs(\hat{R})$$

Example 4.6 Ratio corrections for variable plot size

A numerical example of corrections for different lengths of a strip-transect is given by Norton-Griffiths (1975). The data are those from an aerial survey for several species of African "game". Only wildebeest are considered here. The data are as follows:

<u>Transect</u>	<u>x_i</u> <u>Area (km²)</u>	<u>y_i</u> <u>No. counted</u>
1	8.2	58
2	13.7	44
3	25.8	175
4	25.2	141
5	21.9	151
6	20.9	144
7	23.0	131
8	19.2	135
9	21.4	104
10	17.5	111
11	19.2	130
12	20.8	136
Totals	236.8	1460

The total area of the study region was 2829 km², so the population estimate is:

$$\hat{Y}_R = \frac{\Sigma y_i}{\Sigma x_i} X_T = \frac{1460}{236.8} 2,829 = 17,440 \text{ wildebeest.}$$

There were 126 possible strips in the area, so that $N = 126$, $n = 12$, and calculations from Eq (4.19) are:

$$s^2(\hat{Y}_R) = \frac{N(N-n)}{n(n-1)} [\Sigma y_i^2 + \hat{R} \Sigma x_i^2 - 2 \hat{R} \Sigma y_i x_i]$$

$$= [126(114)/12(11)][193,262 + (6.16)^2 4,935 - 2(6.16)30,561]$$

$$= 436,580.$$

The standard error is $(436,580)^{1/2} = 661$, which is quite small compared to the estimate.

It may be noted that the sample size (12) is a good deal less than the 30 recommended as a rule of thumb for using ratio estimation. However, this very likely is an instance where the ratio estimate is nearly optimal, i.e., the relationship goes through the origin, and the variance of the counts likely increases with the area of the transects. Hence, it seems quite reasonable to neglect the $n = 30$ rule. Density per unit area is estimated by:

$$\hat{R} = 1460/236.8 = 6.16 \text{ wildebeest per km}^2.$$

4.14 Double sampling

The major problem with ratio estimation in ecological studies is just that there are various situations where the method is potentially useful, but a total for the auxiliary variable is not known exactly. Many of these situations do not seem to fit neatly into the present methodology of survey sampling, but it does seem that double sampling comes close enough to provide a useful framework for examining the problems and a useful starting place for much-needed research. The basic idea is just that of the ratio estimation scheme. We have a random variable of primary interest (Y) and an auxiliary variable (X) known to be well-correlated with Y . The missing item is a known total for this auxiliary variable (X_T).

In the instances of interest here, measurements of the auxiliary variable (X) are either very inexpensive to obtain, or are readily available for a large sample taken over the study region. A convenient example is that used to describe ratio estimation; the use of strip transects. We now suppose that the total area of the region under study is not known. If the area is mapped, then it is obviously an inexpensive process to make a large number of measurements of the lengths of potential transect lines from the map. One can thus come very close to estimating the total area (X_T) by working with the map. If we denote this estimated total as X'_T , then double sampling proceeds in just the same manner as ratio estimation, i.e.,

$$\hat{Y}_R = \frac{\sum y_i}{\sum x_i} X'_T$$

but it is now necessary to make some allowance in variance estimation for the fact that the total (X'_T) of the auxiliary variable is not known exactly. Eberhardt and Simmons (1987) conducted some monte carlo studies to suggest when double sampling might still be useful under this limitation.

If the study region is mapped, there are usually better ways to measure the total area (e.g., by planimetry). However, various nontrivial examples can be considered. The survey may be concerned only with a particular cover type, which is not mapped. If the work on the actual sample transect is quite time-consuming, then it may be well worthwhile to measure only the width of the cover-type on a large number of "auxiliary" transects. These widths then provide an estimate of X_T .

Another situation where double-sampling may be useful is where detailed measurements need to be made on individual plots by some time-consuming process. One example is in estimating total oven-dry biomass of, say, nonwoody vegetation. The time required for clipping, drying, and weighing vegetation severely limits the number of plots that can be so dealt with. Double-sampling might well be utilized by using stem counts as an auxiliary variable, since this can be done on a rather large sample of plots at low cost. A similar prospect exists when chemical analyses are to be done on vegetation, but in this case it may be desirable to use weights on a large sample of plots as the auxiliary variable.

An essential feature of these examples is that accurate measurements of the auxiliary variable can be made in each instance. This appears to be the basis for the present theory of double sampling as given, for example, by Cochran (1977: Ch.12). Unfortunately, there are a great many very useful potential applications in ecological studies that do not seem to quite "fit" the existing theory. These are situations where the auxiliary variable is an estimate of some kind, and is subject to either sampling error, bias, or both. The biomass of vegetation example provides a convenient case. Rather than stem counts, the investigator may choose to use an ocular estimate of biomass on a large sample of plots as an auxiliary variable. With some experience (best gained by guessing weights on a sample of plots and then clipping and weighing), he may become very proficient at visual estimation. The problem now is that the auxiliary variable is subject both to the "chance" errors inherent in visual estimation and to any persistent tendency to consistently overestimate or underestimate.

Another illustration may be taken from aerial censusing of animals. Practically all of the available experience shows that aerial observers tend to miss a substantial fraction of the animals on a sample unit (very often a strip transect). Nonetheless, since aerial surveys can be relatively inexpensive, efforts may be made to "calibrate" the surveys by using some accurate method to enumerate the number of animals actually on a subsample of the plots surveyed from the air. If it can be supposed that these "ground-truth" counts are truly without error, then it can be argued that the requirements of double-sampling are met. The aerial survey now provides the auxiliary variable (X), while the ground count provides the accurate census (Y) that is wanted. However, the auxiliary variable (aerial count) is clearly going to be subject to sampling errors, due to a large variety of causes. Hence we no longer have quite the same situation as when the auxiliary variable can be measured without error. It may be feasible to completely survey the study area from the air. However, this is still not a known total, as a repeat survey flown under identical conditions will without much question yield a different total count.

Many readers will have recognized another problem that was passed by above. This is that the "accurate" measurement (Y) is seldom achievable in census work. Usually the best that can be managed is an estimate that is believed to be unbiased, but is clearly subject to sampling error. We thus have both Y and X subject to sampling errors. This circumstance may bring in some major problems in statistical analysis. These problems are particularly difficult in regression analysis, and remain unresolved for a number of circumstances of importance to ecologists and biologists. Ricker (1973) reviewed the situation for problems in fisheries research and management.

There is thus a need to exercise some caution in applying double sampling in situations where the auxiliary variable is subject to sampling errors, particularly when regression equations are used. In many practical applications in ecology it seems that the ratio approach may be quite satisfactory so we will usually rely on it here.

If it is clear that the relationship does not pass through the origin and if the variance appears relatively constant around a regression line, then it is likely that the regression method should be used. However, in the many cases where it is necessary to assume sampling errors in the auxiliary variable, the usual elementary textbook test for significance of the intercept cannot be trusted. Hence it may be best to depend on judgements as to the nature of the relationship and the pattern of variability in choosing between ratio and regression methods.

4.15 Cluster sampling and subsampling

Cluster samples are likely to be useful in field studies whenever the item of interest is primarily associated with some natural sampling unit. An example might be some species of insect found only on a particular species of plant. Any interest in enumerating the insects, or in studying some other measurement, such as the percent of insects parasitized, requires attention to the fact that they come in clusters. In point of fact, this distinction is often ignored in practice, and it can be safely said that measures of variability obtained without considering the clustering effect will usually be very seriously underestimated. Of course, in the example here described one might reasonably use a ratio estimate, counting the number of plants and sampling some part of them for insect abundance.

In some cases it is possible to deal with clusters that are all comprised of the same number of individual sampling units. This is a natural way to approach large-scale area samples, where the "primary sampling unit" may be taken to be a square mile (section). One may want to use much smaller plots (square-meter or 0.01 ha, perhaps) for the actual measurements, but to enumerate the variable of interest on several such plots in each square mile in the sample. One approach is then to draw a random sample of n square miles from the overall area, and to locate m plots (the subsample) in each of the selected primary units. This is usually termed two-stage sampling. An important consideration in such schemes is determining how many subsamples (m) and the total number (n) of primary units to take to minimize the overall variance (or maximize precision) for a fixed over-all cost.

We will not attempt to detail the procedures for optimum use of subsampling methods, but it is worth mentioning one scheme for calculating the overall variance of an estimate, and thus confidence limits. This is just to use the subsample results for each primary sampling unit to estimate a total for that unit. That is, if there are m plots in each unit, one just obtains the total for those m plots and multiplies it by the reciprocal of the sampling fraction to get an estimate for the primary unit. The primary unit totals can then be used directly as random variables to compute a variance for the survey total. This variance will reflect both components of variability -- that for subsampling (within primary units) and that for differences among primary sampling units. What one loses, of course, is any information on the

optimum subsampling rate. The same scheme can be used when the clusters (primary units) are of different sizes (i.e., contain different numbers of sampling units, as, for example trees in woodlots). The shortcoming here is that if the clusters vary considerably in size, that difference will contribute greatly to the overall variance. This point is most important if there is additional information on the cluster sizes in the population, but as was mentioned above, one might then be able to use ratio methods.

Subsampling schemes are often conveniently used to combine systematic and random sampling. In the example above of randomly selected sections (square miles) which are then subsampled, it is often possible to reduce the labor involved if the primary units (sections) are selected at random (usually in a stratified sampling plan) and a series of plots located systematically along a transect within each section as subsamples. It is highly desirable that the transect starting points be randomly selected to avoid any bias due to edge effects or such things as old fencelines in the sections.

Subsampling schemes can involve several stages, and various complexities of estimation. One might for example, use a stratified random sample of square miles, locate subsampling plots in each randomly drawn section, and then elect to examine only a random sample of individual plants on each plot for the variable of interest, which might in turn involve measurements subject to error. Obviously, the statistical analysis of such data can be quite complicated. One way to simplify matters a great deal is to resort to jackknifing or bootstrapping.

Sampling in two (or more) stages is also worth considering when there is uncertainty about the accuracy of the method for making measurements, as is so often the case in estimating the abundance of animal populations. It is usually the case that population density will vary considerably over large areas, and the investigator may have a reasonably good notion of how density varies with habitat and so on (or this may be a major item of interest). It is then logical to use a stratified random sampling scheme to locate primary sampling units on which the actual measurements of density will be attempted. This does not, of course, reduce any uncertainty in the actual measurement method, but it does keep the area differences from compounding matters.

Example 4.7 A cluster sampling example

One simple example of cluster sampling was mentioned in Example 4.5 (stratified sampling). The "primary sampling units" (square miles) were selected at random, and then subsampled with a cluster of eight small plots. All that is needed for analysis of the resulting data is just to multiply the total for the eight subsamples by a "raising factor" or "blow-up factor", which is simply the reciprocal of the sampling rate. In the example used, the individual plots were each $1/50$ of an acre, hence the necessary adjustment factor is: $640/(8/50) = 4000$. Once this is done, the remaining analysis proceeds as though no subsampling had taken place. Skeptics may need to do a little algebraic manipulation at this point. When subsampling rates are not constant, things become somewhat more complicated, and a sample survey text should be consulted for details. However, if the subsampling rate does not vary greatly, the same procedures can be used without elaboration. All that happens is that one overestimates the variance, in most situations. But if the subsampling rate varies considerably and/or is related to size of

the primary sampling unit, then by all means consult a textbook on sampling or a statistician.

Supposing constant size of the primary sampling unit, and a constant subsampling unit (the case most likely in ordinary applications) the main question to be settled is "What is the best subsampling rate?". As usual, answers depend on relative costs. That is, a particular effort (hence cost) is required to survey an individual sampling unit (i.e., one plot in the example), while a separate cost is engendered by the time and travel going from one primary sampling unit to the next. For a given total expenditure for the entire survey, the optimum subsampling rate is that which minimizes the overall variance given the above two costs

Since natural populations exhibit a somewhat frustrating tendency for variances to change nonlinearly with size of the sampling unit (plot size), a simple equation for subsampling rate is not available. What's really needed is a "variance law", i.e., a relationship between plot size and variance. To obtain such a relationship, one has to run a special study using several plot sizes. Then it becomes possible to incorporate costs and get on with the business at hand by consulting Cochran (1977, Ch. 9). As we noted earlier, the kind of measurement (weights, counts, etc.) and the organism under study influence the "variance law" substantially. Hence there are two choices open at this point. One is to run a fairly expensive preliminary field study, and thus to manufacture your own "variance law". The second choice is to resort to the literature in the appropriate field, seeking papers in which several different plot sizes have been used. A number of references along these lines appear in Eberhardt (1978a). However, it is clear that this is an area needing rather more research attention in ecology.

Example 4.8 Cluster sampling involving proportions

One of the commonest errors in the ecological literature is an uncritical acceptance of the binomial distribution as an appropriate model for analysis of proportions in data collected in clusters. It is the appropriate model if, and only if, a simple random sample of individuals can be obtained. In practical problems one almost always collects observations as clusters. When this is the case, the clustering effect must be taken into account in order to obtain a meaningful variance. Very rarely do we encounter a population so well mixed that clusters are indeed equivalent to simple random samples, so that such an example is likely to be more of a curiosity than anything else.

The simplest way to deal with cluster sampling for proportions is to treat the individual observations as random variables. In this instance, the appropriate form of the ratio estimator is:

$$\bar{p} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i}$$

where y_i denotes the number of individuals in the i^{th} cluster possessing the attribute of interest, and x_i is the total number of individuals in the i^{th} cluster, while n is the number of clusters.

The appropriate variance estimate here is (Cochran 1977:65):

$$V(\bar{p}) = \frac{\sum(p_i - \bar{p})^2}{n(n-1)}$$

where $p_i = y_i/x_i$, i.e., the observed proportion in the i^{th} cluster. (We here neglect the finite population correction which can be inserted as a multiplier $(1-f)$ if needed).

An interesting set of data to illustrate behavior of proportions in clusters comes from a paper by Johnson and Chapman (1968). This was a study to estimate the number of fur seal pups on a "rookery" on the Pribilof Islands, off Alaska. A large sample (4,965) of pups were marked (in groups) and then clusters of 100 were examined (for the proportion marked) at randomly selected sampling stations. The estimate of the total number of pups on the rookery was obtained from

$$\hat{N}_1 = \frac{M}{(\bar{p})}$$

where N is the population estimate, M is the number marked (4,965) and \bar{p} is the mean proportion marked, calculated as in the above example.

Two ways of estimating the variance were used. One is based on the "delta method", and is:

$$V(\hat{N}_1) = \frac{M^2 v(\bar{p})}{\bar{p}^4}$$

where $v(\bar{p})$ is obtained as in the above example. The second method is that of "interpenetrating" sampling, in which the sample is subdivided randomly into a number of subsamples. A separate estimate of the population size \hat{N}_i is made from each subsample and these are then averaged for the final estimate, i.e.:

$$\hat{N}_2 = \frac{1}{r} \sum_{i=1}^r \hat{N}_i$$

and:

$$v(\hat{N}_2) = \sum \frac{(\hat{N}_i - N)^2}{r(r-1)}$$

It should be noted that the two estimates of the total population will not necessarily be identical, nor will the variance estimates be the same for the two methods.

The observed numbers of marked pups in clusters of 100 (recorded on two sampling dates) were:

August 26, 1961 (25 samples)	2, 0, 1, 6, 4, 33, 62, 49, 55, 38, 52, 77, 85, 54, 27, 17, 3, 3, 3, 2, 2, 1, 0, 0, 4.
August 28, 1961 (58 samples)	0, 0, 0, 4, 0, 0, 0, 12, 4, 8, 60, 48, 72, 72, 76, 80, 56, 44, 50, 56, 56, 28, 60, 36, 44, 44, 28, 52, 72, 28, 72, 60, 60, 84, 76, 52, 84, 48, 52, 60, 40, 12, 8, 12, 4, 8, 44, 16, 0, 8, 0, 0, 4, 12, 8, 0, 0, 0.

The interpenetrating or replicated samples were defined as follows:

Subsamples 1, 2, 3: Every third observation of August 26, beginning with observations 1, 2, 3, respectively.

Subsamples 4-10: Every seventh observation beginning with observations 1, 2, 3, 4, 5, 6, 7, respectively.

Since there were 25 observations on 26 August, this procedure yields subsamples of size 9, 8, and 8, respectively, while the 58 observations on 28 August yield two sets of size 9 and 5 of size 8. These data lead to the following estimates for the interpenetrating sampling:

Subsample N_i

1	20,497
2	24,219
3	20,060
4	17,455
5	16,674
6	17,732
7	14,391
8	12,490
9	13,066
10	14,821

Total	171,405
-------	---------

Averaging gives $\hat{N}_2 = 17,140$ with $v(\hat{N}_2) = 1,353,000$, while $\hat{N}_1 = 16,550$ with $v(\hat{N}_1) = 2,950,000$.

4.16. Some additional sampling techniques

There are a number of additional techniques students should know about. Multistage sampling was used in Examples 4.4 and 4.5 where subsamples of the primary sampling units were actually enumerated. As pointed out there, it isn't necessary to consider the subsampling in obtaining a variance estimate. All that is needed is to use the subsample data to make estimates for the primary sampling units and treat those values exactly as one would if the entire unit had been tallied. However, it may be desirable to consider the

“within sampling unit” variability in order to do a more efficient job of designing the survey. This requires more complex equations which are given in many books on sampling [e.g., Cochran (1977), Thompson (1992)].

Another useful technique uses unequal probabilities in selecting samples. This approach is exemplified by the line intercept technique described in Chapter 5, and may be useful in any circumstance where the probability of selection may vary from unit to unit, either naturally or for convenience or improved efficiency. Texts or references to sampling techniques may refer to the Hansen-Hurwitz estimator. This is a method for using unequal sampling probabilities (see any of the sampling texts for details).

A relatively recent development is known as adaptive sampling. This may be a very useful approach when items of interest tend to be clustered, but in such a manner that there is no readily defined unit that contains all of the elements of a cluster. The technique provides a means for expanding the sampled area around primary units where a concentration of the items of interest is encountered, without biasing the results (which occurs with certainty if one simply expands the area to include more individuals). Details appear in Part IV of Thompson (1992) and a more extensive (and more theoretical) treatment appears in Thompson and Seber (1996).

Another potentially valuable approach is generally known as “kriging” after the South African mining engineer, Krige, who developed the initial approach in searching for profitable sites for mining for gold or other minerals. The approach is now used in petroleum exploration. In both of these examples drilling exploratory holes can be very expensive and time-consuming. The methodology thus utilizes spatial correlations among the available samples to estimate abundance or density on an area. A natural descriptive phrase thus is “spatial sampling”, and there are many instances where this may be useful in ecology. Thompson (1992:Part V) gives a useful summary and references to the extensive literature.

4.17 Exercises

4.17.1 Using a table of random numbers

Drawing a sample with the aid of a table of random numbers is not very complicated, but the student should try drawing a sample of 10 individuals from a population of 20, and another sample of 10 from a population of 1000 (the “populations” can be just the numbers 1-20, and 1-1000). Two approaches to starting points in the table may be considered. One is to somehow make a “random” start, (e.g., by closing one's eyes and touching a point on a page to select random coordinates in the table for a starting point) the other is to mark off sets of digits as they are used, going on through the table as different occasions for its use come up. The latter course is preferable for repeated surveys of the same areas. Note that samples of 10 out of a small population (like 20) may yield one or more repetitions of random numbers. Notice, too, that one has to use a two digit column of numbers, and many must be rejected with a population of 20. This seems to be even more of a problem with the population of 1000, since one should use 4 digits in order to permit the number

1000 to have a chance to be drawn. However, it is simple to arbitrarily assign the number 1000 to the 3 digit sequence 000 and thus use three columns (001 to 999, plus 000 for 1000). When working with EXCEL it is convenient to use the RANDBETWEEN() function, as that avoids the need to use a table of random numbers.

4.17.2 Determining sample size

Suppose that we want 95% confidence limits of about $\pm 15\%$ for the data in Example 4.1. What sample size is required if $N=1000$? Calculate sample size for $\pm 10\%$ for $N = 1000$.

4.17.3 An exercise in allocation

As an exercise in allocation, use the values of s_h^2 actually obtained in the caribou survey in Example 4.5 to calculate a new allocation and compare it with that actually used.

Another way to guess at the S_h to use for allocation is to assume the coefficient of variation (s/\bar{x}) is constant. Calculate the c.v.'s for each stratum, and try a "typical" values for allocations. Are there substantial differences between the various schemes? Comment on the results.

4.17.4 Computations for mortality survey

Compute \bar{y}_{st} and the total mortality estimate for Example 4.4 along with 95% confidence limits. It is often convenient to use $2[V(\bar{y}_{st})]^{1/2}/\bar{y}_{st}$ as "percentage limits" on survey results. Compute that value and compare it with the same result for example 4.3.

4.17.5 Stratified sampling in a vegetation study

A survey designed to estimate biomass of non-woody vegetation in a sagebrush stand (Eberhardt and Rickard 1963) provides an example of a different approach to stratification and illustrates some of the potential flexibility of sampling methods. In this example, proportional allocation was used in order to avoid advance preparations other than marking out the area well enough to avoid recounting individual plants. Two investigators worked together, one classifying and tallying each sagebrush plant into one of five strata, while the other checked off each plant on graph paper on which certain squares had previously been randomly selected as representing a plant to be sampled (it was thought that about 1/30 of the bushes should be sampled, so three numbers from 1-90 were designated as meaning "sample" and a table of random numbers was used to produce the sampling chart). When a "winner" turned up, the bush was subdivided into from two to five parts, and one of the parts was randomly selected. If that part was too large for weighing, it was subdivided and a random selection again made. The selected portion was then clipped, oven-dried and weighed.

Data from the survey are tabulated below. The "subsampling fractions" show approximately how much of an individual plant was actually removed. In the first record in the table, about 1/4 was removed, while in the second case there were two sub-divisions, and roughly 1/10 [(1/5) times (1/2)] was actually removed. Thus to estimate total weight for a given plant one would multiply by 4 or 10. Stratum IV contained the largest plants and the two plants actually sampled were sampled at rates of 1/56 and 1/28, respectively. Of course the divisions were not exact but any errors in subdividing will enter into overall variance of the survey estimates. There was actually a fifth stratum, but only one plant was sampled, so it has been left out of the tabulation.

As an exercise, the student should work out an estimate of mean oven-dry material and its variance for the entire sagebrush stand using the data in Table 3.1. Calculate an optimum allocation for a sample of the size used here (25), and compare with the proportional allocation (neglect the fpc). Calculate coefficients of variation. Comment on the results.

Results of stratified sampling of a sagebrush stand.

<u>Stratum</u>	<u>Number of bushes in stratum</u>	<u>Subsampling fractions for sampled bushes</u>	<u>Oven-dry weight of sample (g)</u>
I	169	1/4	0.60
		1/5,1/2	1.90
		1/2,1/2	2.05
		1/2,1/2	1.05
		1/2	1.20
II	309	1/4,1/2	1.60
		1/5,1/4	3.20
		1/5,1/5	1.45
		1/3,1/3	4.05
		1/3	2.05
		1/3,1/2	1.45
		1/4,1/4	2.40
		1/3,1/2	1.65
1/3,1/4	0.60		
III	301	1/5,1/4	1.85
		1/5,1/2	2.85
		1/5,1/4	7.15
		1/5,1/3	2.15
		1/3,1/4	4.10
		1/4,1/5	3.50
		1/4,1/3	5.25
		1/5,1/3	5.60
1/3,1/3	1.55		
IV	57	1/8,1/7	6.05
		1/7,1/4	3.60

4.17.6 Try jackknifing to calculate a standard error for Example 4.6. Compare your result with that given in the Example (661). Also, calculate bias estimates for \hat{R} using jackknifing and bootstrapping. Use 200 bootstraps. There is only one jackknife estimate of bias available, but you can run the bootstrap repeatedly and see how the bias changes. Comment on your results. Don't forget to consider the magnitude of the bias relative to the estimate.

4.17.7 Bootstrap the data for August 26 ($n=25$) from Example 4.8 and compare your results with the ratio estimators $N\text{-hat}(2)$ and $V(N\text{-hat}(2))$ given in the example, and with the binomial variance estimate given below. Do 200 bootstraps and calculate $\text{Bias}(\text{boot})$ from eq.(3.2). Run repeatedly and see how $\text{Bias}(\text{boot})$ varies. Is there an indication of appreciable bias? Recall that when simple random sampling of individuals is assumed:

$$v(p) = pq/(n-1)$$

where n here is 2500. The difference in the two estimates reveals why the binomial formula should never be used with cluster samples.

4.17.8 Try jackknifing the "interpenetrating sampling" results of Example 4.8, and compare the variance you get with that given in the example. Explain the results.

5.0 TRANSECT METHODS

5.1 Introduction

One of the main handicaps faced by the practicing population ecologist is that there really is no wholly reliable census method for most wild animal populations. In spite of the intensive theoretical work done on capture-recapture methods, there yet remain various unresolved issues. Hence, as frequently remarked here, it is essential to use more than one method, and to do as much cross-checking and testing of assumptions as possible. It is, of course, easy to recommend such a course, but very difficult to follow it. For census methods, the only sure test of the underlying assumptions may in fact be to secure an absolute population count. Even then we are left with the question of sampling errors--an observed discrepancy may simply be due to chance alone.

Since the capture-recapture methods require at least one outright capture of a sample of animals, followed by one or more repeat observations (which may be visual only), they are necessarily expensive to use in practice. Catch-effort and change-in-ratio methods can only be used effectively on a harvested population. These limitations lead to a need for a method based solely on visual observation since it is often relatively inexpensive and rarely poses any threat to the population. For these reasons, and no doubt because of wider recognition of the difficulties with other methods, there has recently been a considerable interest in transect methods.

One of the brighter prospects for the future of transect methods is that it may be possible to avoid the pitfall posed by the "equal probability of selection" assumption required to apply elementary probability models. Unfortunately, some of the early work on transect methods included an equally untenable assumption, that individual animals are randomly and independently distributed over the study area. I prefer to adopt the working axiom that this is never the case, even when tests for departure from a Poisson distribution are "not significant." I will cheerfully abandon that viewpoint whenever the power of a test of randomness can be shown to be suitably large. Presently, a random distribution of individuals may have to be assumed for various features of secondary importance, such as obtaining an approximate notion (really a lower limit) of a variance for an estimate. However, bootstrapping offers promise for better variance estimates, with less difficulty.

Avoiding the assumption of a random spatial pattern of individuals requires that we substitute random location of transect lines. Systematically spaced lines are much easier to use and have other practical advantages. Not the least of these is the fact that randomly located lines may fall very close together so that running one such line can influence animals on a nearby line. Some ways to avoid this problem are discussed below. Although we will not try to go very deeply into the issue here, it should be remarked that the choice between random and systematic sampling for transect methods is not as simple a matter as for, say, plot sampling. For plot sampling, two features are of paramount importance. One is to avoid a systematic pattern that is correlated with a similar pattern in the material being sampled. The other is that variances obtained from systematic samples usually overestimate the true variance.

Some of the transect estimators depend on a probability model that holds strictly true if and only if the transect lines are indeed randomly located. This may place a considerable premium on adhering to random sampling. Possibly the effect of departing from that model may eventually turn out to be of minor importance. Until more work has been done of the "robustness" of the estimators, we will adhere to the requirement of random sampling, when this is at all feasible. If very large areas are to be covered, it may not be practicable to use anything but a systematic arrangement of transect lines. However, the most apparent problems with systematic samples apply to situations of a much smaller scale, not when lines are very widely spaced.

5.2 A classification of transect methods

Terminology for transect methods is not well-established. We will adhere to a usage that includes three main classes. The strip-transect is essentially a long narrow plot, on which it is basically assumed that all of the individuals present can be seen and tallied. As such, there is no important difference from plot sampling. Some modifications tend to make it more interesting and worth special attention. These include censusing marine mammals at sea, when individuals may submerge for varying periods of time, and thus escape enumeration. In many transect applications the mobility of individual animals is neglected. This is not feasible for those species that are observed when in motion, such as small birds. Thus another kind of modification needs to be considered.

In most census methods individual objects are regarded as points scattered around the map. Sometimes this abstraction either is not practicable or is inefficient. The investigator may be directly concerned with such quantities as the canopy coverage of shrubs or the volume of logs left lying in a cutover area. There is then an advantage in measuring the size of the object intercepted by the line; hence the descriptive term of line-intercept method.

The third class is perhaps best known, and includes the methods in which decreasing visibility of objects with distance away from the transect line has to be taken into account. We include all such methods under the general heading of line-transects. Some writers use the same term to apply to both strip-transects and line-intercepts. The terminology adopted here has the advantage of being reasonably explicit in descriptive terms.

Where mobile animals are concerned, one important distinction lies in whether or not the animal responds conspicuously to the observer's approach. One can then measure the flushing-distance, i.e., the straight-line distance between observer and animal at the time the animal "jumps" or "flushes," i.e., leaves cover. This is also designated in the literature as the radial distance or as the sighting distance. It is essential, however, to also measure the angle or the right-angle distance (i.e., the distance between the track line and the animal).

When detection depends mainly on the observer locating the animal or other object without the help of a flushing-response, there is reason to believe (cf. Robinette et al., 1974) that the flushing-distance models may not hold, and may lead to biased estimates. As alternate approach in such cases is based on

use of right-angle distances. It is thus useful to consider a dichotomous approach (Fig. 5.1) to the several classes of methods. It should help readers keep the various circumstances and conditions leading to the several classes of methods in mind. A recent development in which the observer remains at a point and estimates distances to surrounding objects has been known as the variable circular plot, and is largely treated by methods used for right-angle transects.

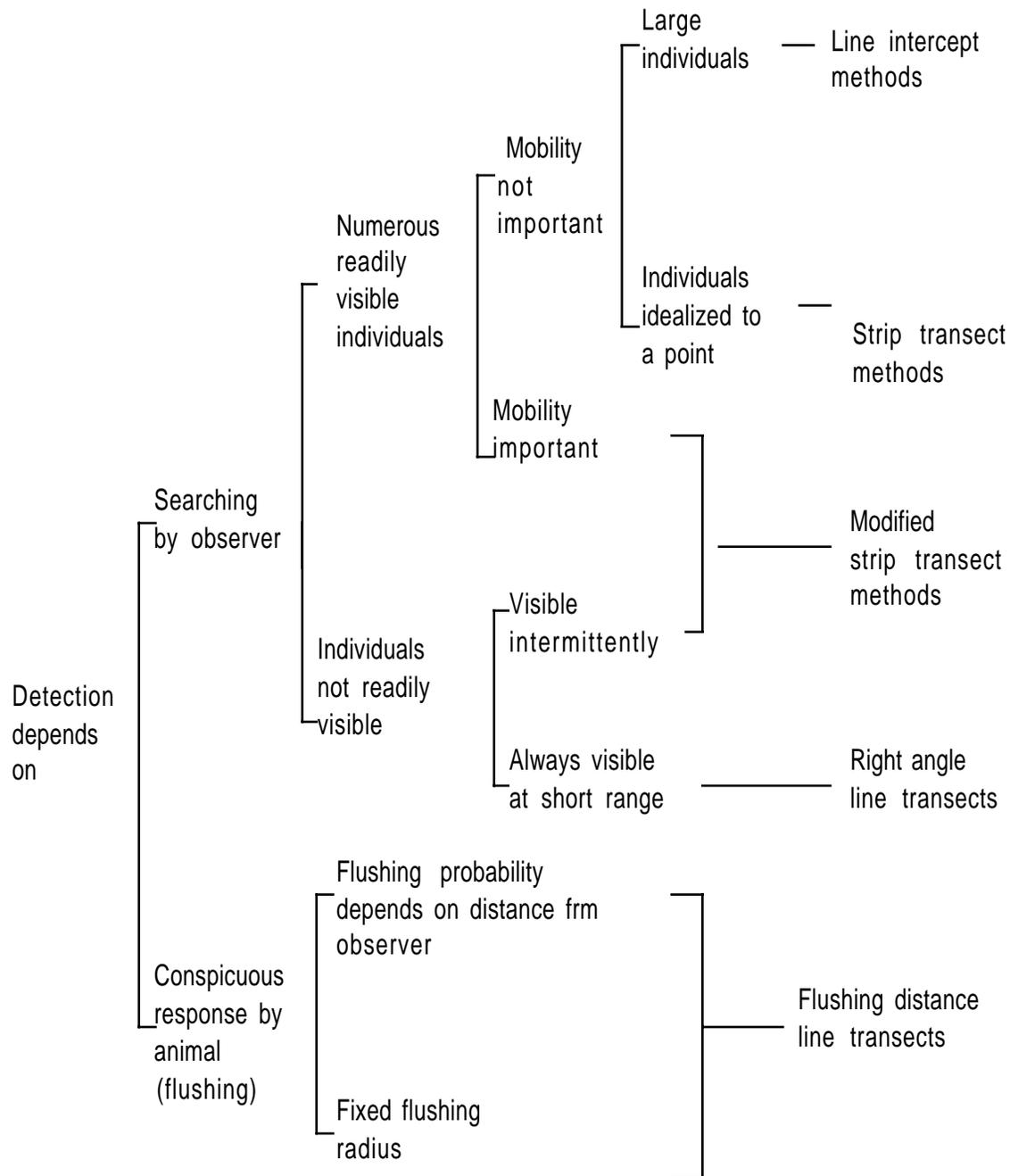


Fig. 5.1. A classification of transect methods.

5.3 The line-intercept method

The line-intercept technique has been used by plant ecologists for many years as a means of estimating "canopy-coverage." In that instance, the basis is simple and direct. All that is necessary is to measure the fraction of the total length of a given transect line that actually intercepts shrub canopies. The arrangement can be depicted as in Fig 5.2, which represents a rectangular study area having dimensions W and L , with a single transect (of length, L) intercepting two shrubs, for one of which the appropriate canopy-coverage measurement (l_i) is indicated. The technique can also be used for tree canopies by sighting upwards to find the margins of the canopy.

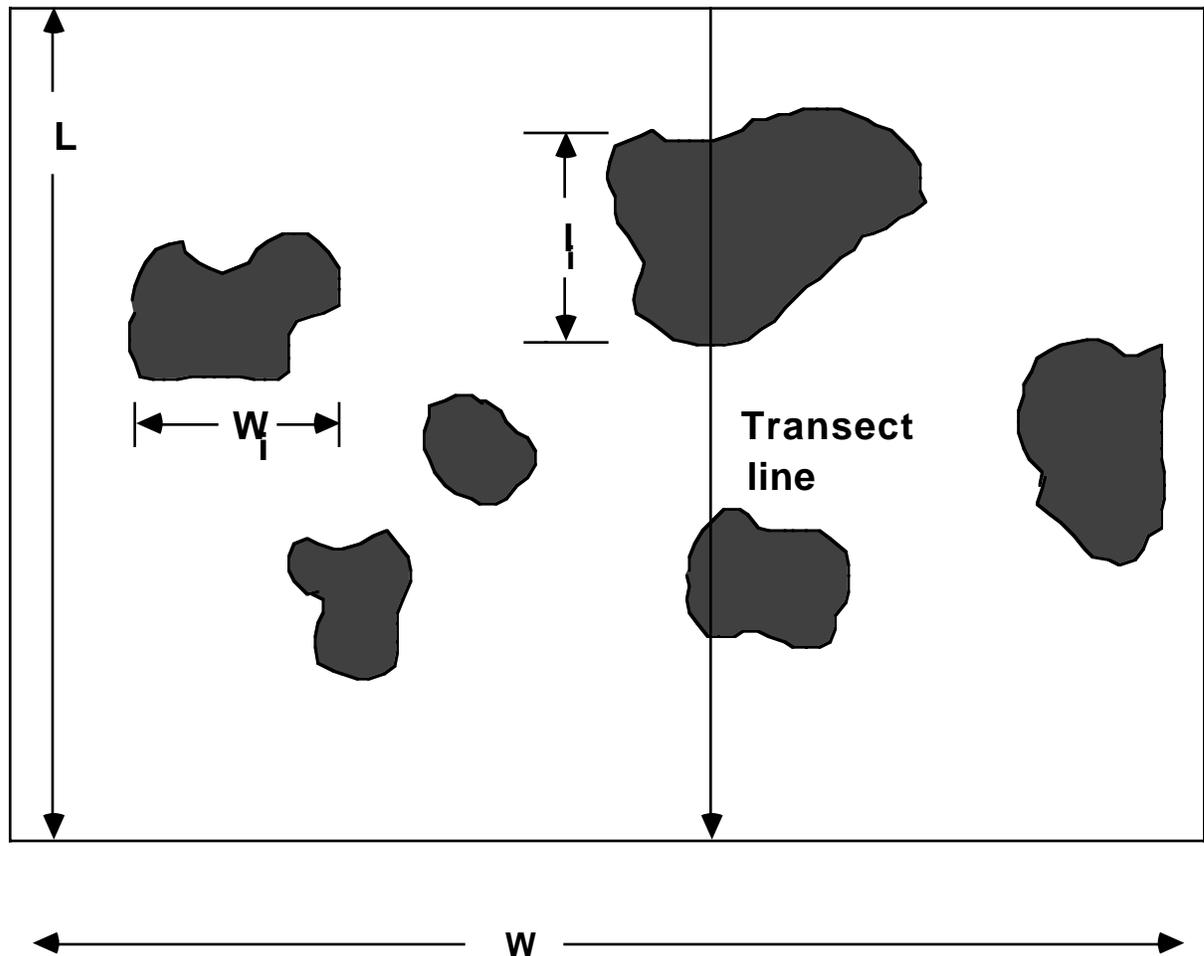


FIG. 5.2 Dimensions used in the line-intercept method. The shaded areas represent shrub canopies.

An unbiased estimate of canopy coverage is just the sum of the l_i observed on all of the transect lines divided by the total length of transects used. Unless there is some sort of regular pattern in the arrangement of the shrubs, very likely a systematic spacing of lines should not cause trouble in this situation. We repeat, however, that the basis of the results given in this chapter lies in randomized location of transect lines.

Two minor points may cause some trouble in actual practice. One is that the boundaries of the study area may intersect some shrubs. A useful rule, that should be decided on before the area is laid out, is to include such plants if they occur on, say, the northern and eastern boundary and exclude them on the other two boundaries. The second common problem is that many useful natural study areas (e.g., habitat types) are very irregular in shape. An easy way to deal with this kind of situation is to proceed as in Fig. 5.3. All that is needed is a baseline W that runs the full length of the area, and to utilize transects of variable length (L_i) with this length measured only within the study area. The calculations are illustrated in Example 5.1.

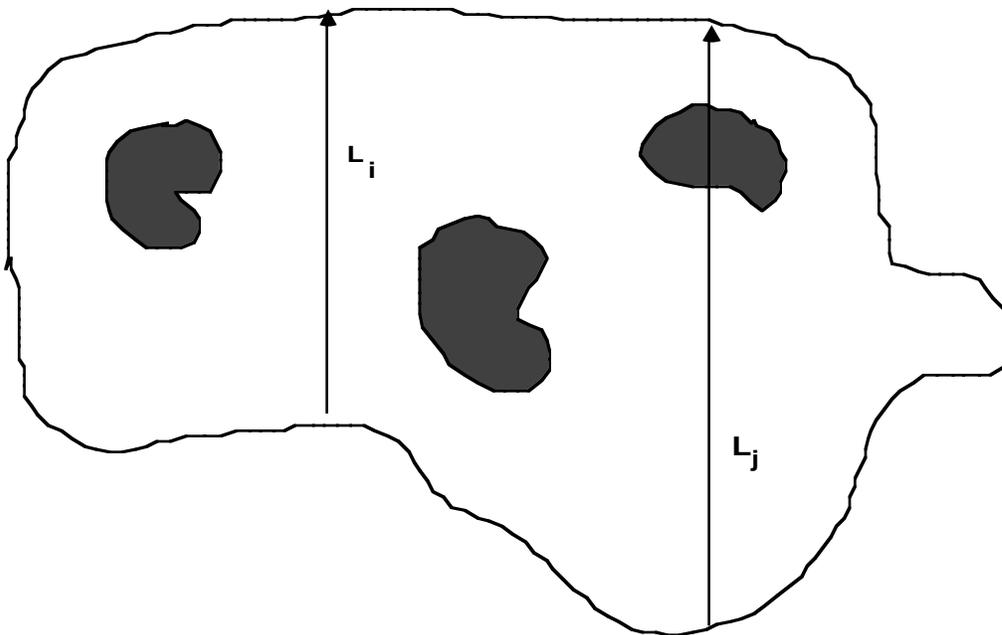


Fig. 5.3. Line-intercepts on an irregularly shaped area.

Estimates of the numbers or density (number per unit area) of individual plants have usually not been made by the plant ecologists in their use of the method. However, there is a simple way to obtain an unbiased estimate of density, although it "costs" an additional measurement. A biased estimate can be obtained without an extra measurement, and will be described first. McIntyre (1953) investigated the use of the measurement l_j for density estimation, and proposed several possible procedures. In using the length of the transect interception (l_j), he considered that the shrubs could be represented by a population of circles of varying diameter. Given random interceptions, it is then easy to derive a theoretical expression for length of intercepts which leads to the equation for density:

$$D = \frac{2}{\Pi nL} \sum_{i=1}^m \frac{1}{l_i} \quad (5.1)$$

in which n stands for the number of transects of length L and m for the number of shrubs actually intercepted (for each of which l_i is measured). As already noted, the transects do not have to be of the same length. The only change is to replace nL by ΣL_i in the denominator above.

Some trials with an artificial population led McIntyre (1953) to suggest that his method might well be useful for objects other than circles. Very likely the best results will be obtained for objects with smooth boundaries and few indentations or "scalloped" edges. This is because Eq. (5.1) uses reciprocals of the l_i , so that a few very short measurements will have a disproportionately large effect on the estimate. To avoid this problem, McIntyre recommended using the longest chord parallel to the transect line (and another equation). However, some better procedures are given below. Eq. (5.1) should mainly be used for an approximate notion of density when canopy coverage is the main purpose of the survey and an extra measurement is not justified.

An interesting alternative to McIntyre's approach can be described as "needle sampling" (DeVries 1974). It was originally developed for inventorying logs lying on the ground in cut-over areas. Instead of a circle, the object now is defined as a "needle" (which can be inscribed in a variety of only roughly elongate objects) and the famous results of "Buffon's needle problem" used to obtain a density estimate. The chief drawback is that the needles need to be oriented randomly, an assumption that may well be questioned in practice. More details appear in Example 5.2.

Example 5.1 Censusing prairie-dog dens

Line-intercepts were used to estimate the number of dens in a prairie-dog (Cynomys ludovicianus) colony by Eberhardt (1978b). The colony was elliptical in shape, with a long dimension of about 700 m and a maximum width of about 500 m. A systematic sample was used, with 9 transects spaced 66 m apart, and running across the narrower dimension of the area. The earth mounds at each den served in the same manner as shrub canopies in the usual application of the line-intercept method. For each mound intercepted by the transect line, measurements of the length of the interception (l_i) and the mound width (w_i), as shown in Figure 5.2. It should be noted that w_i is taken so as to measure the probability of interception for the mound, i.e., it is the distance between transects that just touch the right- and left-hand extremities of the mound.

The individual observations appear in Table 5.1, which also includes the distance between mound centers, or, at the ends of transects, the distance to the edge of the area grazed by the prairie dogs. This was regarded as the boundary of the study area. Calculations of density are thus for the grazed area immediately surrounding the mounds. Calculations on the basic data are summarized in Table 5.2. Proportion of the area covered by mounds is easily estimated, being just the total length of intercepts divided by the total length of transect lines. Thus for the first transect, it is:

$$p_i = 6.12/228.69 = 0.027.$$

For the entire area, the proportion covered is just the sum of all intercepts divided by the sum of transect lengths:

$$\hat{p} = \frac{\sum T_i}{\sum L_i} = 115.36/3578.9 = 0.0322.$$

This is a ratio estimate, for which a variance estimate is given in Chapter 4. Here T_i represents the total length of intercepts on the i th transect, i.e., $T_1 = 6.12$ m. The numerator could just as well be

written $\sum_i \sum_j l_{ij}$, where j denotes observations on a transect and i denotes the transect, but using the transect totals makes it easy to see that this expression has the same form as the ratio estimators of Chapter 4.

The finite population correction is neglected here since a small fraction of the population of mounds was actually tallied. Letting y = total interceptions (T_i) and x = transect length (L_i), the calculations are:

$$[CV(\hat{p})]^2 = \frac{1}{n} [c_{yy} + c_{xx} - 2c_{yx}] = \frac{1}{9} [0.1426 + 0.1507 - 2(0.1176)] = 0.00644$$

The estimated standard error for \hat{p} is then just $(0.00644)^{1/2}$ (0.0322) = 0.0032. The coefficients of variation are appreciably larger than recommended (in Chapter 4) for use of the ratio method. We thus propose that a simple approach might be used here, i.e., compute a variance directly from the proportions covered of the individual transects. This gives $p = 0.0329$ (averaging the transect values), with a standard error of 0.0028, so there is little difference from the ratio estimate.

Since widths (w_i) of the mounds were tallied, McIntyre's method for estimating density, Eq. (5.1), should not be used here. Apart from the factor of $2/\pi$ in the equation, calculations would proceed in exactly the same way as those utilizing widths, given next.

Although I recommend random sampling because most of the estimation procedures in transect work are based on random transect locations, the present example is one in which a systematic sample was taken. This was done mainly to study the pattern of spatial distribution of dens. With a systematic layout and distances between dens (Table 5.1), one can study the spatial arrangement of the dens. This is much harder to do with random transect locations, since random samples, especially relatively small ones, frequently leave sizeable gaps in spatial coverage.

In the present example, we can proceed in essentially the same manner for either random or systematic transect locations. The rationale differs somewhat, and needs to be mentioned for each case. It may be noted, too, that neither Eq. (5.2) or (5.3) is usable here since the transect lines are of variable length and the total area is not known. When the transects are randomly located, each individual transect yields an independent estimate of density, which can be calculated from Eq. (5.2), with $n = 1$. Using data from transect No. 1 (Table 5.1), we get (calculations in meters):

$$D = \frac{1}{L_i} \sum_{i=1}^{m_1} \frac{1}{w_i} \\ = \frac{1}{228.7} \left[\frac{1}{3.81} + \frac{1}{3.12} + \frac{1}{0.86} + \frac{1}{0.81} + \frac{1}{1.12} \right] = 0.0169 \text{ dens/m}^2$$

The same procedure can be used for each of the other transects, and the remaining question is one of how to combine 9 independent estimates (assuming, for illustration, that the transects had been randomly located along a baseline, as in Fig. 5.3). Averaging, and

computing a variance from the individual transect values, is both straightforward and legitimate.

The above procedure is often not very efficient, since it does not take into account the variation in transect lengths. A logical way to do this is just to weight the individual density estimates by the transect lengths, i.e., to calculate:

$$\hat{D}_{\text{overall}} = \frac{\sum_{i=1}^n L_i \hat{D}_i}{\sum_{i=1}^n L_i}$$

which is a ratio estimate just as in Chapter 4.

For a systematic sample, a somewhat different rationale might be used. This is because the uniform spacing of the transects permits viewing the area as being broken down into a number of strips of equal width. We then calculate an estimate of the number of dens in each strip, sum these, and divide by the total area (obtained by summing up the area of the individual strips). The procedure turns out to give exactly the same result as above, since a constant strip width is introduced in both numerator and denominator, and thus cancels out.

Using the data of Table 5.2 gives the following density estimate.

$$\hat{D}_{\text{overall}} = 70.013/3578.9 = 0.0196$$

The coefficient of variation is again estimated as in Chapter 4, without the finite population correction (for the reasons discussed above):

$$[CV(\hat{p})]^2 = \frac{1}{n} [c_{yy} + c_{xx} - 2c_{yx}] = 1/9[0.1698+0.1507-2(-.042)]=0.0042$$

The standard error of the estimate is $(0.0449)^{1/2} (0.0196) = 0.0042$. These results are very similar to those for intercept length. The above example might be converted into a prairie-dog census method if the number of prairie-dogs inhabiting a representative sample of dens could be estimated.

Table 5.1. Spacing (d_i), intercept lengths (l_i) and mound widths (w_i) for 9 line-intercepts in a prairie-dog "town". Transect no. 1 was the westernmost transect. Spacing (d_i) in meters, other measurements in cm. The first d_i is distance to first mound from margin of the area and the last d_i is distance from last mound to the other margin of the study area.

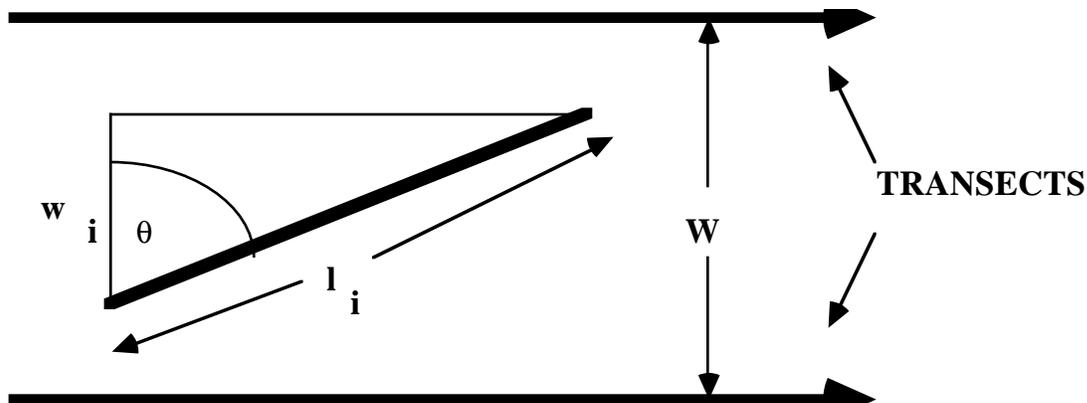
# 1			# 2			# 3		
d_i	w_i	l_i	d_i	w_i	l_i	d_i	w_i	l_i
2.84	381	127	16.06	130	74	138.0	124	107
46.3	312	114	6.62	290	236	39.69	117	104
39.69	86	152	27.4	132	96	67.1	218	124
5.67	81	84	6.62	109	74	9.45	168	58
20.79	112	<u>135</u>	35.91	401	224	1.89	160	74
<u>113.4</u>		612	124.74	282	160	34.02	368	157
228.69			41.58	274	<u>213</u>	27.4	198	<u>38</u>
			<u>28.35</u>		1077	<u>9.45</u>		662
			287.28			326.97		

Table 5.2. Summary of line intercept data for a prairie-dog "town"

Transect number	Length of transect L_i	Proportion of area covered by mounds intercepted	Number of mounds	Sum of intercepts T_i	Density estimates D_i
1	228.7	0.027	5	6.12	0.0169
2	287.3	0.037	7	10.77	0.0131
3	327.0	0.020	7	6.62	0.0126
4	566.1	0.034	13	19.12	0.0174
5	414.8	0.044	14	18.14	0.0247
6	423.4	0.031	14	13.15	0.0303
7	704.1	0.026	14	18.20	0.0132
8	375.2	0.031	10	11.54	0.0186
9	252.3	0.046	10	11.70	0.0364
	3578.9		94	115.36	

Example 5.2 "Needle" sampling

As with most similar sampling problems, this one is most readily conceptualized in reverse of what happens in practice. That is, we lay out the sampling scheme and then introduce, at random, the objects to be sampled. Here we suppose a systematic sampling pattern of parallel transect is laid out, and long, narrow objects of length l_i ("needles") are randomly distributed over the area. Let the spacing between the objects be W , and assume for simplicity, that $l_i \leq W$, i.e., that none of the "needles" is longer than the interval between transects. The relevant measurements appear in the figure below. A "needle" of length l_i is thrown randomly onto a field of parallel transect. The probability that it intercepts a transect depends on w_i , which in turn depends on the angle (θ) that the needle happens to assume.



Dimensions used in "needle" sampling.

We can write the probability of interception for a needle of given length (l_i) as:

$$P = \Pr\{\text{interception}\} = \frac{w_i}{W} = \frac{l_i \cos \theta}{W}$$

Since the angle (θ) is assumed to be randomly determined, it has a uniform distribution between 0 and 90° , or between 0 and $\pi/2$ in radians. Hence the frequency distribution of θ is:

$$f(\theta) = \frac{2}{\pi} d\theta$$

The expected value of p is then:

$$E(P) = \frac{2l_i}{\pi W} \int_0^{\pi/2} \cos \theta d\theta = \frac{2l_i}{\pi W}$$

If n needles are observed to intercept the transects, a simple estimate of the total number (N) of needles in the sampled population is:

$$\hat{N} = n/\bar{p} = \frac{n\pi W}{2} \sum_{i=1}^n \frac{1}{l_i}$$

and the estimated density of needles is:

$$D = \frac{n\pi}{2L} \sum_{i=1}^n \frac{1}{l_i}$$

The main problem with the method is that it is seldom safe to assume that the needles are randomly distributed. We thus recommend measuring w_i directly, and utilizing the equations given in the text for density estimation based on w_i . Students who want a demonstration of the method can readily construct one with a handful of kitchen matches scattered on a hardwood or tiled floor.

5.4 Length-biased sampling

The main issue in estimating shrub density from canopy measurement is one that is common to a very much wider class of sampling problems. Cox (1962, 1969) has used the highly descriptive term "length-biased sampling" to characterize procedures in which the probability of sampling a particular element in the population is proportional to some dimension of that element. Such a sample is by no means representative of the population, being very much biased towards individuals having the greater "lengths." In the present case, it is readily evident (Fig. 5.2) that the probability that a given shrub will be included in a sample taken by the intercept method depends on how "wide" it is with respect to the baseline (W) of the study area. The relevant measurement on the shrub is thus w_i (Fig. 5.2). It should be noted that w_i is the distance between tangent lines drawn parallel to the transect at the right- and left-hand extremities of the canopy.

The probability that a given shrub will be intercepted by the transect line is just w_i/W , on either Fig. 5.2 or 5.3. By measuring w_i accurately, one can thus determine the exact probability that a given shrub intercepted by the transect would be observed, before the transect line was selected. Given the probability of interception for each element observed in the sample, a straightforward argument can be constructed to derive a density estimate (Eberhardt 1978b). The principal equation is:

$$D = \frac{1}{nL} \sum_{i=1}^m \frac{1}{w_i} \quad (5.2)$$

As in Eq. 5.1, m is the number of objects intercepted (and measured) on n transect lines randomly placed in a rectangular area of dimensions W and L .

When the study area is not rectangular, a baseline W can be constructed as indicated in connection with Fig. 5.3, and density estimated from:

$$D = \frac{W}{nA} \sum_{i=1}^m \frac{1}{w_i} \quad (5.3)$$

where A represents the area of the study plot expressed in the same units (e.g., square meters) as the linear measurements (w_i and W). A useful approach when the area is not known is given in Example 5.1, which also illustrates variance calculations. Estimates of N , the total population are, of course, readily obtained from Eqs. 5.1 to 5.3 by multiplying by the area.

Lucas and Seber (1977) have derived equations comparable to those above, but use a different transect layout. They require that the transects be of short length, and both randomly located and randomly oriented with respect to the baseline. They obtain theoretical variance formulas for some circumstances. However, in the present state of theoretical and practical knowledge, it seems advisable to use variances estimated from replicated or interpenetrating sampling, as in Section 5.12 (below), or by the ratio method of Example 5.1.

The above method can be extended to deal with objects other than shrub canopies, and to aggregations of animals or patches of vegetation, so long as the identity and boundaries of each such "object" can be uniquely defined. It is also possible to substantially enlarge the area for interception of a given object. A method for doing this is well-known to foresters as Bitterlich's method. An "angle-gauge" is used to determine whether or not the apparent diameter of a tree is greater than a fixed angle, and thus whether or not the tree should be included in a sample. Readers not familiar with the method can simulate the field operation by extending an arm with the thumb in an upright position. If portions of an object (tree, rock, sign, etc.) protrude on both sides of one's thumb, then that object is "in" the sample. If the observer now moves away from the object until its margins just barely protrude beyond the sides of the "gauge" (thumb), then that position delineates the boundary of the interception area (Fig. 5.4). Circular objects like trees will have a circular boundary, but irregular objects will have an asymmetric boundary.

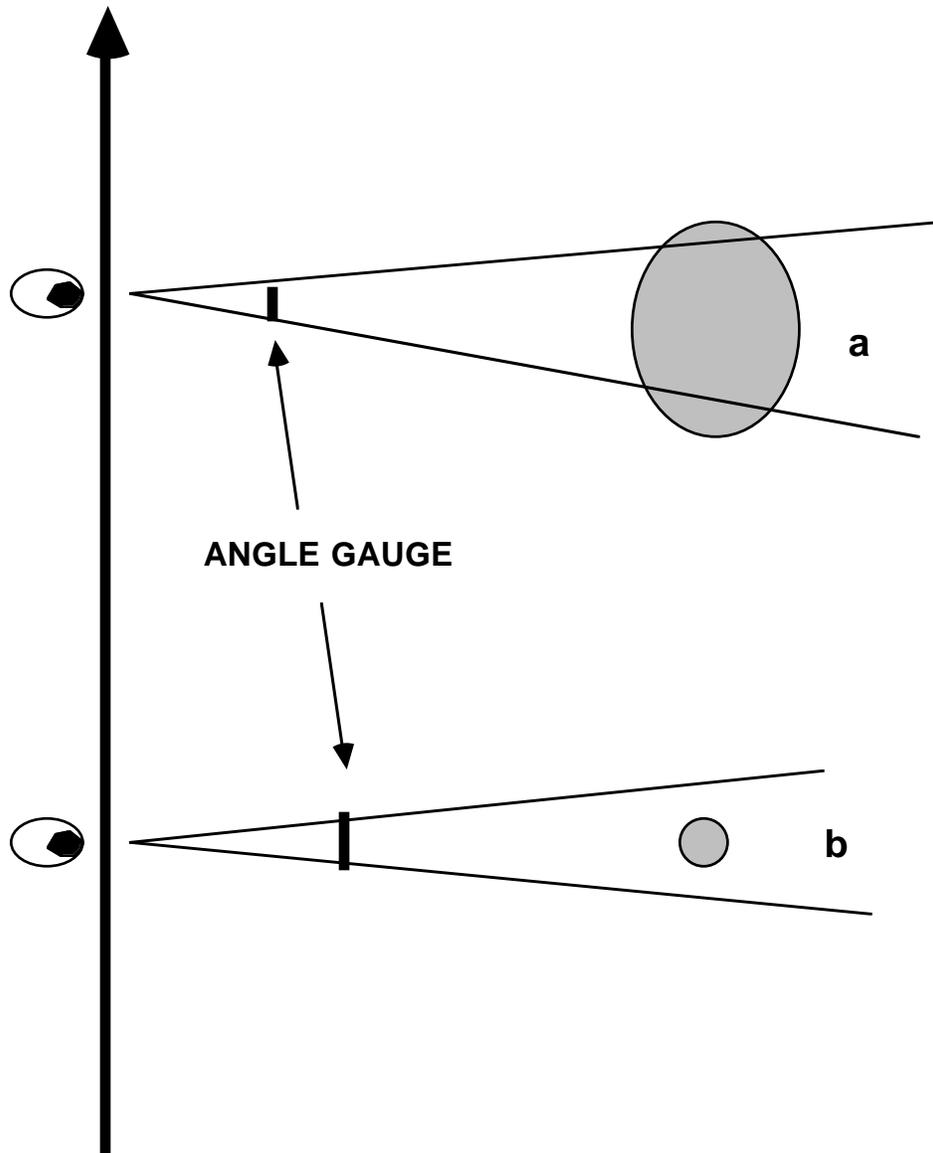


Fig. 5.4. Use of an "angle-gauge". Objects are "in" the sample when, as in (a), the sides protrude beyond the gauge. When the gauge blocks the object from view, as in (b), then it is not included in the sample.

Although Bitterlich's method is normally used only at fixed sampling points, it can be utilized as a transect method, as was proposed by Strand (1958). However, this will usually only be practicable for relatively rare objects, since "intersection" has to be determined by use of an angle-gauge as each object comes into a right-angle position on the transect line. The method might thus be most useful for something like a survey of den-trees in wildlife management.

Density may not be the main objective in some studies. When the volume, weight, or some other measurement is to be estimated, a simple ratio method can be used, and illustrated in Example 5.3.

Example 5.3 Auxiliary measurements

Often the primary objective of a study will be to estimate some attribute other than density, or in addition to density. Thus foresters are usually also concerned with basal area and volume of timber, while ecologists often want to estimate the biomass (total Weight) of vegetation. Methods for securing such estimates by ratio estimation were given in Chapter 4. A related method based on line-intercepts can readily be derived. Let X_i be the "auxiliary" measurement, such as weight or volume of the i^{th} object intercepted. A well-known way to estimate the average value of a sample of such objects is simply to "weight" each object inversely as the probability that it is included in the sample. Since this probability is proportional to w_i , we get the

$$\bar{X}_j = \frac{1}{m} \sum_{i=1}^m \frac{x_i}{w_i}$$

simple result:

If the above estimate is regarded as the estimated average on the j^{th} transect, then variable transect lengths can be adjusted for just as was done in Example 5.1, i.e.,:

$$\bar{x}_{overall} = \frac{\sum_{j=1}^n L_j \bar{x}_j}{\sum_{j=1}^n L_j}$$

and the same approach can be taken to obtaining a variance estimate (ratio method).

5.5 Flushing-distance line transects

In the line-transect method, the objects being censused are considered to be dimensionless points, and the probability of detection is assumed to be measured by use of distances between observer and object. Some trigonometry is involved, based on the distances and angle illustrated in Fig. 5.5. Just which measurements are taken will depend considerably on the particular field situation. The essential measurements for most purposes are r , the sighting-distance (also called radial distance or flushing-distance), and x , the right-angle distance. From simple trigonometry, any pair of the possible measurements can be used to calculate the others. However, precautions need to be taken to avoid measurement errors. I strongly recommend against visual estimation of either distances or the included angle (θ).

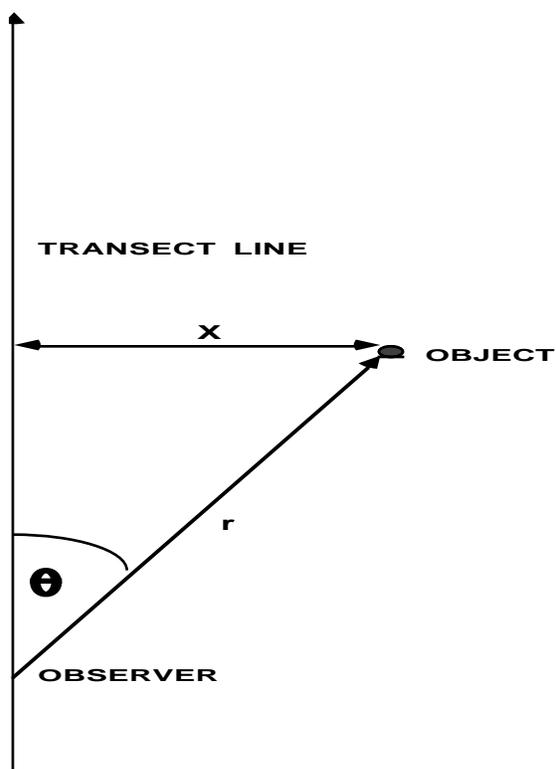


Fig. 5.5. Relevant measurements for the line-transect method. The vertical arrow shows the observer's path along a transect line.

In the flushing-distance model, the distance (r) between observer and animal at the time the animal flushes is the essential measurement. Since a test of the validity of the model is based on x/r (which is the sine of θ), these distances need to be measured as accurately as possible. When the right-angle distance method (described below) is used, only the distance x is utilized. A model for evaluating relative errors in measurements is described in Example 5.4.

Two basic flushing-distance line transect models have been proposed (Eberhardt 1968b). In one model it is assumed that the flushing-distance is fixed, i.e., that the individual animal flushes as soon as the observer crosses the boundary of a circle with radius equal to this fixed distance. This model is due to Hayne (1949), who noted that the fixed distance does not need to be assumed to be a permanent characteristic of the individual animal. The necessary assumption is that each animal on a census area has a fixed flushing-distance during the time when a given randomly located transect is run. In many circumstances it seems quite likely that the flushing-distance will depend very much on characteristics of the particular location in which an animal is resting.

The fixed-distance model permits a simple and direct analysis, proceeding in the same manner as for the line intercept method. The shrub canopy is now replaced by a circle of radius r , and it is assumed that the flushing distance (r) is measured accurately for each animal seen. It is also assumed that animals flush independently, i.e., that startling one animal does not change the behavior of the others. Analysis of the fixed flushing-distance

model then requires only noting that $w_i = 2r_i$, that is, w_i is the diameter of a circle of radius r_i . In analogy with Eq. (5.2), we now have:

$$D = \frac{1}{2nL} \sum_{i=1}^m \frac{1}{r_i} \quad (5.4)$$

with n the number of transects and m the number of observations as before. Irregular-shaped census areas can also be dealt with in the same manner as with line intercepts.

In some cases, groups of animals may flush together, as with broods of grouse, or flocks of small birds. If it can be shown, from field data, that flushing radius and group size are independent, it may be possible to use Eq. (5.4) to estimate the density of groups and multiply that estimate by average group size. If group size and flushing radius are correlated, one can still estimate the number of groups, but the average of group sizes is a biased estimator of the population mean.

As Hayne (1949) indicated, the expected flushing angle is 32.7° . A variety of field studies have yielded average angles that are close to this value for animals that "flush." Robinette et al. (1974), working mostly with animals that do not flush and inanimate objects, obtained wider mean angles. The underlying theory (cf. Eberhardt 1978b) shows that the frequency distribution of the ratio (x/r) of right-angle distances (x) to flushing-distances (r) should be that of the uniform distribution. Hence a simple chi-square test (Example 5.5) can be used to check on the validity of the model. If the test shows significant deviations from the hypothesis of a uniform distribution of x/r , then the best advice presently available is to utilize right-angle distances, as described below.

In the second model it is assumed that the instantaneous probability of flushing is a function of the current distance between observer and animal. It seems quite reasonable to assume flushing probability to increase steadily as the observer approaches, being nearly zero at a long distance and approaching unity in the immediate neighborhood of the animal. One might expect that an animal registers a variety of auditory and visual cues from an observer's approach, and that the cumulative effect of those cues results in an increased probability of flushing. Such a model is conveniently labelled the variable-distance model.

It does not seem likely that the two models can be distinguished on the basis of field observations. Either will lead to a frequency distribution of flushing distances, being based on a population distribution of flushing radii in the fixed-distance model, and on realizations of the probability model in the variable-distance case. Details of the theory appear in Eberhardt (1978a), and lead to the conclusion that Eq. (5.4) should be used for animals that flush. The theory also shows that flushing-distance (r) and flushing-angle (θ) should be independently distributed. Hence a useful further check of conformity to the flushing-distance model is to plot r and θ to see if there is any suggestion of association. Spearman's rank correlation coefficient might be used to test for correlation between r and θ (see, for example, Snedecor and Cochran 1967).

The fixed-distance model can be used to show that there appears to be a loss of efficiency (i.e., a larger variance results) if right-angle distances are used when the flushing-distance model holds. It should be noted that the current authoritative reference on "distance sampling" (Buckland et al. 1993) has dropped the idea of using sighting distances. They remarked that "Hayne's (1949) method is poor if $\bar{\theta}$ is not approximately 32.7° and may not perform well even if $\bar{\theta}$ falls close to this value, i.e., is not a robust method." Consequently they use only right-angle distances.

Example 5.4 Errors of measurement in line transects

Wherever possible, the relevant measurement for line transect estimation should be measured directly and as accurately as possible. However, it may at times be necessary to calculate the appropriate measurement by trigonometry on the pairs of the measurements of Fig. 5.5. Anyone doing this should be aware that the effect of incremental errors may vary considerably, depending on the particular pairs used. Suppose θ and x (Fig. 5.5) are measured and r is calculated as $r = x \sin \theta$. Then we note that $dr = x \cos \theta d\theta$, so that an incremental error ($d\theta$) in measuring θ results in a corresponding incremental error (dr) in the estimate of r . The absolute relative error in r is:

$$\left| \frac{dr}{r} \right| = \frac{\cos \theta}{\sin \theta} d\theta$$

when $\theta = 50^\circ$, $\frac{\cos \theta}{\sin \theta} = 11.43$, while for $\theta = 45^\circ$, $\frac{\cos \theta}{\sin \theta} = 1$, and for $\theta = 60^\circ$,

$\frac{\cos \theta}{\sin \theta} = 0.58$. Consequently, errors at small angles can have rather serious effects.

The above approach can be used to evaluate other arrangements, and a logical extension would be to explore the effects of errors on the final estimate by incorporating the theoretical frequency distribution. Doing so in detail calls for a knowledge of likely incremental errors ($d\theta$) at various angles, but this has not been investigated yet, to my knowledge. However, since the theoretical frequency distribution of angles is proportional to $\cos \theta$ (Eberhardt 1978b), it is obvious that errors at small angles ought to be avoided.

Too often, field data show evidence of gross errors. These appear in histogram plots of angles and distances as a tendency for measurements to pile up at angle like 0° , 30° , 45° and 90° , and for distances to be similarly grouped. "Trial runs" or pilot surveys are useful devices for catching such tendencies and training observers.

5.6 Right-angle distance line transects

When detection depends on the observer, it is unlikely that the flushing-distance (now sighting-distance) models can be expected to hold. The major summary of field experience is that of Robinette et al. (1974) and suggests that these models do not hold for animals that do not flush and for some inanimate objects. One prospect that needs study is that the manner in which observers scan ahead as they move along the transect may well influence the data. For the present, the safest course in circumstances where detection depends on the observer is to resort to use of right-angle distances,

and, as noted above, only the use of right-angle distances is recommended in the recent literature (Buckland et al. 1993). This may well require larger samples because of an added component of variability in using right-angle distances. Hence the need for research to determine whether sighting-distance models might be used, if suitable precautions are taken (in particular, advance surveys should show that the mean sighting angle is very close to 32.7 degrees). As noted above, only the use of right-angle distances is recommended in the recent literature (Buckland et al. 1993).

Supposing that the conservative course is chosen, i.e., that the right-angle distances are to be used, there then is the question of how to estimate density from such data.

A convenient frame of reference is that of Eberhardt (1968b). We again suppose that the study area is rectangular in shape as in Fig. 5.6 with a baseline of length W . It is assumed that virtually all of the observations made from a given transect line (represented by the solid line in Fig. 5.6) fall within a distance Z on either side of the transect line, and thus within the shaded area of Fig. 5.6. Hence if Z is small relative to W we can neglect most boundary problems. As suggested before, one can adopt the convention that observations made outside the study area on two boundaries will be included, and those outside of the other two boundaries will be neglected. So long as Z is quite small relative to W this approach should serve to deal with irregular shaped areas. To simplify the presentation, we now "fold" the left-hand side of the shaded area over onto the right-hand side and depict the actual observations of positions of observed individuals as in Fig. 5.7. If we then project these positions down onto a baseline, as shown by lines in Fig 5.7, we can analyze the data in terms of right-angle distances alone. The mathematical results then used (Eberhardt 1968b) are those of Parzen (1972). However, instead of an "intensity function," we use a "visibility curve," $g(x)$, as in Fig. 5.8. The essential features are that the probability of sighting an animal directly on the transect line shall be unity ($g(0) = 1.0$), and that the curve decrease smoothly away from the transect line. Further theoretical details appear in Burnham and Anderson (1976), Eberhardt (1968, 1978b) and in Buckland et al. (1993).

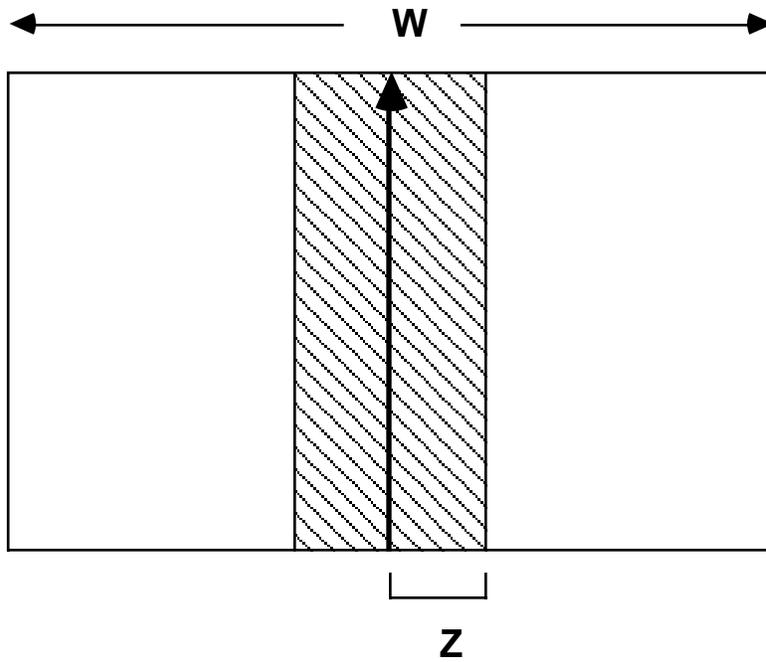


Fig. 5.6. Restricted area (shaded) used in many right-angle line transect methods.

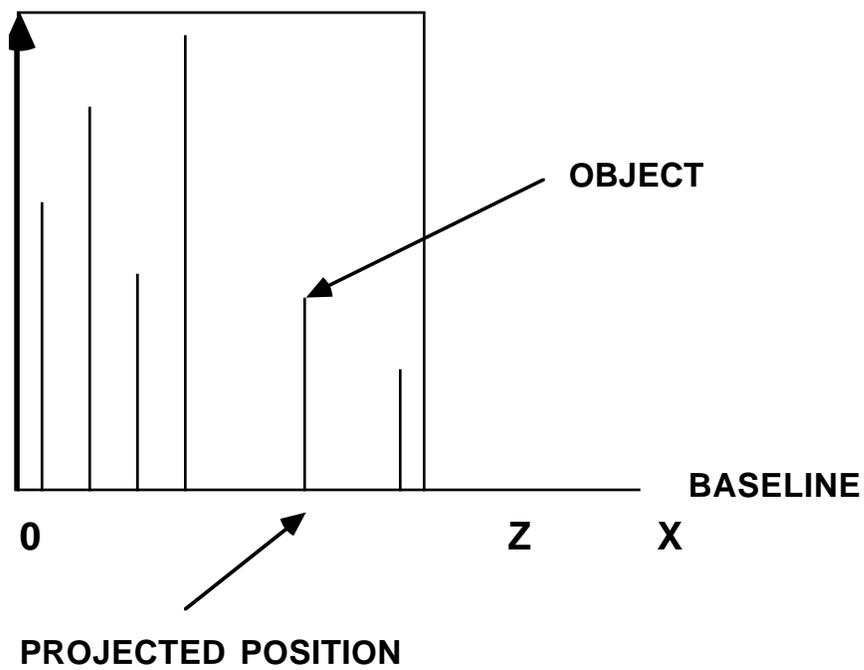


Fig. 5.7. Projection of observed positions on to a baseline.

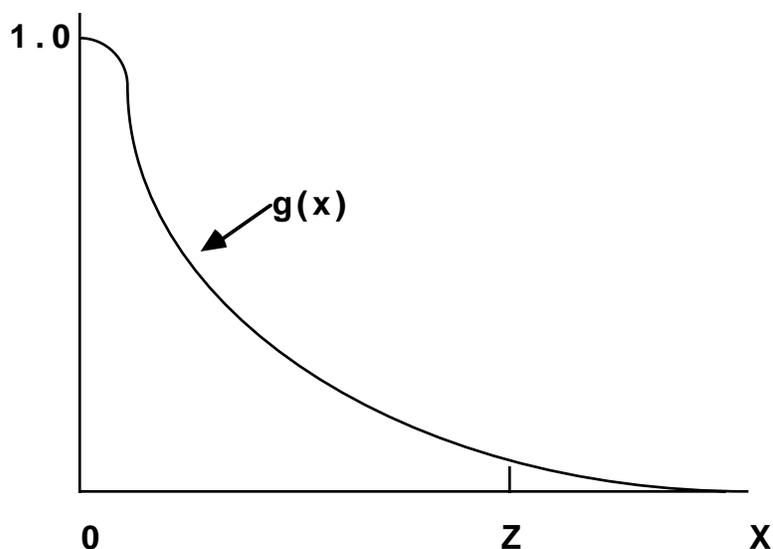


Fig. 5.8. A "visibility curve", $g(x)$, showing truncation imposed by neglecting observations beyond z .

The visibility curve of Fig. 5.8 is then the underlying model that generates the actual right-angle distances associated with a particular transect, represented by projecting to the baseline of Fig. 5.7. The position of Z in Figures 5.6 to 5.8 is arbitrary. It needs to be such that most, but not all, of the actual observations fall to the left of Z , when the entire set of data from a given study are considered. The actual selection of Z will be discussed below.

The visibility curve is not itself a frequency distribution, but it can be converted to such a distribution if it is divided by a constant that is the integral of $g(x)$, as shown by Burnham and Anderson (1976) and Eberhardt (1978b). Thus we have:

$$f(x) = \frac{g(x)}{\mu} \quad (0 \leq x \leq \infty) \quad (5.5)$$

where

$$\mu = \int_{x=0}^{\infty} g(x) dx .$$

A simple example of a visibility curve is the negative exponential as used by Gates et al. (1968) and Gates (1969). They found that it fitted data on flushing of ruffed grouse (*Bonasa umbellus*) quite satisfactorily, and gave estimating equations for both flushing-distance and right-angle distance data. An objection to this curve, however, is that it drops off at a constant rate. As suggested by Eberhardt (1968b), a more logical curve would be one that is nearly flat near the transect line, dropping off sharply some distance from the line, and then "tailing off" more gradually. Such a curve accommodates both the realistic assumption that a narrow strip census is feasible (i.e., that nearly all animals will be seen on a narrow strip centered on the transect line) and the observational fact that a few animals are seen at some considerable distances from the transect line. One curve fitting this requirement is the "reversed logistic" proposed by Eberhardt (1968) and described in more detail by Eberhardt (1978b).

The negative exponential curve has one parameter while the reversed logistic has two. If one parameter of the latter curve is very small, it becomes virtually indistinguishable from the negative exponential. This property was used by Eberhardt (1978b) to explore the effect of small deviation from the negative exponential on the resulting density estimates. The simulations conducted by Eberhardt (1978b) yielded biases (overestimates) of 17 percent and 50 percent if the true model were one of the two reversed logistics but the negative exponential were assumed to be the appropriate model. Consequently we do not recommend assumption of the negative exponential model.

A variety of other models have been proposed in the literature. The half-normal (Hemingway 1971) is a one-parameter model having the shape suggested above as appropriate. Anderson et al. (1978) have proposed a log-linear model, while Pollack (1978) presents an exponential power series model. Both of these "families" of models include the negative exponential and half-normal and provide considerable flexibility. The immediate problem is a lack of published experience covering a variety of field data. We will thus not try to make any specific recommendations about the use of particular frequency distributions.

One of the several recent developments in frequency distribution models is the "Fourier Series" estimator of Crain et al. (1978). It provides a highly flexible model that may be expected to give very good fits to field data. Both theoretical and simulation studies were employed by Crain et al. (1978) to show that the method has relatively small bias and high efficiency. That work has been followed up in detail, with several new models, and the results published in "Distance Sampling" by S. T. Buckland, D. R. Anderson, K. P. Burnham, and J. L. Laake (1993). Computations are available in the program DISTANCE which is available on the worldwide web along with a comprehensive manual and the full text of the book by Buckland et al.

Example 5.5 Testing flushing-angles

A simple test is available to check whether observed angles are in conformity with the underlying theory. The test is actually based on the distribution of $\sin\theta$, and holds for either the fixed or variable flushing-distance model (Eberhardt 1978b). It is, however, most readily derived for the fixed flushing distance model. From Fig. 5.5, $\sin\theta = x/r$. Consider a fixed flushing radius of r . Given that the animal is flushed (i.e., that the transect passes through a circle of radius r about the animal), and that transects are randomly located, it is evident that x will take on any distance between 0 and r with equal probability. Hence the distribution of x/r is uniform over the interval 0 to 1. A simple test is then a chi-square test. Divide the interval from 0 to 1 into equal sub-intervals, with the number selected so that the smallest expected number is about 5, and tally the observations of x/r by intervals. An example (from Eberhardt 1978b), appears in Table 5.3.

There are 84 observations, and 10 subintervals were used, so that the expected number in each interval is 8.4. The chi-square test is then:

$$\text{Chi-square} = \sum_{i=1}^{10} \frac{[\text{observed number} - 8.4]^2}{8.4}$$

It is worthwhile to tabulate individual deviations and chi-square calculations (as in Table 5.3), so that any aberrant observations can be identified if the test shows statistical significance. In the present example, the chi-square value (10.73) is well below the 95 percent significance level (18.31) for 10 degrees of freedom. Students should note that 10 degrees of freedom are used here, because the expected value is obtained independently from the data.

Table 5.3 Chi-square test for uniformity of $\sin \theta$ data for a census of the side-blotched lizard.

Interval ($\sin\theta = x/r$)	Number of observations	Deviations from expected number	Chi-square value
0.00-0.10	10	+1.6	0.30
0.10-0.20	7	-1.4	0.23
0.20-0.30	8	-0.4	0.02
0.30-0.40	15	+6.6	5.18
0.40-0.50	10	+1.6	0.30
0.50-0.60	10	+1.6	0.30
0.60-0.70	6	-2.4	0.68
0.70-0.80	4	-4.4	2.30
0.80-0.90	9	+0.6	0.04
0.90-1.00	5	-3.4	1.38
	84	0.0	10.73

5.7 Density Estimation

The generally accepted estimator for right-angle line transect models is (Seber 1982), Buckland et al. (1993):

$$\hat{D} = \frac{m}{2L} \left(\frac{1}{\hat{\mu}} \right) \quad (5.6)$$

where m is the number of objects observed, and L is the (total) length of transect on which the m objects are observed. The estimate of the reciprocal of μ is calculated from the observed distances. This is done by noting that, in Eq. (5.5), $f(0) = 1/\mu$. Thus the main objective of the various methods is to obtain an estimate of the frequency of observations "on" the transect line, or $f(0)$. Consequently, an equivalent form of Eq. (5.6) is just:

$$\hat{D} = \frac{m}{2L} \hat{f}(0) \quad (5.7)$$

Looking back to Eq. (5.4), it may be observed the Hayne's (1949) estimator is of this form, except that $f(0)$ or the reciprocal of μ is estimated from the average reciprocal of flushing distances, i.e.:

$$D = \frac{m}{2nL} \left[\frac{1}{m} \sum_{i=1}^m \frac{1}{r_i} \right] = \frac{1}{2nL} \sum_{i=1}^m \frac{1}{r_i}$$

and that n transects, each of fixed length L , were used.

Since, as we have already remarked, the "state of the art" and the theory of line transect are now described in detail by Buckland et al. (1993), we will not attempt to review all of the currently used methods. Some examples appear in Example 5.6 and I recommend consulting the current literature for recent improvements. I suggest use of the "distribution-free" methods of the next section as a check on any other method used.

Example 5.6 Density estimation for line transects

Calculations for two of the methods will be illustrated on the set of data in Table 5.4. These data come from actual observations made in a line transect study (Eberhardt 1978b) of the side-blotched lizard (Uta stansburiana). An artificial grouping of the data into 8 transects has been used here as a device to illustrate variance calculations. Since these data appear to conform to the theoretical model for animals that flush, it may be possible to use Hayne's method, Eq. (5.4). It may be remarked here that the "flush" exhibited by these animals is a dart for cover, and that nearly all sightings result from this cue, as basking animals are not readily seen before they move.

Using Eq. (5.4) gives the results of the summary table (Table 5.5). The equation is used with $n = 1$ for individual transects, i.e.,:

$$\hat{D} = \frac{1}{2L_i} \sum_{i=1}^m \frac{1}{r_i}$$

The individual transect results can be combined with the ratio estimate of Example 5.1:

$$\hat{D}_{\text{overall}} = \frac{\sum_{i=1}^n L_i \hat{D}_i}{\sum_{i=1}^n L_i}$$

Variance calculations proceed in the same manner as for line intercepts (Example 5.1)

Table 5.4. Line transect data from a lizard study. Flushing (r) and right-angle distances (x) for individual transects.

#1		#2		#3		#4		#5		#6	
r	x	r	x	r	x	r	x	r	x	r	x
46	46	91	51	137	29	60	43	51	23	91	16
82	26	42	25	21	17	67	34	42	14	74	58
59	10	36	32	84	25	51	18	109	57	57	10
42	36	126	88	62	25	68	37	120	43	101	40
40	35	43	15	79	37	55	13	60	32	74	72
		100	96	80	0	55	33	39	32	46	13
				70	0	168	90	81	32	46	15
				95	35	90	73	67	18	99	37
				95	32	78	25	55	0	87	0
				61	41			165	75	58	11
				58	42			269	33		
				24	13			269	25		
								85	35		
								168	98		
								50	0		
								83	4		
								42	27		
								75	10		
								104	0		
#7		#8									
r	x	r	x								
153	48	85	79								
112	45	112	55								
126	34	94	0								
61	45	78	15								
53	43	158	68								
78	0	153	72								
53	17	153	74								
59	49	42	27								
78	64	42	27								
150	146										
128	34										
114	38										
90	54										
93	24										

To illustrate the use of right-angle distances, we use the half-normal distribution. This requires the assumption that right-angle distances from the transect line have the relative frequency given by:

$$f(x) = \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

This is just the familiar normal distribution, but with $\mu = 0$, i.e., centered on the transect line. Also, the distribution is multiplied by a factor of 2 in order to permit "folding-over" half of the distribution, and thus considering observed distances as though they all fell on one side of the transect line. Recalling that the general form

of the density estimator for line transects calls for an estimate of $f(0)$, we get:

$$f(0) = \frac{2}{\sqrt{2\pi} \sigma}$$

The parameter (σ) is estimated just as it is for the normal distribution, i.e.,:

$$\sigma_i = \left(\sum_{j=1}^m \frac{x_j^2}{m} \right)^{1/2}$$

With the exception that the divisor is m , rather than $m-1$, since in this case the mean is known (i.e., is zero). Inserting the above expression for $f(0)$ in Eq. (5.7) gives:

$$\hat{D}_i = \frac{m}{L_i \sqrt{2\pi} \hat{\sigma}_i}$$

The individual transect estimates appear in Table 5.4, and are combined just as with Hayne's method above:

$$\hat{D}_{\text{overall}} = \frac{\sum_{i=1}^n L_i \hat{D}_i}{\sum_{i=1}^n L_i} = \frac{0.822}{500} = 0.0016.$$

Table 5.5. Summary of line transect data for a lizard study.

Transect <u>number</u>	Length <u>L_i</u>	Number of <u>observations (Hayne)</u>	<u>L_i D̂_i</u>	<u>L_i D̂_i</u> (half-normal)
1	30	5	0.050	0.061
2	50	6	0.052	0.040
3	60	12	0.109	0.169
4	80	9	0.066	0.076
5	100	19	0.129	0.197
6	80	10	0.074	0.113
7	60	14	0.083	0.100
8	40	9	0.056	0.066
	<u>500</u>	<u>84</u>	<u>0.619</u>	<u>0.822</u>
Density estimates			0.00124	0.00165

5.8 A "distribution-free" method

The terms, "parametric models" and "non-parametric models" have been used in the literature to classify line transect methods. We prefer to avoid that classification because the procedures thus far used mostly do involve parameter estimation. Hence we prefer to label the method presented here as "distribution-free," since it does not require the specification of a particular frequency distribution or "visibility curve." Burnham and Anderson (1976) suggest some other approaches that do not depend on a specific frequency distribution.

The method presented here is one originally devised by Cox (1962, 1969) and adopted for right-angle line transects by Eberhardt (1978b, 1979). A physical analogy, "length-biased sampling" was described in Section 5.4, in

reference to line-intercept methods. In that situation the physical size (length) of an object determines the probability that it will be intercepted by a randomly located transect line. However, a much larger class of situations may be included if one considers what Patil and Rao (1978) have described as "weighted distributions." They derive an equation of the form of Eq. (5.5) by supposing that the true frequency distribution cannot be observed directly, and that the observed frequency distribution is somehow "weighted" in the observation process.

In line transect work, the weighting function is what we have called a visibility curve above (cf. Fig. 5.8). Given random location of transect lines, the probability that an object will actually exist at a right-angle distance, x , from the transect lines is given by a uniform distribution. That is, theoretically, any distance is equally likely. However, the distances we actually observe depend on the visibility curve. Hence objects directly on the transect line are seen with certainty ($g(0) = 1.0$), while those at a considerable distance are seen very infrequently. Hence, formally, Eq. (5.5) should be written as:

$$f(x) = \frac{g(x)dx}{\mu} \quad (5.8)$$

so that dx represents the uniform probability that an object exists at any distance x from the transect line, and $g(x)$ is the "weighting function."

The main value of all of this is theoretical, in that it lets us extend the rather concrete notion of a line intercepting an object to the more abstract notion of a visibility curve. Further details and applications to a wide range of problems can be found in Patil and Rao (1978) and in the references cited in that paper.

Cox's method depends on tallying observed distances within fixed intervals away from the transect line. Thus all of the observations within a distance, Δ , on either side of the line are added up and used to estimate the true proportion of all observations, denoted $p(0,\Delta)$, that fall in that interval. Hence if there are k_1 observations within the distance Δ , we estimate $p(0,\Delta) = k_1/m$. Similar estimates are constructed for $p(\Delta,b\Delta)$, the next pair of parallel belts (Fig. 5.9) and $p(b\Delta,d\Delta)$. Cox's original method used only two intervals, but an extension to three or more intervals is readily obtained (Eberhardt 1979). However, it appears that the variance of the resulting density estimate increases as the number of intervals is increased (Eberhardt 1979), so we will limit the present discussion to two intervals.

An estimator for two intervals is (Eberhardt 1979):

$$\hat{\left(\frac{1}{\mu}\right)} = \frac{(b^2 - 1)\hat{p}(0,\Delta) - \hat{p}(\Delta,b\Delta)}{b(b-1)\Delta} \quad (5.9)$$

where Δ is the width of the inner interval and $b\Delta$ is the width of the inner two intervals (Fig. 5.9). The quantities $\hat{p}(0,\Delta)$ and $\hat{p}(\Delta,b\Delta)$ are estimated as described above, i.e.,

$$\hat{p}(0,\Delta) = \frac{k_1}{m} \quad \text{and} \quad \hat{p}(\Delta,b\Delta) = \frac{k_2}{m} \quad (5.10)$$

where there are k_1 objects observed within the belts of the width Δ on either side of the transect line, k_2 are seen within the two intervals (right- and left-hand sides of the transect line, Fig. 5.9) demarcated by Δ and $b\Delta$, and m is the total number of objects observed regardless of distance from the transect line.

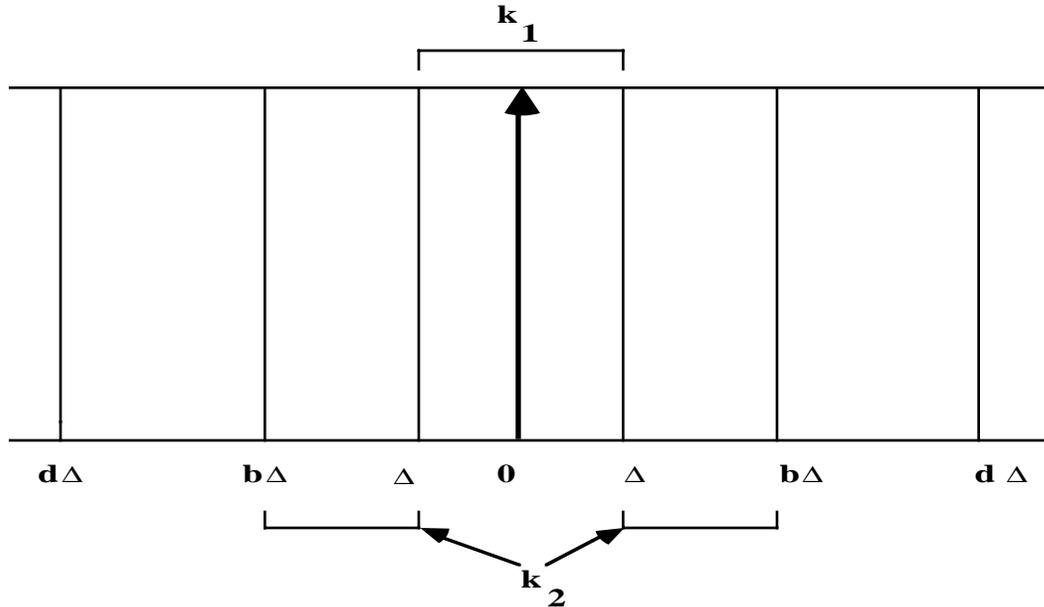


Fig. 5.9. Intervals or "belts" used in Cox method.

Having estimated the reciprocal of μ , all one needs to do is insert that estimate in Eq. (5.6) to estimate density per unit area, where the units are those in which right-angle distances (x_j) and length of transect (L) are recorded. Cox (1969) used $b = 2$, so that the inner and outer intervals are equal (i.e., they are both of width Δ). Since m appears in the numerator (Eq. 5.6) and in the denominators of $p(0, \Delta)$ and $p(\Delta, b\Delta)$, it effectively cancels out in the calculations. Hence, if one combines Eqs. (5.6), (5.9) and (5.10), the result is:

$$\hat{D} = \frac{(b+1)k_1 - k_2}{2Lb\Delta} \quad (5.11)$$

where we have used the result that $(b^2 - 1) = (b + 1)(b - 1)$. One apparent consequence of this simplification is that m (the total number of objects) is not required for density estimation. However, that quantity is essential in studying variability of the estimates, and thus should be recorded, except in special circumstances. One such situation may be in cases where identification of objects beyond a distance of $b\Delta$ is uncertain. Calculations are discussed in Example 5.7.

An interesting variant of Cox's method is the case where only one interval, of width Δ (on each side of the transect line), is used. This reduces Eq. (5.11) to:

$$\hat{D} = \frac{k_1}{2L\Delta} \quad (5.12)$$

We then have simply a strip transect, or "Kelker's method," in which it is assumed that all of the objects are observed within a strip of width 2Δ . It may

be noted that if the number of objects observed in the two belts is equal, i.e., if $k_1 = k_2 = k$, then Eq. (5.11) becomes:

$$\hat{D} = \frac{(b+1)k - k}{2Lb\Delta} = \frac{k}{2L\Delta}$$

so that we again have a strip transect.

Another variant, of interest to ornithologists, is Emlen's (1971) method. It turns out (Eberhardt 1978b:15) that Emlen's method essentially reduces to use of Kelker's method, or a strip transect. Details appear in Example 5.8. Since the assumption that all animals (or objects) are seen in the inner strip is an important and uncertain item, we recommend that two intervals be used in practice.

Variance estimation for the Cox method is a subject that needs more research, particularly research supported by field data. The weak point in present theoretical approaches is that they assume that the number of individuals observed (m) is Poisson-distributed, which essentially amounts to assuming random distribution of individuals. Since this is not likely to occur in practice, the present variance estimates are likely to be too low, i.e., underestimates. For the Cox method with two intervals, a convenient expression of the variance is obtained as (Eberhardt 1979):

$$[C.V.(\hat{D})]^2 \doteq \frac{(b+1)^2 k_1 + k_2}{[(b+1)k_1 - k_2]^2} \quad (5.13)$$

where b , k_1 , and k_2 are as defined above, and $C.V.(\hat{D})$ stands for the coefficient of variation of the density estimate. For practical purposes, if $C.V.(\hat{D})$ equals, say 0.25, one can propose approximate confidence limits on an estimate as being the estimate ± 50 percent (i.e., we round the usual 95 percent normal curve "Z-value" of 1.96 to 2.0).

An alternative procedure for variance estimation is to employ the "replicate sampling" idea, i.e., to break the total sample down into random subsets, calculate a density estimate from each such subset, and obtain the variance estimate from the resulting set of independent density estimates.

For planning purposes, a rough approximation (Eberhardt 1978b) is:

$$C.V.(\hat{D}) \doteq \left(\frac{4}{m}\right)^{1/2} \quad (5.14)$$

Two examples on actual data (Eberhardt 1979) suggest this equation underestimates the results of Eq. (5.13) by roughly 10 percent. Seber (1973) and Eberhardt (1978b), using different approaches, suggest a comparable result for flushing-distances (sighting-distances, radial distances) to be:

$$C.V.(\hat{D}) \doteq \left(\frac{2}{m}\right)^{1/2} \quad (5.15)$$

We thus have an indication that the variance using right-angle distances and the Cox method may be much as twice that obtained for flushing distances.

Example 5.7 The "Cox" method

The right-angle distances of Table 5.4 can be used to illustrate the Cox method, as given by Eq. (5.11). The main problem lies in selection of the two intervals Δ , and $b\Delta$. My recommendation (cf. Eberhardt 1979) is to include about 80 to 90 percent of the observations inside $b\Delta$. If we let $b = 2$ and $\Delta = 35$, then $b\Delta = 70$. Although density can be calculated directly from Eq. (5.11), in this case it would be desirable to make the intermediate calculation represented by Eq. (5.9), for use in comparison with the results of Example 5.6 above. This is because the transect lengths used in the example are artificial, so the best comparison is to estimate $(1/\mu)$ or $f(0)$. Hence the entries under sums of $L_i D_i$ from Table 5.5, are best compared with the estimate obtained from the Cox method. Note, however, that these quantities need to be doubled for comparison. Calculations with the Cox method can be carried out transect by transect, and it is probably worth doing so on Table 5.4 just to see how the estimates behave. With small samples, however, it is preferable to make a single estimate for the entire area (i.e., combine all of the observations in Table 5.4).

Example 5.8 Emlen's method

A method due to Emlen (1971) became quite popular with ornithologists. It depends on a "coefficient of detectability" which is determined by an intensive study on one area, and then used to adjust counts in other areas. The basic approach is to use the observed data to determine where visibility drops off. If we let this point be Δ , the assumption is that all birds are seen between the transect line and Δ . Suppose k_1 birds are counted in this strip, and that we want to estimate the number of birds expected to be found between the transect line and some outer boundary, R . The logical estimate is just $(k_1/\Delta)R$. Emlen divides the total number of birds (k) seen between the transect line and R by this projected number and calls this the "coefficient of detectability":

$$C.D. = \frac{k\Delta}{k_1 R}$$

This clearly amounts to an estimate of the proportion seen of the birds present between R and the transect line. Emlen then proceeded to divide the number seen (k^*) on a new area by the coefficient of detectability, and regarded this as a population estimate for the new area:

$$\hat{N} = \frac{k^* k_1 R}{k \Delta}$$

Actually R is selected so that a fixed transect length (1 mile) gives \hat{N} as the number of birds per 100 acres. Hence \hat{N} is really a density, and we note that if units of feet are used $2RL = 100(43,560) = C$, so we can write $R = C/2L$ and express \hat{N} as:

$$\hat{N} = \frac{Ck^*}{k} \left[\frac{k_1}{2L\Delta} \right]$$

Since the quantity in brackets is just Kelker's estimate [Eq. (5.12)], Emlen's procedure turns out to have the following steps:

(1) Estimate density on one area by Kelker's method, while counting all of the birds visible on 100 acres (k).

(2) Count all of the birds visible on 100 acres (k^*) on a new area.

(3) Use the ratio k^*/k to project the Kelker estimate of the first area to the second area.

Several limitations of the method seem apparent. One is that the visibility curve is assumed to have a particular form, i.e., all birds are seen out to some particular distance. A second limitation is that it is assumed that we can locate that distance from observed data. A third, and major problem, is that it is assumed that the visibility curve remains constant from area to area, and time to time.

These several limitations can be avoided simply by taking one precaution. That is to record separately all birds seen between the transect line and the distance Δ on the second area. One then can use Eq. (5.11) with $b = R/\Delta$ and get a direct estimate of density independently on the two areas.

5.9 Assumptions underlying line transect methods

A decision to use a particular line transect method needs to take into account the underlying assumptions. The list given here is based on seven assumptions given by Gates et al. (1968) and Seber (1973,1982), but is rearranged to show just which assumptions are required for a given line transect method. We assume that randomly located transect lines are utilized, and thus drop one restrictive assumption, that of random location of the objects being censused, which is not required for randomly located transect lines (Eberhardt 1978b).

The first three assumptions are basic and whether or not they are met will depend on behavior of the observer and of the animal being censused. They are:

- (1) No animal (or object) is counted more than once on a given transect line.
- (2) When flushed, each animal is seen at the exact position it occupied when startled by the observer's approach. Obviously this does not apply to animals or objects that are fixed in place during the census.
- (3) The response behavior of the population on a census plot does not change during the course of running a given transect.

Definitions of the response behavior serve to distinguish the various methods. One of these is achieved by defining a visibility curve as follows: The probability that an animal, or object, being seen, given that it is at a right-angle distance x from the transect line is a simple function, $g(x)$, such that $g(0) = 1$ (i.e., animals, or objects directly on the transect line are observed with certainty). These assumptions then suffice for right-angle line transects. Assumption (3) now means that the visibility curve does not change during the course of running a given transect line.

A fourth assumption serves to define conditions for the fixed flushing-distance line transect. This assumption defines the response behavior of the animals:

(4) Individual animals have fixed flushing radii, during the course of running a given transect, and flush if, and only if, an observer comes within this characteristic distance (r). It is of course, also necessary to assume that r is accurately observed and recorded.

An alternative assumption, plus some others, leads to the variable-distance line transect:

(4a) The animals are homogenous with regard to their inherent response behavior.

(5) The sighting of one animal is independent of the sighting of another.

(6) The instantaneous probability of flushing is a function, $f(r)$, of the radial distance, r , between animal and observer.

These several assumptions lead to the theory of the variable distance model (Eberhardt 1978b). It may be remarked that one could assume a particular mathematical model for $f(r)$ and proceed to derive efficient estimates for density estimation under such a model. This has been done in some of the literature (e.g. Gates et al. (1968) and Gates (1969)). We have previously mentioned two tests that should be applied to observed data before the flushing-distance method and Eq. (5.4) is used (cf. Example 5.5). We will return to discussion of some aspects of the above assumptions in a subsequent section on sampling design.

5.10 Strip transects

The simplest case of a strip transect occurs when the objects being censused are readily visible and sufficiently abundant to permit using a restriction on width of the strip covered. The method then amounts simply to a sample survey using long, narrow plots. The methods of Chapter 4 can then be applied. A basic assumption is that all of the objects on the plot are tallied.

When there is a reason to believe that not all of the objects on the plot are seen, then it may be necessary to introduce a visibility-curve. We have preferred to treat such situations under the heading of right-angle line transects (as in Fig. 5.1). However, this is mostly a matter of preference, and one could classify such situations as "strip transects using visibility-curves." This may be a more natural-seeming description in circumstances where a finite boundary exists on strip width. An example is the study of Anderson and Pospahala (1970), who counted duck nests on dikes. The width of the dike then provided a natural boundary on strip width. However, the methods of estimation will remain essentially those treated here as right-angle line transect, except that the total number of objects tallied (m) will be those tallied within the strip, and the visibility-curve, $g(x)$, is truncated (cut-off) at the strip boundary.

An important issue in such situations is that a visibility-curve has to enter the calculations of density in some manner. Several publications have used a procedure to correct for reduced visibility away from the transect line. This consists of summing the total observations from the entire survey for various intervals out from the transect line. With the symbols used above, one would thus have k_1 observations in the interval $(0, \Delta)$, k_2 in $(\Delta, b\Delta)$ and so on. It is then assumed that all of the animals are seen in the innermost interval, and the fraction seen in the other intervals is calculated from the observed data, i.e., k_2/k_1 , k_3/k_1 , etc. These rates are then used to adjust daily (or weekly, etc.) observations in the outer intervals, supposedly correcting them for the fraction missed. However, this procedure simply results in adjusting all of the intervals to equal the central one (to k_1 observations), so one may as well use only that interval and not bother with the rest. The same kind of procedure has also been used to adjust for numbers seen by time of day when certain observation periods give the highest counts.

While the Cox method (Sec. 5.8) does not require postulation of a specific visibility-curve, it does take the existence of such a curve into account in the estimation procedure. All of the other methods actually used thus far do specify a particular curve or "family" of curves.

Strip transects have been widely used in aerial surveys, largely of terrestrial animals. It is now well-established that not all of the animals on the strip are seen by the aerial observers. Caughley (1974) has summarized evidence on this point. Caughley et al. (1976) have conducted some experiments designed to explore the effects of strip width, altitude and speed on the numbers of animals counted. They go further, and use multiple regression equations to attempt to correct for these variables. However, I do not recommend the use of such equations, because a very uncertain sort of extrapolation is utilized--going from the observed data to zero strip width, speed, and altitude.

Two alternatives seem worth consideration. One is to utilize such experiments to arrive at a standard set of observation parameters, and to then regard the observed data as an index. When accurate counts by another method are feasible, one can then attempt to go further by "ground-truth" correction. A variety of special precautions need to be taken in aerial surveys, and are described in a publication by Norton-Griffiths (1975).

As mentioned in Sec. 5.8, the Cox method might be applied to aerial surveys in the form of Eq. (5.11), using two strips. This approach is particularly attractive in that it will not ordinarily be possible to attempt to record right-angle distances. About all that is likely to be feasible is to record observations in two intervals, demarcated by markers on windows and struts. Since most such surveys are conducted by observers who look out side windows of small aircraft, a particular drawback to this arrangement needs to be noted. This is that the visibility-curve is not likely to be that of Fig. 5.8, with certainty of observation of animals directly on the transect line. This is because the transect line is directly under the aircraft, and not readily viewed by the observer.

Unless a specially fitted-out aircraft is available, with provision for a "bow" observer to look directly forward and down, the only alternative seems

to be to attempt to locate, by experience, the right-angle distance that can be viewed effectively and comfortably by observers. The observers should then concentrate on "guarding" (covering) a fairly narrow strip starting at that point. This strip then becomes the interval $(0, \Delta)$ and frequent glances up and out are used to tally animals in the outer strip $(\Delta, b\Delta)$. The critical point is to concentrate enough effort on one line which is considered to be "the" transect line (normally there will be one such line on either side of the aircraft). In larger aircraft, it may be feasible to assign two observers to a side. One does nothing but scan the "track line," while the second observes the outer strip $(\Delta, b\Delta)$.

5.11 Modified strip transects

Three modifications of strip transects have been mentioned above. One includes corrections for decreasing visibility with distance from the transect line, and we have elected to cover this situation under right-angle line transects. A second is the case where animals, largely marine mammals, are visible only intermittently. The third has to do with animals that are in fairly constant motion, as with some small birds.

One basis for dealing with animals that submerge, and thus are not always visible on a transect, assumes a constant diving time (u) and a constant period on the surface (s) between dives. This is not particularly satisfactory, since both quantities may vary, and needs modification. In shipboard counts of ringed seals (*Phoca hispida*), McLaren (1961) assumed that all surfaced seals could be seen out to a fixed distance (r) from the vessel. This, too, is not a very reasonable assumption, as quite certainly there will be a decrease in visibility with distance. It might, however, be acceptable if this distance (r) is kept reasonably short. If the average probability (\bar{p}) that a seal within a strip of width $2r$ will be counted can be estimated, then the observed count (m) within the strip can be translated to an estimate of density as:

$$\hat{D} = \frac{m}{2rL(\bar{p})} \quad (5.16)$$

i.e., the number present is estimated as m/\bar{p} and this is divided by the area of a strip of length L .

McLaren (1961) gave the probability (\bar{p}) of seeing an individual seal, given that it is in the strip, as:

$$p = \frac{t}{s+u} + \frac{s}{s+u} \quad (5.17)$$

where t denotes the duration of the period when a surfaced seal would be visible to an observer. This varies according to the right-angle distance from the vessel, since the radius of visibility (r), shown in Fig. 5.10 limits the time an animal can be seen at a given right-angle distance. If v denotes velocity of the vessel, then (see Fig. 5.10):

$$t = \frac{y}{v} = \frac{r \sin \theta}{v}$$

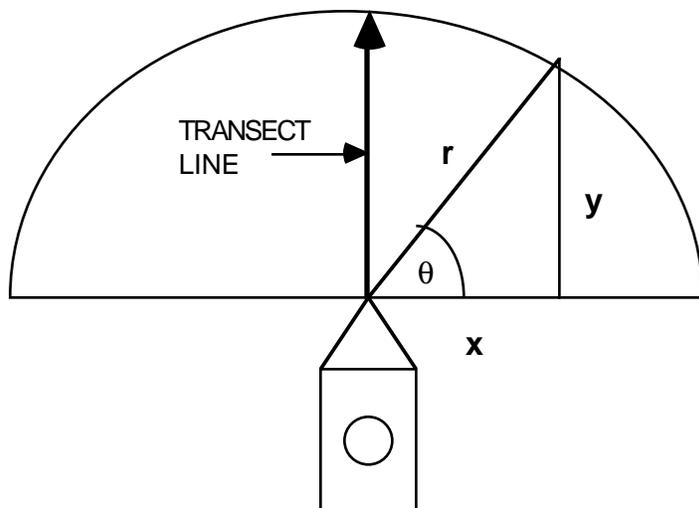


Fig. 5.10. Dimensions used in shipboard surveys of marine mammals. The arrow denotes the ship's course along a transect line, while the observers scan the semi-circular area of radius r .

McLaren assumed that no point could be kept under observation for as long as u minutes, that is, $t < u$. Given this restriction, Eq. (5.17) can be obtained by considering the two successive intervals representing a dive (u) and the succeeding time on the surface (cf. Exercise 5.13.1). If we make the usual assumption of random transect locations (and thus a uniform probability of a seal being present at a given right-angle distance), then Equation 5.17 can be "averaged" to obtain:

$$\bar{p} = \frac{\pi}{4} \frac{r}{v(s+u)} + \frac{s}{s+u} \quad (5.18)$$

This differs from McLaren's (1961) result (see Eberhardt 1978b).

A similar approach has been employed for censusing whales visually, except that a narrower width of field forward of the vessel is scanned. Also, much longer detection distances are postulated, due to the greater visibility of "spouting" or "blowing" by the whales. Doi (1974) developed an expression for the probability of detection. He also assumed a constant diving time, so that the same question of the effect of a variable diving time arises in connection with his results. An important difference in the two approaches is that McLaren assumes $t < u$, i.e., that submerged seals may go undetected, even if they are in the zone of maximum possible detection directly ahead of the vessel. Doi, however, postulated a zone within which the observation time is long enough that any whales were seen with certainty. Doi also introduced a correction factor (K) for the prospect that observers may fail to see some whales, even though they do surface and "blow." This factor is:

$$K = 1 - \left(1 - \frac{\theta_p}{2\theta_1}\right)^s$$

where θ_p represents the visual angle of the observer, θ_1 is the angle searched (on either side of the vessel) and s is the number of observers. Buckland et al. (1993) described another method for whales called "cue counting".

The second modification to be considered here is that required to deal with animals, such as small birds, that are in motion during the course of the survey. Yapp (1955) proposed an approach based on the kinetic theory of gases. The mathematical aspects were later reviewed by Skellam (1958). The two equations involved are:

$$\hat{D} = \frac{z}{2rv} \quad (5.19)$$

where D = density of the population, z = number of encounters per unit time, v = average velocity of the animals relative to the moving observer, and r = range or radial distance within which an animal must approach the observer to effect an encounter.

$$v^2 = \bar{u}^2 + \bar{w}^2 \quad (5.20)$$

where \bar{u} = average velocity of the organisms and \bar{w} = average velocity of the observer.

An important assumption is that the behavior of the animals is not influenced by the presence of the observer. Eq. (5.19) is based on the assumption that the area in which encounters take place is a circle of radius r . This, then is the same sort of troublesome assumption encountered before in this chapter. If we let $z = m/T$, where m is the number of animals observed

during the total time of the survey (T), and also assume $\bar{u} = 0$, then:

$$\hat{D} = \frac{m}{2r\bar{w}T} = \frac{m}{2rL}$$

where L represents the total distance traveled by the observer. We thus have the usual equation for a simple strip transect. One can, of course, let

$\bar{w} = 0$, i.e., assume that the observer sits still and base results on the average velocity of the organisms:

$$\hat{D} = \frac{m}{2r\bar{u}T} \quad (5.21)$$

This has some attractive features, in that the radius (r) can probably be determined with reasonable accuracy under such circumstances, and a motionless observer is less likely to influence behavior of the animals. A drawback is in the "representativeness" of the spot selected for observation. No doubt random selection of several spots would help on this score. However, if the radii vary, as they likely will, then questions of the effect of density vs. cover type may need to be considered.

An important problem with the above method is that of measuring the average velocity of the animals (\bar{u}). Clearly this cannot be done during the survey, at least not if the observer is also moving. However, if the observer sits still he might then use a stop watch to time movements of animals and thus estimate their velocity.

Little use seems to have been made of Yapp's method, so that it is difficult to provide an evaluation based on experience.

5.12 Survey design

Although much of the discussion of line transect methods is couched in terms of results obtained on a single transect line, practical use of the method will frequently require combining results from a number of separate lines into a single sampling unit. This will be especially true in terms of variance calculations, since precautions need to be taken to avoid individual sampling units on which no animals are observed. As we have already indicated, we believe variance estimates based on theory should mainly be used for such purposes as comparing methods of estimation, appraising bias, and the like. Another important use is in obtaining approximations suitable for estimating sample sizes in planning a survey, as in Eqs. (5.14) and (5.15).

In the actual analysis of survey results, we recommend variances be estimated directly from the survey estimates, as illustrated in the several examples. However, the investigator should not wait until the data are all in hand before considering how this is to be done. The arrangements for analysis of the data should instead be decided at the survey design stage.

Usually the survey will require a number of days for completion so that a worthwhile precaution is to arrange the sampling plan so that the transects run in the same sub-area are spread out over the total time interval during which the survey is conducted. Thus "replication in time" is introduced into the survey, and it may be useful, in analysis of the data, to try to evaluate any trends in time. When this kind of arrangement is feasible, it will be important to randomize the locations of successive lines falling in the same sub-area. In fact, this may well be the best way to use randomized sampling, in that the separation in time will usually eliminate the need for concern about having two transects fall close together. When large areas must be dealt with, it will usually not be possible to use a scheme of this sort.

The above scheme may be illustrated by reference to Fig. 5.11, which shows a study area divided into three subunits, denoted by vertical lines in the figure. One randomly located transect line (L_1 , L_2 , and L_3) is shown in each sub-area for each day on four successive days. To obtain a variance estimate on the basis of "interpenetrating," or "replicate" sampling, one simply calculates an estimate of density for each day, and uses that estimate in the variance calculation. That is, density is estimated as

$$\hat{D} = \frac{1}{n} \sum_{i=1}^n D_i \quad (5.22)$$

and variance as

$$s^2(\hat{D}) = \frac{1}{n(n-1)} \sum_{i=1}^n (D_i - \hat{D})^2 \quad (5.23)$$

where in this case $n = 4$. Note that the variance given is that of a mean, i.e., $s^2(\hat{D})$ is usually described as the standard error.

For larger areas, travel time will be rather too costly to permit use of single transects in each sub-area. Alternatives are to use several randomly located lines in each sub-area, or to use a systematic arrangement with a randomly selected starting point. Suppose three lines in each unit are to be used. Then each base line length (W_i) is divided into three segments of length $W_i/3$. A random location is selected in the first and the remaining units spaced out by the interval $W_i/3$. It may be noted that the baselines (W_i) of Fig. 5.11 are not of the same length. This is because of the irregular shape of the region--it is best to try to keep the areas of the subunits about equal. Differing lengths of transect line can be handled as described in Example 5.1. Note that the three systematically placed lines discussed above should be treated as one transect line in the analysis.

Stratified random sampling (Chapter 4) may well be desirable in line transect work. Example 4.6 illustrates use of stratification with a strip transect. Unfortunately stratified sampling has not been used much with line transects, so we have little experience to draw on for planning. One prospect is that the use of variable sampling intensity (by strata) will call for making individual population estimates for each stratum. Obtaining separate variance estimates for each stratum may thus require fairly intensive sampling in each stratum.

A very important feature of survey design is to review the underlying assumptions (Sec. 5.9), and to consider whether the proposed design is likely to result in violation of one or more of the assumptions. I have repeatedly recommended random sampling, as this is the basis for the present theory. A practical alternative is a systematic sample with a random start. With animals that are highly mobile, one has to avoid a sampling pattern that places lines near enough together that individual animals might be seen twice on the same systematically arranged sampling unit.

The various assumptions that have to do with response behavior obviously require good knowledge of the species and situation. Some species behave in ways that make them doubtful candidates for line transect censusing. When right-angle distances are used, the "behavior" of the observer is of crucial importance. Some design arrangements can help reduce the effect of observer differences. For example, if several observers are used in a single aircraft, for an aerial survey, they should rotate through the viewing positions fairly often (in small aircraft this may be practicable only on landing). This practice helps "average out" observer and position effects.

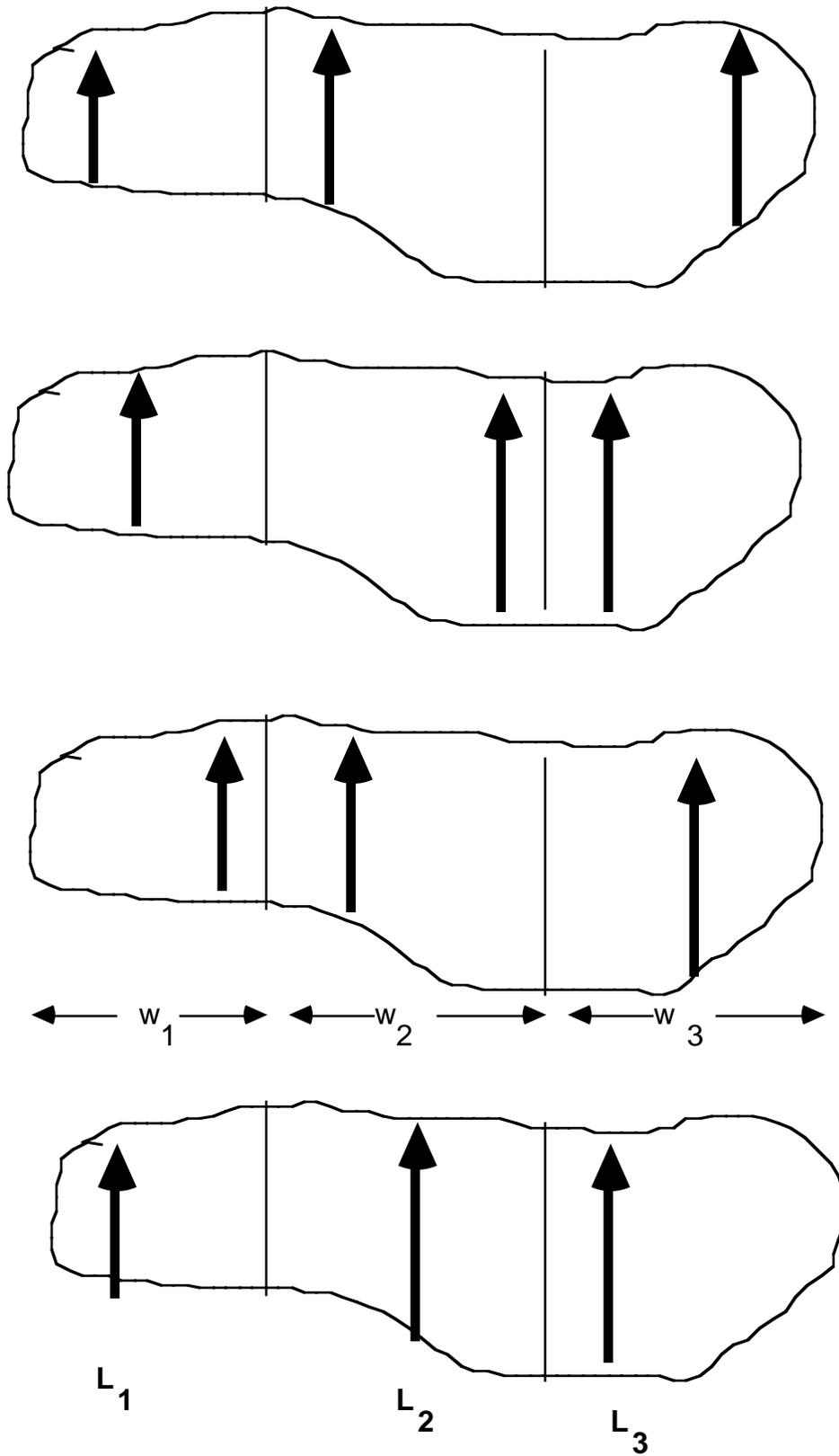


Fig. 5.11. Randomization of transect lines within sub-areas on successive census days.

Since circumstances of line transect surveys are quite variable, it is not possible to review each situation with regard to the assumptions. The investigator needs to understand them, and to take precautions wherever possible. Another example is that lines should not be run close to sharp breaks in cover, topography, etc. One can, of course, simply not census on the side of a line that parallels such a break. Sometimes it is possible to arrange that the lines go at right-angles to such "discontinuities," and this should help. A similar reasoning dictates that transect lines should not run the "long way" of an elongated study area. Methods to minimize errors of measurement and data-recording are of course essential in survey design and planning.

A variety of recommendations concerning transect methods in censusing marine mammals appears in Eberhardt et al. (1979). Some of these may be useful in other circumstances. As noted earlier, much more detail on recently developed methods appears in the book by Buckland et al. (1993).

5.13 Exercises

5.13.1 Calculate $CV^2(p)$ from Table 5.2 using ratio estimation, and calculate a standard error for p using the individual proportions. Show your calculations.

5.13.2 calculate the weighted average density from Table 5.2 using ratio estimation, and its standard error. Show your calculations.

5.13.3 Carry out calculations for the Hayne and half-normal methods for the lizard data of Table 5.4. Show calculations.

5.13.4 Do the calculations for Example 5.7.

5.13.5 Estimate the D_i for Exercise 5.13.3 and tabulate these along with the D_i from Exercise 13.4. Compute coefficients of variation treating each transect as an independent estimate. Compute correlations among the three sets of data. Also compute the coefficient of variation for the Cox method given in Eq. (5.13), combining the data from the several transects.

5.13.6 Components of variance

It was remarked in Sec. 5.5 that the fixed distance model could be used to show that a larger variance results if right-angle distances are used for estimation, rather than the flushing distance. One way to appreciate this is to recall that the basis for estimation for the fixed distance model depends on doubling the flushing-radius to determine the probability of observing that individual. Doing this with the right-angle distance introduces an additional component of variability due to the fact that the observed right-angle distance (x) falls randomly between zero and the flushing-distance (r). Students with some training in mathematical statistics may want to try calculating coefficients of variation for x and r , assuming x to be uniformly distributed on $(0,r)$ and that r has some underlying distribution, say $m(r)$. One can then find the two C.V.'s in terms of the first 3 moments (μ_1, μ_2, μ_3) and obtain a notion of the relative difference in efficiency.

5.13.7 Deriving a sighting probability

Students should attempt to derive Eq. (5.17), assuming s and u are fixed and that $t < u$. Nothing beyond elementary probability considerations is involved, but a diagram helps.

5.13.8 A seal census

McLaren gives data as follows. Ship's speed 0.12 nautical miles per hour, visibility limit 0.32 miles, $s = 1$ minute and $u = 3$ minutes, and 43 seals were counted on a given transect. He does not give the transect length. Students should convert Eq. (5.16) to represent number of seals (N) in the area scanned by the observer and carry out the relevant calculations.

6.0 THE ANALYSIS OF VARIANCE

6.1 One-way Analysis of Variance

We will begin with the simplest case, the one-way analysis of k sets of observations (the t -test considers $k=2$). The ANOVA is usually presented as a table showing the sums of squares, degrees of freedom, mean squares and the associated F -tests of significance (named for the pioneer worker and originator of the test, R. A. Fisher, who devised much of the methodology in the 1920's and 1930's).

The basic calculations depend only on algebraic identities yielding the Sums of Squares. These hold for any k sets of numbers so there are no assumptions involved in the basic calculations. We bring in various assumptions in order to develop statistical tests of significance.

We will subject one set of data (the pheasant count data of Table 6.1) to several different forms of ANOVA to demonstrate the mechanics of calculations for various arrangements of data. The assumptions involved in F -tests will be discussed more fully later, after we examine the basic calculations.

Table 6.1. Pheasant call count data reported by S.M. Carney and G. A. Petrides
Journal of Wildlife Management 21:393, 1957

STATIONS	OBSERVERS					
	A	B	C	D	E	F
1	39	33	33	32	29	27
2	46	36	32	30	35	35
3	45	36	44	31	31	23
4	15	25	29	18	18	14
5	17	14	14	9	14	7
6	27	24	26	14	20	15
7	24	19	15	13	19	15
8	22	22	22	13	16	13
9	28	35	33	32	26	28
10	26	24	23	26	22	17
11	12	13	5	9	8	8
12	8	11	9	9	12	7
13	6	5	9	4	10	3
14	7	6	7	2	7	6
15	7	11	11	6	10	6
16	9	11	9	6	10	4
17	1	2	4	2	5	0
18	5	4	4	4	6	2
19	3	2	3	1	3	1

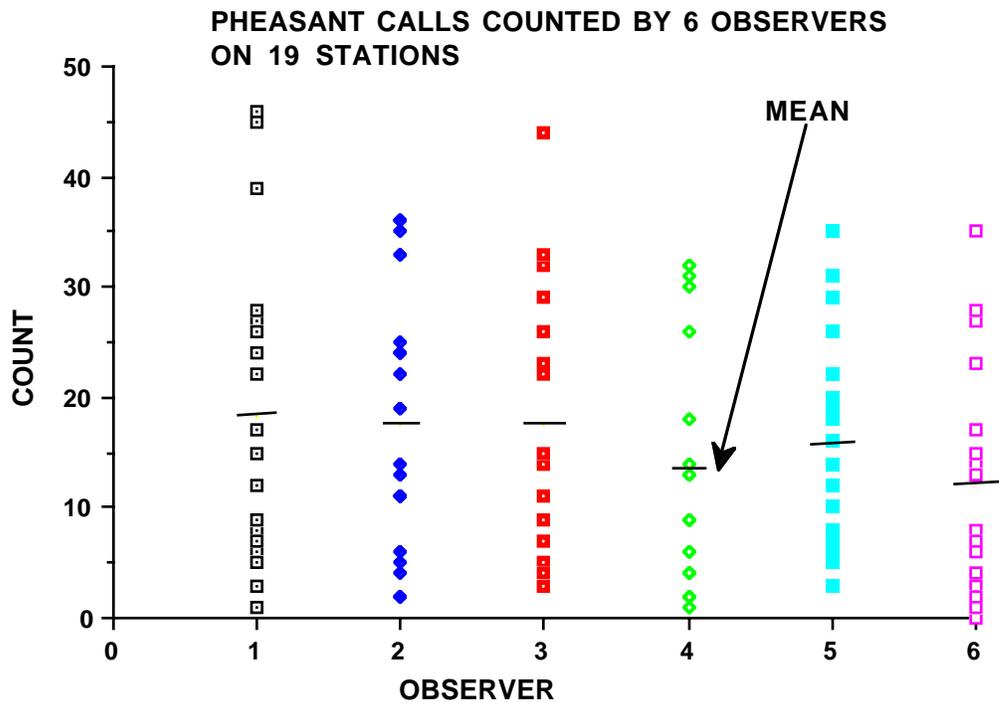


Fig. 6.1 Counts of calling (crowing) pheasants at 19 stations counted simultaneously by each of 6 observers.

The short horizontal lines in Fig. 6.1 mark the mean counts for each of the 6 observers. If it is assumed that the data all come from the same population or process, then the apparent differences in means arise as a matter of chance. Then any particular cluster of points will occupy roughly the same position as any other cluster. On the other hand, if at least some of the populations (or processes) do have quite different means, then the clusters of plotted points will not occupy quite the same positions. Three ways in which the clusters of points can differ are: (1) one or more clusters are shifted up or down from the others (a "scale" or "location" difference), (2) the spread of the individual clusters may differ, and (3) the shape of the clusters may differ. Sample variances provide a measure of the spread of the data, being calculated for the n_i observations from each observer as:

$$s_i^2 = \frac{\sum (y_i - \bar{y}_i)^2}{n_i - 1} \quad (6.1)$$

so that data with a wide spread (scatter) of points will have a large variance. If the clusters have the same shape (this can't be reliably checked without very large samples) and spread (same variance), then a simple shift up or down scale can be detected by comparing the variability of individual clusters with that of the whole set of data. If the clusters are shifted well apart, obviously an overall variance will considerably exceed that of the individual clusters. Offhand, it doesn't look as though the several sets of pheasant data differ much. Another example, based on natural logarithms of counts of "signs"

(mounds and dens) of pocket gophers at different locations appears in Fig. 6.2, which does seem to suggest real differences between sites.

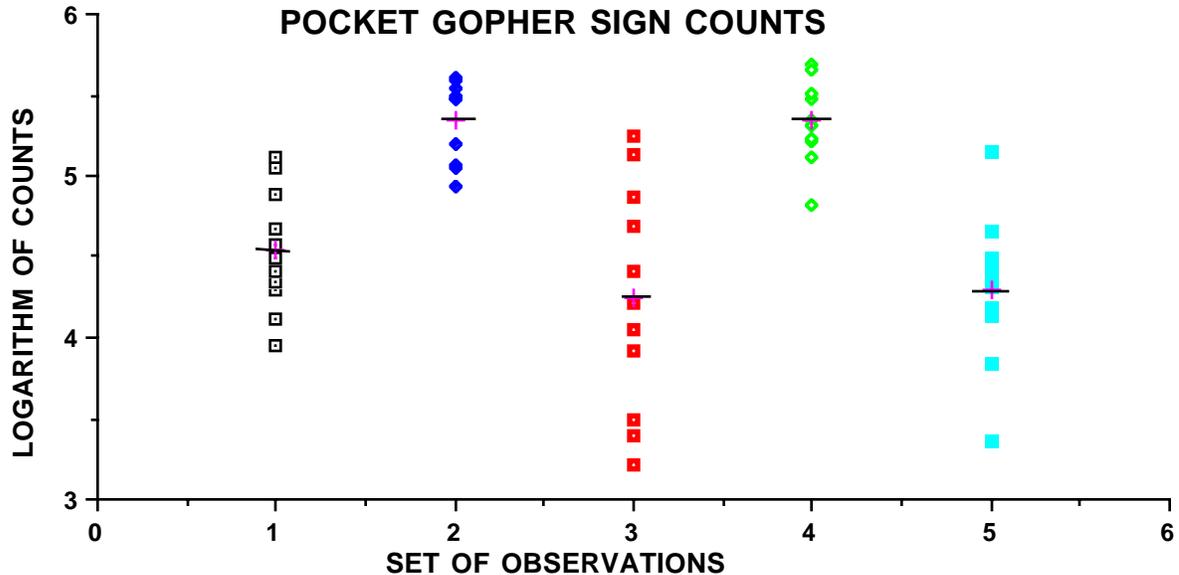


Fig. 6.2 Logarithms of counts of pocket-gopher signs at different locations and/or years (Reid, Hansen, and Ward, *Jour. Wildl. Manage.* 30:330,1966).

A comparison between an overall variance and that of individual comparisons can be constructed by examining the sum of squares making up the overall variance. That is, let y_{ij} be the j^{th} observation in the i^{th} column of tables of data and n_i be the number of observations in that column. Denote the overall mean by \bar{y} and a column mean by \bar{y}_i . Then the overall variance is written as:

$$\frac{\sum_i \sum_j (y_{ij} - \bar{y})^2}{\sum_i n_i - 1} \quad (6.2)$$

Considering only the numerator of eq. (6.2) (the sum of squares) for the present, we can rewrite it as:

$$\begin{aligned} \sum_i \sum_j (y_{ij} - \bar{y})^2 &= \sum_i \sum_j [(y_{ij} - \bar{y}_i) + (\bar{y}_i - \bar{y})]^2 \\ &= \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 + \sum_i \sum_j (\bar{y}_i - \bar{y})^2 \end{aligned} \quad (6.3)$$

$$\sum_i \sum_j (y_{ij} - \bar{y})^2 = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 + \sum_i n_i (\bar{y}_i - \bar{y})^2 \quad (6.4)$$

$$\text{TOTAL S.S.} = \text{WITHIN S.S.} + \text{BETWEEN S.S.}$$

This results because a little algebra shows that the cross-product term vanishes (students should do the algebra for themselves). The first of the two resulting terms is just the sum of the components that would be used to calculate a separate variance for each column and is thus denoted the "within"

(within columns) sum of squares. The second component represents the variability "between" columns. These quantities are usually displayed in an Analysis of Variance (ANOVA) table (let $\sum n_i = n$):

Table 6.2 Analysis of Variance for a one-way design.

<u>Source</u>	<u>Sum of squares</u>	<u>Degrees of freedom</u>	<u>Mean squares</u>
Between groups	$SSb = \sum_i n_i (\bar{y}_i - \bar{y})^2$	k - 1	$\frac{SSb}{k - 1}$
Within groups	$SSw = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2$	n - k	$\frac{SSw}{n - k}$
Total	$= \sum_i \sum_j (y_{ij} - \bar{y})^2$	n - 1	

The "mean squares" (MS) are estimates of variances, and under the hypothesis of no difference between the populations (processes) represented by the columns of the Figures above, these estimates should be equal. Arriving at the divisors (degrees of freedom) can be remembered by the following devices: (1) there are k means being considered in the "between" groups so the usual practice for estimating a variance prevails, i.e., divide by k-1, (2) within each group a variance would be estimated by Eq. (6.1). A logical way to pool these within-group variances is to weight by the degrees of freedom, i.e., calculate:

$$\frac{\sum_i (n_i - 1) s_i^2}{\sum_i (n_i - 1)} \quad (6.5)$$

which gives the between-groups value used above (Table 6.2).

Whether the two variance estimates are significantly different or not is tested by the "F-ratio", which is:

$$F = \frac{SSb/(k-1)}{SSw/(n-k)}$$

Values denoting significant deviations are widely tabulated in textbooks in statistics and are now printed out by the various computer programs used to calculate ANOVAS. The advent of such computer programs has made it very easy to do the calculations. The serious disadvantage of these "canned" programs is that virtually anyone can calculate complex analyses without having any real idea what the results mean. Students thus need to actually work out the calculations for the above examples so as to understand how they are carried out. This is easy to do on a spreadsheet, such as EXCEL.

Inasmuch as EXCEL will conduct one-way ANOVAS, we can first use that function (Anova: single factor) and then calculate the sums of squares directly

on a separate spreadsheet, as a way to understand what's going on. Thus the EXCEL one-way program produces a listing of sample sizes, sums, averages and variances for each column in the table of pheasant call-count data, followed by an ANOVA table of Sums of Squares, degrees of freedom, mean squares, F value, P-value (probability of significant difference between groups, and "F-crit" (the significant value of F at the $\alpha = 0.05$ level).

To check these results directly, one needs only to insert two columns between each of the existing columns of data, calculate column means (\bar{y}_j) for the data and the overall mean (\bar{y}), and use these to calculate "within" and "total" sums of squares and add them up to get the values produced by the program. The "between" sum of squares is calculated directly from the definition given in Table 6.2 above using column means and overall mean.

6.2 Two-way analysis of variance

One-way ANOVA usually does not involve much in the way of a study design. The comparisons are likely to be obvious, and the only complication that may arise is if it is desired to compare subgroups of the k sets of observations. We will return to such comparisons later on. The "higher-order" forms of ANOVA are more versatile and thus more powerful. More planning is thus involved, and we need to distinguish between the various possible approaches. The simplest of the more complex ANOVA's is the two-way analysis without replications. As the name suggests, it is based on a two-way table. There are k sets of data, each appearing in r rows, so that there are rk observations. The pheasant call-count data provide an example, where we now consider the rows (stations) as a factor in the analysis. This is done by calculating a row sum of squares, and incorporating it in the ANOVA table. It is worthwhile to depict the data as an table of x_{ij} with k columns and r rows as follows (some authors use r rows and c columns; others a rows and b columns - - notation is not consistent in statistics books). It is useful to border the table with row and column means. The dot notation (e.g, $\bar{x}_{.1}$) is used to signify that the average is taken over a row or a column ($\bar{x}_{.1}$). A double dot notation ($\bar{x}_{..}$) is used to designate the overall means (sometimes this appears with two bars over x). **Note that we have switched from y_{ij} to x_{ij} .** Both notations are common; it is worthwhile to use x_{ij} from now on because y_{ij} will be used as the "independent" variable in regression analysis later on.

		COLUMN MAIN EFFECT						
		1	2	3	i	k		
	1	x_{11}	x_{21}	x_{31} ...	x_{j1} ...	x_{k1}	$\bar{x} \cdot 1$	
	2	x_{12}	x_{22}	x_{32}	x_{j2}	x_{k2}	$\bar{x} \cdot 2$	
ROW EFFECT	3	x_{13}	x_{23}	x_{33}	x_{j3}	x_{k3}	$\bar{x} \cdot 3$	
		.						
		.						
	j	x_{1j}	x_{2j}	x_{3j}	x_{ij}	x_{ki}	$\bar{x} \cdot i$	
		.						
		.						
	r	x_{1r}	x_{2r}	x_{3r}	x_{ir}	x_{kr}	$\bar{x} \cdot r$	
		$\bar{x} \cdot 1$	$\bar{x} \cdot 2$	$\bar{x} \cdot 3$	$\bar{x} \cdot i$	$\bar{x} \cdot k$	$\bar{x} \cdot \cdot$	

The sums of squares (S.S.) are obtained in the same way as in the previous example, that is, we expand the Total S.S. to form the other sums of squares:

$$\text{Total S.S.} - \text{Columns S.S.} - \text{Rows S.S.} = \text{Residuals (Error) S.S.}$$

$$\sum_{i=1}^k \sum_{j=1}^r (x_{ij} - \bar{x}_{\cdot\cdot})^2 - r \sum_{j=1}^r (x_{i\cdot} - \bar{x}_{\cdot\cdot})^2 - k \sum_{j=1}^r (\bar{x}_{\cdot j} - \bar{x}_{\cdot\cdot})^2 = \sum_{i=1}^k \sum_{j=1}^r (x_{ij} - \bar{x}_{i\cdot} - x_{\cdot j} + \bar{x}_{\cdot\cdot})^2 \quad (6.6)$$

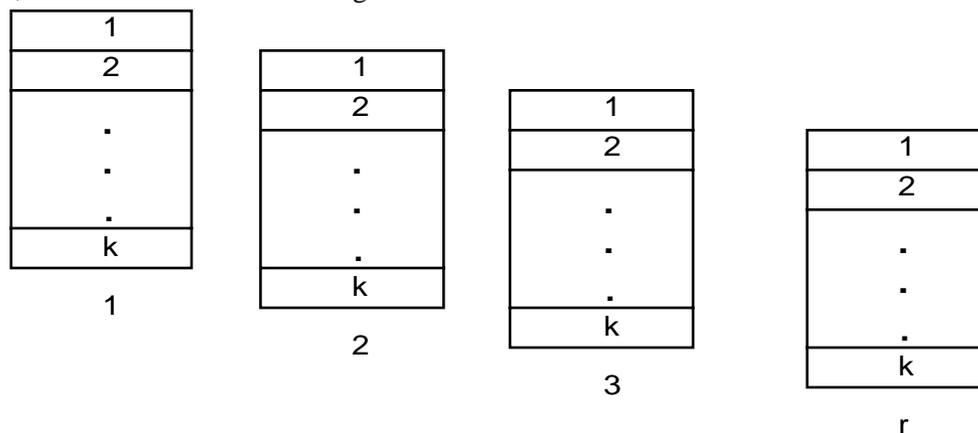
These results are calculated by EXCEL as 2-way ANOVA without replication. The program produces tables of row and column means and variances along with an ANOVA table.

6.3 Randomized blocks designs

The two-way program is listed in EXCEL as being "without replication". However, this is not necessarily true, as the row effects can indeed represent replications. Such an arrangement results from a randomized blocks design. These designs are widely applicable. Suppose we have k treatments to study, and can arrange to test them in r "blocks", where each block is comprised of k units that are relatively uniform in nature. For example, we might want to evaluate the effectiveness of k drugs on weight gain in rats. We might thus obtain r litters of k rats each, and give the different drugs to each of the k rats of each litter (choosing individual rats out of a given litter at random to receive one of the k drugs), and maintain the individual litters together under uniform conditions. The trick is to keep the blocks as uniform as possible so as to minimize "within block" variability so that most of the variance within a block results from the treatments. The method was developed in agricultural experimentation where the blocks are usually plots of ground selected for

their uniformity. Each plot is made up of k subplots, to which some set of, say, fertilizers, is applied. Fig. 6.3 shows how randomized blocks designs are laid out in plots. Note that the blocks may be separated by some distance, being selected for the uniformity of material within a block, which can reduce the "error" M.S. considerably.

The randomized blocks design can be a powerful and efficient approach. Note that the individual units in the blocks serve as true replicates so that the randomized blocks design does have replication. In our pheasant example, the stations are not replicates, so the ANOVA there is "without replication". However, the two cases (with and without replication) use the same calculations. The difference lies in the experimental design -- randomized blocks may be far more efficient in assessing differences. Much of the efficiency depends on the investigator's knowledge of the experimental material -- there is an element of "art" in picking blocks. In long-term studies one can sometimes take advantage of previous year's data to see how uniform the blocks are. Also, "uniformity" studies can be run to measure the variability within blocks. In these studies the same "treatment" (usually no treatment) is applied to all plots, and the ANOVA run to measure variability within and between blocks. One would, of course, like to have a very small "within" mean square, and can tolerate a large "between" blocks M.S.



A randomized block study design. There are k units, assigned to a location at random within each block, and r blocks in all. Every treatment appears in each block (randomly assigned to a position).

Fig. 6.3 Randomized blocks layout. The blocks (often plots in agricultural studies) are laid out to be as uniform as possible within individual plots.

6.4 Two-way analysis of variance with replication

The two-way analysis of variance with replication normally appears with replicates "within cells" in a table of data. We thus need to consider observations with three subscripts, x_{ijk} , as shown in the table below, which has 3 replicates per cell. In general, we may have m replicates per cell where $m \geq 2$, and thus rkm observations in the entire table.

The calculations for S. S. in the ANOVA table now become somewhat more complicated, but take on a general form that can be followed in even more complex cases. The residual S. S. (error term) is always calculated from the replicates within cells, i.e.,

$$\text{Residual S.S.} = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij\bullet})^2 \quad (6.7)$$

where $\bar{x}_{ij\bullet}$ is the cell mean (these are not shown in the table below as they are an average of the m observations in the cell; 3 observations in the table above). An easy way to remember how to calculate residual (error) mean squares when there is replication, is to note that the units within a given cell all get identical treatments, and thus furnish the best estimate of the underlying variability. Hence the error mean square estimates the underlying variance of the experimental units. Any other variance estimate (mean square) may be inflated by treatment effects.

COLUMN MAIN EFFECT

		1	2	3	...	i	...	k	
		x ₁₁₁	x ₂₁₁	x ₃₁₁	...	x _{j11}	...	x _{k11}	
	1	x ₁₁₂	x ₂₁₂	x ₃₁₂	...	x _{j12}	...	x _{k12}	$\bar{x}_{\cdot 1\cdot}$
		x ₁₁₃	x ₂₁₃	x ₃₁₃	...	x _{j13}	...	x _{k13}	
		x ₁₂₁	x ₂₂₁	x ₃₂₁	...	x _{i21}	...	x _{k21}	
	2	x ₁₂₂	x ₂₂₂	x ₃₂₂	...	x _{i22}	...	x _{k22}	$\bar{x}_{\cdot 2\cdot}$
		x ₁₂₃	x ₂₂₃	x ₃₂₃	...	x _{i23}	...	x _{k23}	
	•	•	•	•	...	•	...	•	
	•	•	•	•	...	•	...	•	
	•	•	•	•	...	•	...	•	
		x _{1j1}	x _{2j1}	x _{3j1}	...	x _{ij1}	...	x _{kj1}	
ROW EFFECT	j	x _{1j2}	x _{2j2}	x _{3j2}	...	x _{ij2}	...	x _{kj2}	$\bar{x}_{\cdot j\cdot}$
		x _{1j3}	x _{2j3}	x _{3j3}	...	x _{ij3}	...	x _{kj3}	
	•	•	•	•	...	•	...	•	
	•	•	•	•	...	•	...	•	
	•	•	•	•	...	•	...	•	
		x _{1r1}	x _{2r1}	x _{3r1}	...	x _{ir1}	...	x _{kr1}	
	r	x _{1r2}	x _{2r2}	x _{3r2}	...	x _{ir2}	...	x _{kr2}	$\bar{x}_{\cdot r\cdot}$
		x _{1r3}	x _{2r3}	x _{3r3}	...	x _{ir3}	...	x _{kr3}	
		x _{1..}	x _{2..}	x _{3..}	...	x _{i..}	...	x _{k..}	x _{...}

Another very general S. S. is the Total sum of squares, calculated from the individual observations as before:

$$\text{Total S.S.} = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{...})^2 \quad (6.8)$$

with $\bar{x}_{...}$ the overall mean.

The third general S. S. is the Treatment Sum of Squares, calculated from the cell means

$$\text{Treatment S.S.} = m \sum_i \sum_k (x_{ij\bullet} - \bar{x}_{...})^2 \quad (6.9)$$

If any of the treatments are effective, the treatments mean square will be inflated. Of course, we want to be able to break this overall S.S. down into row and column S. S. . As before, we do this with row and column means:

$$\text{Row S. S.} = mr \sum_{i=1}^k (\bar{x}_{i\bullet\bullet} - \bar{x}_{...})^2 \quad (6.10)$$

$$\text{Column S.S.} = mk \sum_{j=1}^r (x_{\bullet j\bullet} - \bar{x}_{...})^2 \quad (6.11)$$

The two-way analysis with replication contains a new S. S., the interaction Sum of Squares. This is often calculated as Treatment S. S. - Row S. S. - Column S. S. but a direct calculation from the means is

$$\text{Interaction S.S.} = m \sum_{i=1}^k \sum_{j=1}^r (\bar{x}_{ij\bullet} - \bar{x}_{i\bullet\bullet} - \bar{x}_{\bullet j\bullet} + \bar{x}_{...})^2 \quad (6.12)$$

The Total S.S. breaks down into Treatments and Error, and the Treatments S.S. contains Rows, Columns and Interaction S.S. Textbooks usually show the ANOVA table in this form, but EXCEL ignores Treatments, producing only Total, Rows (labelled Samples for unknown reasons), Columns, and Interaction.

In the final ANOVA, the F-test of significance of the interaction mean square ($MS_{\text{Inter}}/MS_{\text{Error}}$) is very important in deciding what can be said about the main effects. This is because a significant interaction mean square suggests that the row and column main effects are somehow correlated, i.e., they "interact". If this is the case, then one cannot discuss the two sets of main effects (row and column) separately, making interpretation of the experiment much more difficult.

Note the similarity of the equation for Interaction S. S. to that for the Residuals (error) S. S. for the 2-way ANOVA without replication. This suggests that the error term in that case is really an interaction term, making it evident that we need to have replications to assess interaction (there is a test, Tukey's test, for a particular form of interaction in a 2-way ANOVA based on one observation per cell. It appears in many statistics texts (e.g., Snedecor and Cochran, Statistical Methods, Iowa State University Press, Ames, Iowa, 6th Edition, 1967).

The pheasant count data of Table 1 were not replicated in a strict sense, which would require repeated counts at the same station by the same observers. One of the reasons this is not done is that calling activity drops off quite sharply after the early morning hours. It is also obvious from Table 1 that there is a gradient over distance. High counts are obtained in the best pheasant habitat, and this route apparently went into marginal habitat beyond Station 10. Inasmuch as the stations are reasonably close together (usually 1 mile apart to keep from counting the same calling individuals a second time), it isn't too much of a stretch of technique to regard adjacent stations as "replicates". To do this, we drop the counts on Station 19, and use successive pairs as replicates (thus counts on Stations 1 and 2 are called replicates, while Stations 3 & 4 are also replicates, etc., giving $r = 9$, while $k = 6$ for $rkm = 9(6)2 = 108$). The analysis of variance table is as follows.

Table 6.3 Analysis of variance of Pheasant count data with $m=2$.

Source	SS	df	MS	F
Rows	11189.91	8	1398.74	52.24
Columns	568.41	5	113.68	4.25
Interaction	605.76	40	15.14	0.57
Within	1446.00	54	26.78	
Total	13810.07			

6.5 Assumptions for the analysis of variance

The discussion thus far has focussed on the mechanics of the analysis of variance, being mainly concerned with developing Sums of Squares for 3 models: (1) one-way ANOVA, (2) 2-way ANOVA without replication (and the special case of randomized blocks designs), and (3) two-way ANOVA with replication. The ANOVA tables present the S. S., their associated Mean Squares, and the ratios of Mean-Squares (F-ratios). As previously noted, the tests of significance (F-tests) depend on the assumptions underlying ANOVA, but the underlying framework - the Sums of Squares, along with the Mean Squares and F-ratios, can be calculated for any set of numbers. No assumptions are required. We thus have a mechanical analysis that says something about variability introduced by treatments, without assumptions about the underlying data.

To consider the assumptions required for tests of significance, we write a model for the observations for a 2-way ANOVA with replications:

$$x_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ij} \quad (6.13)$$

Here, μ represents an overall mean value, α_i and β_j are the main effects (column and row effects), γ_{ij} represents the interaction between the two main effects, and ε_{ij} is the error term. This latter term (ε_{ij}) is assumed to have an "expected value" of zero. That is, when it is averaged over a large data set it should equal zero. Usually it is assumed that $\sum \alpha_i = 0$ and $\sum \beta_j = 0$. We thus have the

x_{ij} made up of an overall mean value (μ) plus an effect for its row (α_i) and its column (β_j). As we noted earlier, an "interaction" is an effect that makes adjacent observations tend to be correlated. When there are no interactions ($\gamma_{ij} = 0$) then the expected value of x_{ij} (effectively x_{ij} averaged over very large samples) can be written as $E(x_{ij}) = \mu + \alpha_i + \beta_j$ and we say that the model is additive. Such analyses are far easier to understand and interpret than are those where interactions are present (γ_{ij} not equal to zero).

It is important to recall that we want to test several hypotheses that state that the main effects and interaction are zero, and the assumptions become important in assuring validity of the F-tests.

Assuming additivity, we can use the reduced model:

$$x_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad (6.14)$$

Two major assumptions underlying tests of significance in ANOVA are:

- (1) the ε_{ij} are independent, i.e., uncorrelated.
- (2) the ε_{ij} are from a normal distribution with mean zero and variance σ^2 . The normal distribution is a symmetrical, bell-shaped curve, with its "spread" (variance) measured by σ^2 [Eq.(1.3)].

Consider a two-way ANOVA with many replications per cell (and no interactions). The mean value in any cell should be approximately

$$\bar{x}_{ij} = \mu + \alpha_i + \beta_j$$

and the x_{ijk} in this cell should have the same variance, σ^2 . Any two cells should have the same variance, σ^2 . This is often described as homoscedasticity, which simply means equal variances.

The assumptions for ANOVA can be listed as:

- (1) additivity ($\gamma_{ij} = 0$).
- (2) independence of the ε_{ij}
- (3) ε_{ij} normally distributed with mean zero and variance σ^2 .

Sometimes (3) is split into 2 assumptions:

- (3) ε_{ij} normally distributed with mean zero
- (4) homoscedasticity - variances in replicates are all equal to σ^2 .

In most applications of ANOVA there simply are not enough replicates within cells to test these assumptions. Given quite large samples in the cells (say 20-30 replicates per cell) it is worth comparing variances. Testing for

normality takes larger samples. Some authors recommend Bartlett's test, but Scheffe (The Analysis of Variance, J. Wiley and Sons, 1959, p. 83) points out that it "is extremely sensitive to nonnormality", and recommends that a preliminary test of homogeneity of variances not be made.

Ecological data often come as counts of some kind, and these tend not to be normally distributed, often having a skewed frequency distribution -- a long "tail" of less frequent observations on one side or the other of the bulk of the observations. Such data can be brought into closer approximation to normality by a transformation. Two of the most commonly used transformations are the square root transformation $(x_{ij})^{0.5}$, and the logarithmic transformation, $\log_e(x_{ij})$. It often turns out that standard deviations of ecological data tend to be proportional to the mean values (coefficient of variation, s/\bar{x} = approximately a constant). The logarithmic transformation tends to "normalize" such data and to make variances more nearly equal on the transformed (i.e., logarithmic scale).

Testing the need for or the effects of a transformation is often recommended, but it is risky to let such tests govern a decision to use or not use a transformation.

It is worthwhile to simulate data based on the assumptions for ANOVA. We start with Eq.(6.14); no interactions, and produce a table of main effects (using α_i and β_j such that they sum to zero) to which we add μ (taken as 5 here). A table of random normal deviates can be produced using EXCEL (used here as $N(0,1)$, i.e. normal with zero mean and unit variance). These are added to the main effects table giving a set of simulated data (3 replicates per cell were used; they have the same main effect value, but different random draws were used to add ϵ_{ij}). The table of data follows, along with the ANOVA table.

		1	2	3	4
	1	6.10	5.89	4.98	5.67
		7.18	4.80	5.51	5.70
DATA		5.73	6.71	4.89	6.80
FOR	2	6.33	6.17	6.75	2.91
ANOVA		6.48	5.89	2.54	4.13
		6.77	4.42	5.56	5.23
	3	6.57	5.36	5.93	3.71
		6.93	4.93	3.08	4.80
		5.68	5.42	2.81	5.31
	4	6.49	4.14	5.80	2.65
		3.39	4.74	5.26	5.82
		4.49	6.39	4.74	3.89
	5	4.18	5.97	5.14	2.61
		6.55	5.18	4.46	3.11

		4.10	3.99	3.93	3.57	
	6	5.69	4.40	4.20	4.25	
		4.15	3.63	4.18	4.78	
		3.97	2.88	3.75	4.49	
Source	SS	df	MS	F	P-value	F crit
Sample	21.20	5.00	4.24	4.08	0.00	2.41
Columns	14.67	3.00	4.89	4.71	0.01	2.80
Interaction	18.29	15.00	1.22	1.17	0.32	1.88
Within	49.86	48	1.04			
Total	104.02	71				

We can repeat the above exercise with interactions. In this case, the γ_{ij} were taken as a fractional power of the product denoting row and column positions of a main effect entry, $\gamma_{ij} = (xy)^{0.7}$. Again a random normal deviate is added to give eq. (6.13). The table of "data" and ANOVA table follow.

		1	2	3	4
DATA	1	7.10	7.51	7.14	8.30
FOR		8.18	6.42	7.67	8.34
ANOVA		6.73	8.33	7.05	9.43
	2	7.96	8.81	10.25	7.20
		8.10	8.53	6.04	8.42
		8.40	7.06	9.07	9.51
	3	8.73	8.87	10.59	9.40
		9.09	8.44	7.73	10.50
		7.84	8.93	7.47	11.00
	4	9.13	8.43	11.49	9.62
		6.03	9.03	10.96	12.79
		7.13	10.67	10.44	10.85
	5	7.27	10.98	11.79	10.75
		9.63	10.19	11.12	11.25
		7.19	9.00	10.58	11.71
	6	9.19	10.09	11.77	13.50
		7.65	9.33	11.75	14.03
		7.47	8.57	11.31	13.74

Source of Variation	SS	df	MS	F	P-value
Sample	78.15	5	15.63	15.05	0.0000
Columns	69.00	3	23.00	22.14	0.0000
Interaction	44.39	15	2.96	2.85	0.0030
Within	49.86	48	1.04		
Total	241.40	71			

As noted previously, some authors recommend testing the deviations (given in Eq.(6.12)) for normality. With the numbers of replicates usually

available, this is not sensible advice. A plot (Fig. 6.4) of all 72 normal deviates used in the simulations was generated from a normal distribution, but is not too reassuring in terms of the assumption of normally distributed errors.

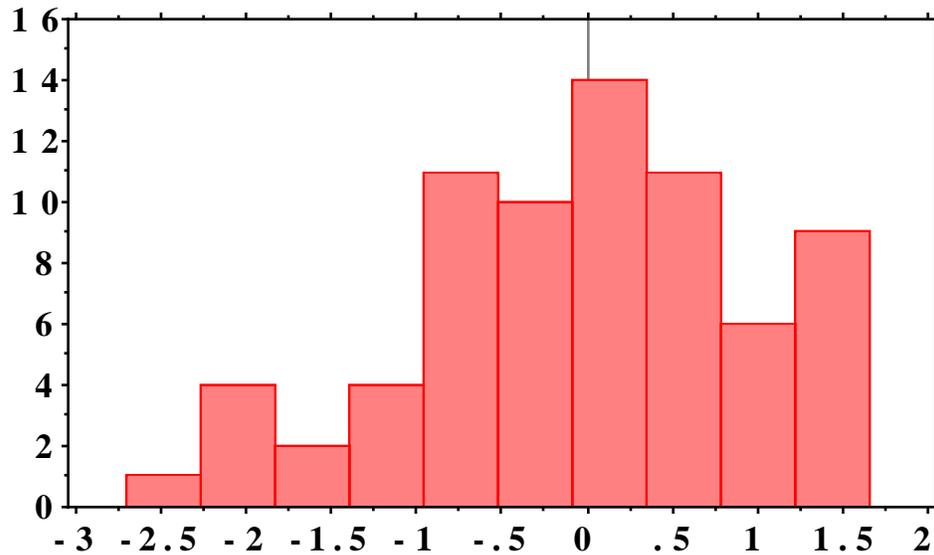


Fig. 6.4. Frequency distribution of 72 deviations from data used in simulations.

6.6 Comparisons in ANOVA

In a brief account like the present one, it is not possible to cover more than a fraction of the features of the Analysis of Variance. The book by Scheffe is a classic account, and should be examined for more details. It is, however, couched in the language of matrix algebra. Another good account is that of Snedecor and Cochran, *Statistical Methods*, the Iowa State University Press, Ames, Iowa. It has gone through at least 8 editions, and is another classic text. Important topics that we have not covered are those of comparisons or contrasts. In most experimental work, the main interest will be in certain comparisons (e.g. Exercise 6.12 on weight gains in rats). In the pheasant data there were 2 observers with experience, so a comparison between experienced and inexperienced observers is of considerable importance. The pocket-gopher data (Exercise 6.5) was collected at different locations and over different years. One would thus emphasize those comparisons. In Exercise 6.8 there is a "control" plot ("check" treatment) which would normally be compared with all other plots. Snedecor and Cochran give good descriptions of how to sort out such contrasts. A short account of two approaches follows.

When comparisons are planned in advance, a t-test can be used to test these specific comparisons for significance. The test depends on computing a linear combination of the observed means:

$$L = \lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 + \dots + \lambda_k \bar{x}_k \quad (6.15)$$

with the λ_i constants adding to zero, i.e. $\sum \lambda_i = 0$. The standard error of L is estimated as:

$$S.E.(L) = s_L = \sqrt{\sum \lambda_i^2} \frac{s}{\sqrt{n}} \quad (6.16)$$

The d.f. for the estimated standard error of L are those used to estimate s , and n is the number of observations used to compute each mean, \bar{x}_i . Scheffe describes comparisons such as L as contrasts. The t-test for comparisons planned in advance is:

$$t = \frac{L}{s_L} \quad (6.17)$$

The λ_i are dictated by the comparison desired. If, for example, 2 means, say \bar{x}_1 ,

and \bar{x}_3 , are compared to a third one, \bar{x}_2 then $L = \frac{\bar{x}_1}{2} + \frac{\bar{x}_3}{2} - \bar{x}_2$ with the λ_i being 1/2, 1/2 and -1, and thus adding to zero. If there are additional means in the overall analysis that are not to be involved in the comparison, then the λ_i for those means are assumed to be equal to zero. The simplest comparison is that for 2 means, with the comparison being $\bar{x}_1 - \bar{x}_2$, so that $\lambda_1 = 1$ and $\lambda_2 = -1$, and

$s_L = 2^{1/2}s/n^{1/2}$ so that:

$$t = \frac{n^{1/2}(\bar{x}_1 - \bar{x}_2)}{2^{1/2}s}$$

with n being the number in each group and s is obtained from the error M.S.

A test due to Scheffe provides a general method for finding significant differences among a full set of means without designating these comparisons in advance of conducting the experiment. The price paid is less sensitivity (broader confidence limits). It uses the same set-up as above, but declares L/s_L significant only if it exceeds $[(k-1)F_{05}]^{1/2}$ where F_{05} is the 5% level of the F-distribution for $k-1$ and $n-k$ degrees of freedom when we are considering a one-way analysis. The test can be used in more complex ANOVAs using $(k-1)$ and the d.f. associated with the error mean square. Scheffe's test also reduces to the t-test when $k = 2$. The test should not be used if the F-test in an ANOVA is not significant, as there are then no significant contrasts in the data. It is important to understand that the S-method can be used to check all significant contrasts in the means, while preserving the chosen α level.

We illustrate the two procedures by using data simulated as in Section 6.5 where the simulations were used to study the assumptions for ANOVA. In this example the column main effects have been changed. The "data" are as follows:

	1	2	3	4
1	5.90	5.39	5.48	5.87
	6.98	4.30	6.01	5.90
	5.53	6.21	5.39	7.00
2	6.13	5.67	7.25	3.11
	6.28	5.39	3.04	4.33
	6.57	3.92	6.06	5.43
3	6.37	4.86	6.43	3.91
	6.73	4.43	3.58	5.00
	5.48	4.92	3.31	5.51
4	6.29	3.64	6.30	2.85
	3.19	4.24	5.76	6.02
	4.29	5.89	5.24	4.09
5	3.98	5.47	5.64	2.81
	6.35	4.68	4.96	3.31
	3.90	3.49	4.43	3.77
6	5.49	3.90	4.70	4.45
	3.95	3.13	4.68	4.98
	3.77	2.38	4.25	4.69
Means	5.3995	4.5498	5.1398	4.6118

The ANOVA (two-way with replications) is:

Source	SS	df	MS	F	P-value
Sample	21.1987	5	4.2397	4.0818	0.0036
Columns	9.1840	3	3.0613	2.9473	0.0421
Interaction	18.2903	15	1.2194	1.1739	0.3233
Within	49.8576	48	1.0387		
Total	98.5306	71			

If we suppose the planned comparison was between means 1 and 3 against means 2 and 4, then:

$$L = 0.25(5.3995) - 0.25(4.5498) + 0.25(5.1398) - 0.25(4.6118) = 0.3444$$

$$\text{and: } s_L = (\sum \lambda_i^2)^{1/2} \frac{s}{n^{1/2}} = 0.5[(1.0387)/(18)]^{1/2} = 0.122$$

$$\text{then: } t = \frac{L}{s_L} = \frac{0.3444}{0.122} = 2.82 \text{ with 48 d.f.}$$

From t-tables ($\alpha = 0.05$) we have $0.005 < P < 0.010$. EXCEL has a function that will compute the probability directly. Enter the statement = TDIST(t,d.f.,tails) where t is the calculated value (2.82 here), d.f. are 48, and "tails" is 2 for a 2-tailed test. This function yields $P = 0.007$. Quite possibly past experience would lead to a one-tailed test for a planned comparison.

We can illustrate Scheffe's S-method for the same comparison. The t-value remains the same (2.82) but we use:

$$[(k-1)F_{05}]^{1/2} = [3(2.80)]^{1/2} = 2.90$$

as the criterion for significance at the 5% level, where 2.80 is the tabulated F-value at $\alpha = 0.05$ with 3 and 48 d.f. Hence, the test is close to the 5% level of significance. We can go on and look for other significant contrasts as suggested by the data while still having $\alpha = 0.05$. This is definitely not the case for the first comparison tested above, which has to be selected in advance of the study. Because it is a general-purpose "data-snooping" tool, Scheffe (1959) suggested his test might be used with $\alpha = 0.10$, rather than the usual 0.05. EXCEL can be used to find the tabular F-value by using the function FINV(P, d.f.1, d.f.2) where P = 0.05 here and, d.f.1 = 3, and d.f.2 = 48. This function gives $F_{05} = 2.798$, and $F_{10} = 2.201$.

6.7 Type I and II errors and "power"

Most ecologists are used to the notion of Type I error, routinely conducting statistical tests, such as the t-test, at the 5% level of significance ($\alpha = 0.05$). They understand that such tests give a 0.05 probability, over the long run of many such tests, of erroneously claiming that the null hypothesis of "no effect" can be rejected when it is in fact true. Many do not seem to realize that there is another side to the issue, which is failing to find a significant difference when it exists (possibly because there were not enough samples or replications to detect an important difference. This is known as a type II error.

This issue of type II error can be discussed in terms of the "sensitivity" of a study, i.e., how small a change or difference will a study of a given size reliably detect? The statistician's answer is usually couched in terms of a power function or the "power of a test". Consider the likely points of view as to the impact on the environment of some new facility. There are usually two sides, those who construct and operate the facility and those with environmental regulatory authority. Suppose that both sides can agree on a study method that has well-known statistical properties and can be applied in the circumstances under consideration. Suppose that they further agree to make certain modifications in the facility and/or operational procedures if a field survey shows a specified degree of change has taken place (an "impact"). What remains is to decide how large a sample should be taken in the field study. But that depends on the amount of protection each party requires against errors damaging to their best interests.

These can be described as follows: (1) The people doing the construction and operation would rather not have the survey results indicate a significant change when the agreed-on degree of change really did not take place (Type I error). Just how strongly they voice objections will depend on the consequences of a determination of "change". If only minor modifications are then necessary they perhaps will agree that a 10% rate ($\alpha = 0.10$) of such "false positive" results is acceptable. However, if the changes require costly retrofitting and expensive operational modifications, they may well want to try to insist on Type I error rates of 1% or maybe even less.

(2) On the other side, the staff of the regulatory agency would not like to fail to recognize a significant impact when one does occur (Type II error). If small

samples are taken, the results almost always will come out not significantly different. Hence the regulators may be guided by rules that require an 80% chance of being sure to detect a real difference (of the magnitude agreed on) when the impact is not of minor environmental consequence. But if very substantial damage to an important resource may be involved, they may well argue for a 99% assurance. All too often, by default or lack of understanding the actual rate may be about 50%, much like settling the issue by flipping a coin and doing no field work.

To make any progress in the ensuing arguments a way of estimating sample sizes is needed. An easy solution is just to take a very large sample. Usually that is either too expensive or impractical (it may also result in environmental damage from the sampling process). A handy formula for approximating sample size for given Type I and II errors is given by Snedecor and Cochran and in the useful book by Cochran (Planning and Analysis of Observational Studies, W. G. Cochran 1983, J. Wiley and Sons). It can be written as:

$$n = \frac{2(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2} = \frac{2(z_\alpha + z_\beta)^2}{\left\{\frac{\delta}{\sigma}\right\}^2} \quad (6.18)$$

where z_α = normal deviate for Type I error, z_β = normal deviate for Type II error, σ^2 = variance (assumed the same in both data sets being compared), and δ = true (unknown) difference between two population means ($\mu_1 - \mu_2$) or two areas being studied, and n = the desired sample size for each population or area (thus $2n$ required). z_α is the familiar value used in confidence limits, i.e., $z_{05} = 1.96$, $z_{10} = 1.64$. Some values of z_β are:

Type II error (β)	Power ($1 - \beta$)	z_β
0.20	0.80	0.84
0.10	0.90	1.28
0.05	0.95	1.64
0.01	0.99	2.33

A major problem is that σ is always unknown, and must be either guessed at or estimated from a preliminary survey (etc.). Thus the right-hand side of eq. (6.18) is frequently used, i.e., one guesses at the ratio of the difference to be detected to σ . Suppose we take $z_\alpha = 1.96$ and $z_\beta = 1.28$ (power = 0.90). Then

$$n = \frac{2(1.96 + 1.28)^2}{\left\{\frac{\delta}{\sigma}\right\}^2} = \frac{21}{\left\{\frac{\delta}{\sigma}\right\}^2}$$

Consequently, if we suppose that the true difference is one-half of σ , $n = 84$, while $n = 21$ if we assume $\sigma = \delta$. Clearly, if we assume a small difference is to be detected, the sample size required may be huge. Using a small sample without

looking at power of the test amounts to operating in ignorance (but still happens a lot).

6.8 Other aspects of ANOVA

We did not go beyond two-way tables with replications. Efficient designs will use more factors in order to get the most information per dollar spent on experimentation. Again Snedecor and Cochran provide good discussions. There are text-books devoted entirely to the design and analysis of experiments, and a bewildering array of prospects. We also did not investigate what are called unbalanced analyses. These are typically two-way analyses with replications where the same number of replications per cell is not present. Sometimes a study is planned with m replicates per cell but some are destroyed or, in the case of experiments with animals, die unexpectedly, etc. In other cases, it may not be possible to get m replicates per cell. Analyses of unbalanced designs can be complex and, in some cases, controversial. It may be noted that the pocket-gopher example is unbalanced, but this is not a problem in one-way analyses.

The models described here are fixed-effects models, where interest is solely in the set of main effects studied in the experiment. Very often we have to consider random-effects models where the effect studied is regarded as a sample from some large population of effects. The analysis then takes a different form. Probably most practical work can be described by mixed-effects models, where one set of factors is fixed and the remainder random. The great advantage of the fixed-effects model is that the analysis of variance is quite "robust" in such cases, i.e., non-normality is not as serious a concern as in the random-effects models, where we assume sampling from a random normal distribution, and depend much more on that assumption for tests of significance. We remarked that significant interactions pose problems of interpretation, but did not note that it may be necessary to use the interaction term as the denominator of F-tests.

The prominence of ANOVA in ecological studies is a bit puzzling. A quick review of a major ecological journal a few years back showed that ANOVA was then the dominant statistical technique used in that publication. However, I suspect that many of the cases really stem from editorial and reviewer insistence on statistical testing. The mere fact that some "significant" result was obtained doesn't really provide much information about the process being studied. Hence I suggest that students use the data-snooping quality of the test due to Scheffe (Section 6.6) whenever possible as a tool for searching out the particular comparisons that really are significant in an analysis. As noted there, Scheffe proposed using a 10% significance level with the test, but that may not be very palatable to editors (who often will not realize that it is up to the investigator to choose the significance level, not an editor or a referee).

6.7 Exercises

6.1 Show that the cross-product terms in eq. (6.3) cancel out, resulting in eq.(6.4).

6.2 Calculate a one-way analysis of variance for the pheasant data using the ANOVA program in EXCEL.

6.3 Using the group variances printed out in the EXCEL output for Exer. 6.2, calculate a value for Eq. (6.5) and locate the corresponding value in the one-way ANOVA table prepared in Exer. 6.2 (the calculations can be inserted to the right of the summary of means and variances in the one-way output tables).

6.4 Copy the data from Exer. 6.2 to a new spreadsheet in the same Workbook and calculate the S. S. for the one-way ANOVA directly from the definitions given in eq. (6.4).

6.5 Repeat the calculations for Exer. 6.2 and 6.4 on the pocket gopher data given below, but first convert the counts to natural logarithms [natural log = LN(number)].

Pocket gopher counts

	Black Mesa		Grand Mesa		
	1962	1963	1964	1963	1964
1	61	237	82	167	62
2	61	271	68	183	172
3	132	253	168	297	106
4	158	157	190	188	87
5	90	244	33	238	89
6	52	180	57	285	81
7	107	269	30	124	29
8	73	138	25	209	75
9	155	159	50	248	65
10	77	237	131	204	46
11	82		108		
12	97				
13	167				

6.6 Run the two-way ANOVA without replications on the pheasant data. Notice that the total sum of squares remains the same as the one-way analysis, and the between-groups S. S. is the same as that for columns in the new analysis. However, the F-tests are now significant, and it is worth considering why this should happen (look at the mean square for error, and compare it with the within-groups value of the one-way analysis. What is your explanation?

6.7 Calculate the S. S. for Exer. 6.6 directly.

6.8 Randomized blocks The data below are from a randomized block study reported by Snedecor and Cochran (Statistical Methods, 6th Ed., p. 300). Do a two-way analysis in EXCEL and compare the M. S. for replication with that for treatments. What do the F-tests suggest?

Treatments	Replicates				
	1	2	3	4	5
CHECK	8	10	12	13	11
ARASAN	2	6	7	11	5
SPERGON	4	10	9	8	10
SEMESAN	3	5	9	10	6
FERMATE	9	7	5	5	3

6.9 Calculate the ANOVA table for Exer. 6.8 directly. It shouldn't take long, and is first-rate practice in using EXCEL.

6.10 Calculate ANOVA for the pheasant data arranged as having 2 replicates as described in the text section on 2-way ANOVA.

6.11 Copy the data from Exer. 6.10 to a new spreadsheet in EXCEL and calculate the S. S. directly from the definitions of eq. (6.8) thru eq.(6.12). This is something of a chore, with the main difficulty being in keeping things straight (use good labelling). Use two spreadsheets. You will need 2 copies of the original counts (one for computing Total S.S., the other for computing error S. S.) and 2 copies of the cell means (to calculate Treatment and Interaction S. S.). If you label things carefully and use the proper multipliers (m, k, and r) you will get the same S. S. as in Exer. 6.10. Patience is necessary, as is accuracy. If you do all of the exercises, it should help in remembering how to use ANOVA.

6.12 Snedecor and Cochran (1967:p. 347) give the following data on gains in weight (grams) of rats fed on six diets. The columns are replicates (individual rats on the same treatment). Run an ANOVA and report the results. Note that a two-way ANOVA with replications is indicated.

	High level			Low level		
	Beef	Cereal	Pork	Beef	Cereal	Pork
1	73	98	94	90	107	49
2	102	74	79	76	95	82
3	118	56	96	90	97	73
4	104	111	98	64	80	86
5	81	95	102	86	98	81
6	107	88	102	51	74	97
7	100	82	108	72	74	106
8	87	77	91	90	67	70
9	117	86	120	95	89	61
10	111	92	105	78	58	82

6.13 Use the planned comparison method on the data of Exer. 6.12 to compare the 2 levels of protein (High and Low). You may also want to look at the breakdown given by Snedecor and Cochran for this example, as it uses orthogonal comparisons to break down the treatment S.S. into 5 individual comparisons. Apply Scheffe's S-method to the data. Discuss results.

6.14 Use the planned comparison method to compare the "check" (control) mean with the other treatment means of the data in Exer. 6.8. Make the same comparison with Scheffe's method. List his criterion for $\alpha = 0.10$ as well as for $\alpha = 0.05$. Discuss results.

7.0 INTRODUCTION TO POPULATION ASSESSMENT

7.1 Viable populations

There has been much interest in recent years in maintaining viable populations of various species. Many of the efforts stem from the requirements of the Marine Mammal Protection Act of 1972 and the Endangered Species Act of 1973. Due to the complexities of natural systems, legislative mandates often cannot readily be translated into programs that can actually be implemented in a field setting. Thus an early question was "what is a viable population?" One attempt at an answer led to definition of a "Minimum Viable Population" (MVP) as the population size that had a 95% probability of surviving for 100 years. Stochastic population models were implemented on a computer and starting population size varied to find an initial population size that resulted in about 5% of the simulated population going extinct in 100 years. A major problem with such models is that the outcomes depend very much on the set of parameter estimates used in the model. If the parameters are such that the average population trend is a decrease, then higher initial population levels are required to meet the criterion than if the expected trend is upwards. Interpreting the available field data provides another pitfall -- three such stochastic models have been published that consider the Yellowstone grizzly bear population, all with shortcomings of interpretation.

When the difficulties in the Minimum Viable Population approach began to become apparent, appraisals using less specific and broader methodology were developed and described as Population Viability Analysis (PVA). Various kinds of models continued to be emphasized for the analysis of populations of endangered species (Soule 1987). These broader appraisals of Population Viability Analysis include the important feature of the possible loss of genetic variability, but there is as yet controversy about the minimum population size required to preserve sufficient genetic variability to maintain the species (Boyce 1992). It does seem to be generally accepted that an occasional exchange between isolated subpopulations is sufficient to maintain genetic diversity, so that the genetic issue may be of major importance when only one small remnant of a species exists.

A crucial uncertainty in modeling any natural population is the poorly understood role of density-dependence. Consider a population containing 50 females in which births and deaths are balanced so that the expected trend is to remain constant ($\lambda = 1.00$). If a stochastic model with no density-dependence is run 1,000 times for 100 years each run, the outcomes may be as in Fig. 7.1, in which less than 2 % of the "populations" were extirpated, thus meeting the MVP criterion of "viable". A major problem with such an approach is that constant rates cannot be expected to continue regardless of population size. If the population is stable at about 50 individuals, then it is widely accepted that increases much over this level will not occur due to "density-dependence", which seems largely to be expressed in large mammals as resulting in lower first year survival when an asymptotic population level is approached. Conversely, if the population falls much below the asymptotic value, then early survival increases so that it tends to return to the asymptotic level. Unfortunately, the model (generalized logistic given in Chapter 13) that seems to best fit data on larger animals does not account for the occasional "overshoot",

when a population exceeds the “carrying capacity” level, so there isn’t any really satisfactory model for populations above carrying capacity. Consequently, an effort to introduce density-dependence in the model used for Fig. 7.1 has to have an arbitrary component for those cases where the population overshoots.

The model used here uses data for an elk population, as presented in Chapter 11. Data for rapid growth of that population come from work by Houston (1982). More recent data on a Yellowstone elk population (Garrott et al., 2003), gives adult female survival rates of about 90%, in contrast to rates of about 99% from the work by Houston (1982). In order to achieve the “stationary” population, first year survival was reduced until the estimated value of λ was unity, with adult female survival at 0.90 (only females were considered in the simulations). Random draws for individuals were used to generate the data of Fig. 7.1. Density-dependence was then introduced through the generalized logistic model of Chapter 13, which seems to work reasonably well for populations below carrying-capacity. When an “overshoot” occurred, rates for the stationary population of Fig. 7.1 were introduced, giving Fig. 7.2 for the outcomes of 1,000 runs. This arbitrary adjustment for “overshoots” means that the right side of Fig. 7.2 is of somewhat dubious utility (but less than 8% of 1,000 simulations were “overshoots”). The important feature here is that density-dependence resulted in no populations going extinct, with a lower limit around 20.

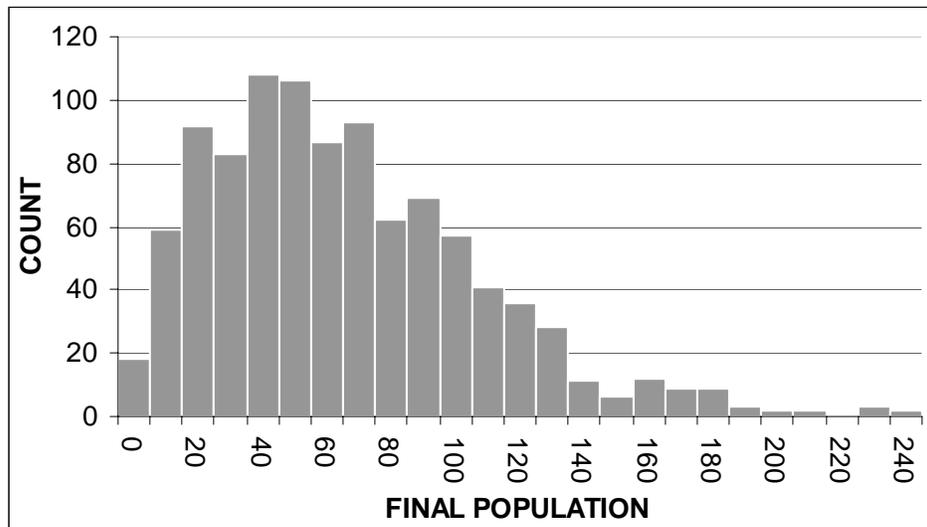


Fig. 7.1. Outcomes of 1,000 simulation runs of a population starting with 50 females with parameters selected to give $\lambda = 1.00$, and no density dependence.

Until much more is known about the mechanisms controlling trends in small populations, the most prudent approach to maintaining viable populations is to concentrate on "population analysis", that is, to determine survival and reproductive rates in an effort to determine whether the population can be expected to increase or decrease in the immediate future. If a decrease seems likely, then management actions need to be directed towards changing the rate most likely to be responsible. First-year or adult survival appear to be responsible in the available examples. The same general principle appears to apply in those cases where the goals are to control an over-numerous population or to secure a maximum sustainable yield from an exploited population.

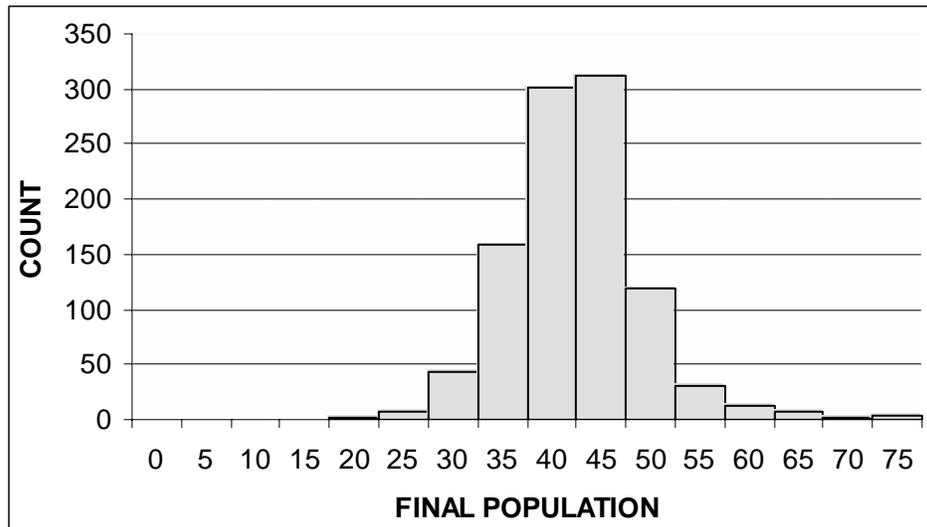


Fig. 7.2. Outcomes of 1,000 simulation runs of 100 years for an initial population of 50 females with a density-dependence function acting on first-year survival and "carrying-capacity" (K) set at 50 females.

Many studies may require 10 years or more to obtain enough data to determine the key issues, and to begin to develop effective approaches to solutions. Often the initial assessments of the perceived problem turn out to be in error, and it may take a long time to correct these initial perceptions in the face of public pressure to "do something". While the "environmental movement" has created a climate in which actions to maintain endangered or threatened species have become possible, the many private organizations dedicated to fostering such actions may hamper progress in particular cases. This usually results from the need to demonstrate their active participation in order to maintain a flow of funds from the public. Litigation engenders publicity and thus funding, but may also seriously limit the ability of a responsible agency to obtain required data.

7.2 Methods for population analysis

Administrators and the general public always ask "How many are there?" when faced with some issue concerning a population. This seems to be a wholly reasonable question, and is one that often has to be answered in one way or another. If sizable removals are made for exploitation or control purposes, an estimate of absolute numbers may be essential, in order to assess the likely impact of the removals. Endangered species are often present in low densities, and may thus be very difficult to census. Consequently, an investigator may expend much of the available resources in an attempt to obtain a population estimate, only to discover that the precision of the estimate is not adequate for the major goal of any study of an endangered species. This almost always has to be one of determining trend of the population. If, as usually seems to be the case, a very large effort has to be expended to get a census estimate of relatively low precision, then repeating the census in another year very likely will not supply a useful measure of trend.

There are basically two alternatives. One is to opt for some measure of trend based on an index of abundance that may be much less expensive to obtain than an estimate of absolute abundance. The other is to obtain reproductive and survival data on which to base an estimate of likely trend. This has a distinct advantage for studies of endangered species, inasmuch as such studies usually also need to try to determine why the species is "in trouble" and what might be done to insure a positive rate of increase. If one relies only on trend data, it likely will take a sizable number of years to establish a trend, and the mere observation of trend will not provide any information on reasons for the change. It is true that the basic cause for a declining trend may be obvious, i.e., loss of suitable habitat. Nonetheless, it may be essential to know how this cause affects rate of change in the population, and this requires knowledge of reproductive and survival rates. If positive steps to reverse a decline can be taken, then the impact of such steps will most likely first be evident in reproductive and survival rates.

Virtually all field data on large populations is inadequate in scope for "textbook" analyses of present status and likely future trends of the population. Each data set has unique features, often in consequence of the unique nature of the particular species, but also because of the difficulties and costs of data collection. A variety of approaches is thus required, including various kinds of approximations and indirect methods of estimating essential parameters of the population.

Most field studies of large populations seek to predict the future trend of a population by assessing data collected over time. In some instances either research goals or a legislative mandate (see, for example, Eberhardt 1977a) may direct attention mainly to the past, often with a goal of evaluating present status of the population or predicting the likely impact of some alteration of the landscape or its uses. A universal need in such studies is to evaluate likely accuracy and precision of the outcome.

Because the methods used in practice all require assumptions that are very difficult to support under field conditions, the only satisfactory demonstration of accuracy may be independent estimates of the same quantity, usually rate of change of the population. This can be achieved by comparing estimates based on trend data (e.g., a log-linear regression of a population index) with those derived from reproductive and survival data. Most studies fall short of this goal through lack of estimates of some essential rate. Another need is thus to indicate something of the minimal requirements for assessing accuracy of a population study.

Uncertainty about fulfillment of assumptions has similar effects on precision for any given kind of estimate (e.g., of population size). Since an overall assessment will require a number of individual estimates, appraising precision of the end-product (e.g., a rate of increase) is also very difficult, and rarely attempted. The problem needs to be faced, however, because improving population analyses will likely depend on combining inferred parameter estimates (e.g., those derived from, say, age and sex ratio data) with direct measurements (obtained, for example, through radiotelemetry). Combining two such sources will usually require weighting by measures of precision. The non-

parametric methods, such as bootstrapping and jackknifing, provide the necessary flexibility.

The basic ingredients for an assessment are measures of population size and estimates of reproductive and survival rates. These may be obtained in a variety of ways, and can be used in an analysis based on either a direct projection of population size or by estimating a rate of change from survival and reproductive rates alone. The major difference between the two is that projections require estimates of population size. We are thus concerned here with three sources of data and two methods of utilizing that data (Fig.7.3). A brief listing of the individual sources and methods given in Fig. 7.3 follows.

Population estimation can be approached by making direct estimates of actual abundance or through indices of abundance. There is a very large literature on methods for estimating animal abundance (Seber's 1982 book is still the major reference) and much theoretical and practical work continues to be published on these methods.

Survival estimates are usually obtained either indirectly through analysis of age structure data or directly through evaluation of data from tags and marks. Use of age structure data alone requires the very restrictive assumptions of constant population size and constant recruitment to the population in the years in which the age structure was developed, along with constant survival. If tags are used, then the assumption of constant recruitment is not needed in estimating survival from tag recoveries. A major advance has been use of radiotransmitters as tags, reducing uncertainties and variability associated with tag recoveries.

Reproductive rate data are usually more readily obtained than information on population size and survival. Often sizable numbers of individuals can be examined for pregnancy and age-specific rates derived directly. However, some species (e.g., bears and whales) do not reproduce annually, so that it may be necessary to use a composite measure of reproductive rate, based on sex ratios at birth, litter sizes, breeding intervals, and so on.

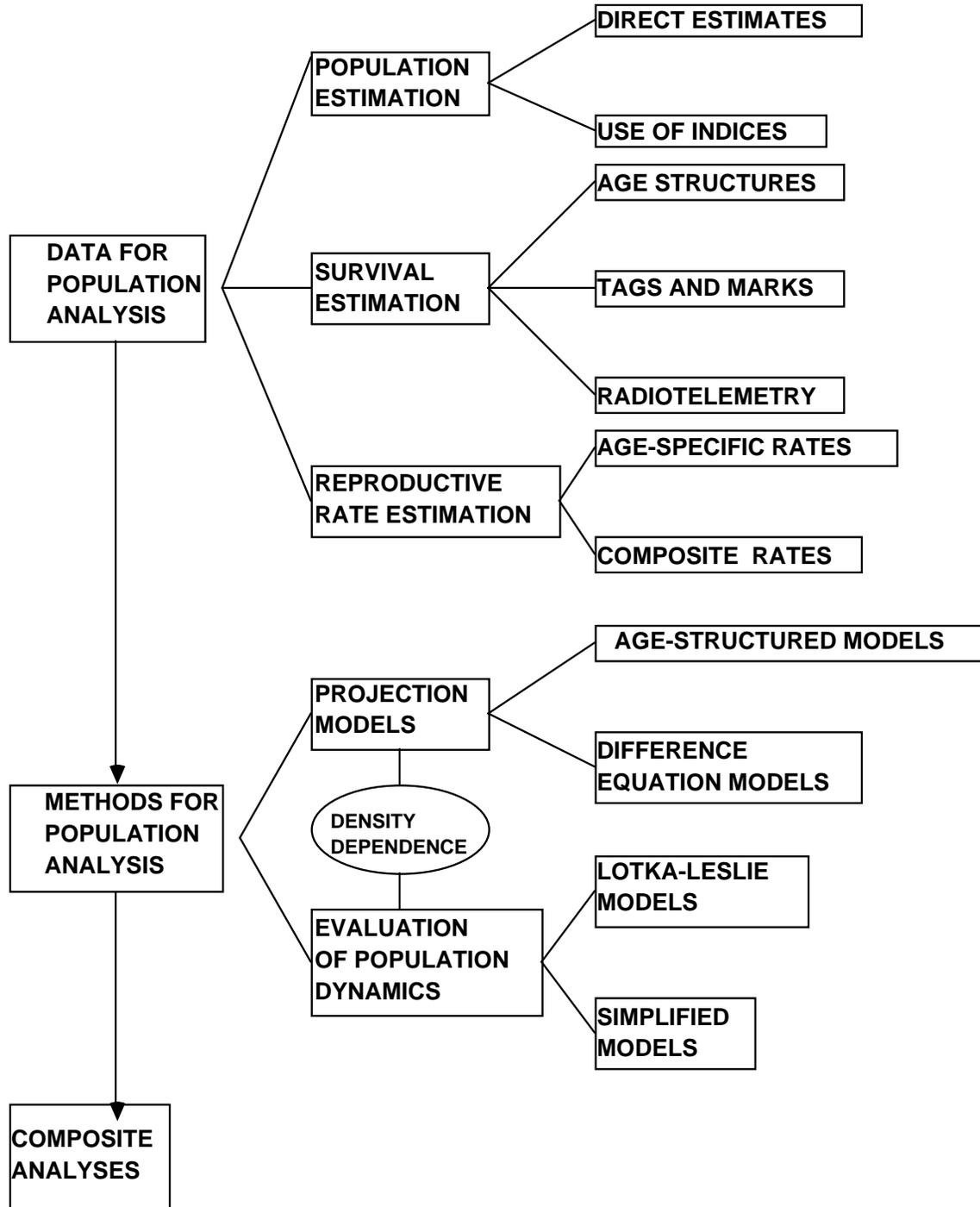


Figure 7.3 An outline of techniques useful in analysis of large vertebrate populations.

Methods for population assessment using the data outlined above can be considered in two classes. *Projection models* require an estimate of population size and apply birth and death rates to project the population from one year to the next. The simplest such models that incorporate provisions for removal of known numbers of individuals are:

$$N_{t+1} = R(N_t - K_t) \quad (7.1)$$

$$N_{t+1} = RN_t - K_t \quad (7.2)$$

where K_t denotes removals at time t and R represents a rate of increase. The two equations differ in terms of whether the removal takes place just before or just after reproduction occurs. When age-specific birth and death rates are known, the equations may be written as *matrix models*, with $R = \mathbf{M}$, the Leslie matrix. When age-specific rates are not available, the equations may be expressed as simple *difference equation models*, with $R = \lambda$, a constant annual rate of increase.

One of the major difficulties in actual applications of projection models is the present uncertain state of knowledge about population regulation. If the prospect of *density dependence* is to be incorporated in the model, then some sort of functional form has to be assumed, operating directly on R for difference equation models and one or more of the elements of the Leslie matrix for age-structured data.

In the absence of population estimates, a direct *evaluation of population dynamics* may be conducted on the basis of age-specific reproductive and survival rates. The classical approach is that of A. J. Lotka, who first demonstrated that constant age-specific birth and death rates result ultimately in both exponential growth (or decline) and a constant relative age structure (the stable age distribution). Lotka's basic equation is equivalent to the characteristic equation of the Leslie matrix, so the models are here described as *Lotka-Leslie models*. In practical applications, there usually is not enough data to use more than a few reproductive and survival rates, so Figure 7.3 includes a class of simplified models based on these rates.

7.3 Population estimation

Although there is a large array of methods for estimating population densities, relatively few methods are actually used on large vertebrates. The principal techniques used in the field are transect methods and those based on harvest data (largely "catch-effort" methods). For the most part, the use of a limited set of methods is a consequence of the very large areas involved and the high costs of marking animals on such areas. The catch-effort methods have mainly been applied to marine mammals, usually give highly variable results, and suffer from several other difficulties (cf. Eberhardt, Chapman and Gilbert 1979). Transect methods are largely used from aircraft or ships and thus are frequently subject to uncertainty as to whether the assumptions required for density estimation are met (Eberhardt, Chapman and Gilbert 1979; Burnham, Anderson, and Laake 1980). As noted above, the major reference for estimating both population size and survival rates is the book by Seber (1982). A monograph by Pollock et al. (1990)

provides some further discussion of underlying models and access to a computer program suitable for estimating parameters of these models. Program MARK maintained at Colorado State University may be one of the most up-to-date of the many programs now available.

In many applications, it is likely that the data on population abundance have to be considered as an index, rather than as an estimate of absolute numbers. There are two difficulties. One is just that very little quantitative work has been done on indices. A more serious problem is that many field applications concern populations subject to harvesting or other removals. Since removals are expressed in absolute terms, it is necessary to also express population size in absolute terms, or to make the very restrictive assumption that removals are directly proportional to abundance. Because the projection models require estimates of absolute abundance, indices may be most useful to check calculation of a rate of change from the Lotka-Leslie models in those cases where the population is not subject to substantial removals.

Optimal use of index data for population analysis requires some sort of calibration to convert the index to an estimate of absolute abundance. Often a method that provides direct estimates is available, but is too expensive or time-consuming for application over the entire area of concern. The appropriate means for calibration may then be the statistical technique known as double sampling, in which small samples obtained through an expensive but accurate technique are used to make ratio or regression corrections to large samples obtained by a relatively inexpensive technique (the index). Unfortunately, the available statistical basis for the method depends on approximations, so that the usual recommendations for sample sizes (Cochran 1977) are larger than can be managed for many animal population studies. Some simulations indicate that smaller samples can be used, and various devices may be used to try to reduce the effort required for calibration (Eberhardt and Simmons 1987).

The principal alternative to calibrating an index may be large scale marking. For the most part, costs are too high to make capture-recapture uses of tagging and marking feasible for large populations inhabiting sizable areas, so that the technique has mainly been used to estimate survival. The high, stable survival rates necessarily exhibited by at least the adult female age classes of many species of large vertebrates suggest the possibility of periodic population estimates based on tagging or marking over a number of years, so that enough marks are built up in the population to yield reasonably precise estimates in a recapture series. DeMaster et al. (1980) reported one such attempt for polar bears, but assumed a constant, known survival rate for estimation purposes. Unfortunately, satisfactory direct estimates of survivorship are very difficult to obtain for this species (and most others). Consequently, it seems likely that progress along these lines will call for the imposition of appropriate restrictions on the Jolly-Seber method (see, for example, Brownie, Hines, and Nichols 1986).

7.4 Survival estimation

Survival estimates are likely to be the most important ingredient in population analysis, yet are often the least satisfactory estimate obtained in actual practice. Their

importance is due to the magnitude of the effect of a small difference in adult female survivorship on rate of change of the population (cf. Eberhardt and Siniff 1977). A few percentage points one way or another spell the difference between a comfortably increasing population and one threatened with extirpation. Subadult survival rates have markedly less effect on rate of change, and may thus be the major factor in population regulation, while adult female survivorship quite possibly provides the last mechanism to come into play in a sequence of events under natural conditions (Eberhardt 1977b).

Poor quality or outright absence of survival estimates in many studies no doubt reflects the high cost of obtaining useful estimates. The least expensive route is via age structure data. If survivorship in adult age classes is constant from year to year, and the same number of individuals are recruited to adult status each year, then the ratio of numbers in successive age classes will ultimately reflect the common survival rate. Averaging over a number of age classes in a sample is necessary to reduce variability in the estimate, and Chapman and Robson (1960) described an efficient estimator for that purpose. Other methods can be very much less efficient (see Eberhardt 1972 for examples).

Unfortunately, the assumptions required for use of survival estimates from age data are so restrictive that they are unlikely to be met in practice, and should be tested whenever the method is used. A null hypothesis of constant annual survival rates plus constant recruitment implies that age structures in successive years should be homogeneous, so that chi-square might be used to test the hypothesis and thus the assumptions. However, such homogeneity is not a sufficient condition to justify estimating survival from age structure data since homogeneity of successive age structures is implied by the stable age distribution (Keyfitz 1968). Such a distribution results from populations changing at a constant rate. Consequently, adequate justification for using survival estimates from age structure data also requires a demonstration that population size has remained constant.

In many instances, a population may tend to increase (or decrease) at a relatively constant rate (examples are given by Eberhardt 1987). The Chapman-Robson method may then still be used, but now under the assumption of constant survival and of recruitments changing at a constant rate. The parameter estimated then becomes s/λ , where λ denotes the "finite population multiplier", i.e., $\lambda = e^r$ or $\lambda = 1+r$ depending on whether continuous or discrete rates of population change are assumed.

The alternative to using age structure data is to mark or tag individuals, and estimate survival rates on the basis of rates of recovery of these individuals in subsequent years. This disposes of the assumption of constant recruitment, by virtue of the fact that the number of marked individuals is now known. The Chapman-Robson estimation procedure remains appropriate with recoveries from marking and tagging (Paulik 1962). However, if marked animals are introduced over a series of years, the way is then opened for a wide range of estimation procedures along with some additional tests of assumptions. An extensive set of models and estimation procedures was developed by Brownie et al. (1978), mainly with reference to applications to bird-banding. A general method of analysis based on numerical solutions of maximum-likelihood estimators was

proposed by White (1983), and illustrated on a set of large vertebrate data. The book of Seber (1982) contains the most extensive coverage of methods, but should be supplemented with the recent monograph by LeBreton et al. (1992) which covers recent developments and lists the many computer programs now available for processing data.

One of the more troublesome aspects of use of marking and tagging for estimating survival is that only a relatively small fraction of marked individuals are ever recovered. This opens the door for a variety of biases (cf. Pollack and Raveling 1982) and results in substantial variability in the estimates. A logical approach to the difficulty is then to use telemetry, so that status of individuals is largely known throughout life of the transmitters. Relatively little quantitative work has yet been done on survival rates estimated from telemetry data. A useful reference dealing with circumstances in which survival of an individual is checked at irregular intervals (Bart and Robson 1982) gives results that should be useful in telemetry studies, while Heisey and Fuller (1985) specifically discussed estimation of survival rates by telemetry, as did Pollock et al. (1990).

A problem in using these methods is that they are based on the assumption that survivorship is constant from day to day and among individuals. The results may thus be suitable for relatively short time intervals, but cannot safely be extended to longer periods, such as a year, due to the prospect that daily rates cannot be assumed constant throughout long periods. An alternative is to do the analyses in terms of years, not days. This will, in most cases, be necessary for long-lived animals in any case, due to infrequency of mortalities. Another issue in need of attention is that of censoring, i.e., individuals may be lost to observation when the radios cease to function or by emigration from the study area. Also, in long-lived species, many of the tagged individuals will remain alive at the end of the study period.

7.5 Reproductive rate estimation

For those species that reproduce annually, there usually is little serious difficulty in obtaining estimates of reproductive rates. The major sources are observations of pregnancy rate and tallies of young per adult female. Normally only pregnancy rates can be determined on an age-specific basis. In any case, survival rates for the youngest age class used need to be defined in terms of the reproductive rate used, e.g., if pregnancy rates are used then early survival includes mortalities in late-term pregnancy.

Species that do not reproduce annually may pose some special problems in estimating reproductive rates, particularly when direct observation is not feasible. It may then be necessary to use a composite rate, based on different sources of data. Litter size may be directly observed, while sex ratios of young may be available only from those individuals that can be caught and handled (often sex ratio is simply assumed equal at birth). The composite rate might then be computed as:

$$m = \frac{(\text{sex ratio})(\text{mean litter size})}{\text{breeding interval}}$$

Direct measurements of breeding interval are likely to require marking or tagging, and may best be done with radiotelemetry to reduce the prospect of missing a reproductive event. Knight and Eberhardt (1985) discuss problems in estimating this composite rate for grizzly bears.

In some situations it may be feasible to substitute proportion giving birth rather than breeding interval since, in an equilibrium situation, the proportion giving birth will approximately equal the reciprocal of breeding interval. When single births are the rule, if sex ratios are assumed equal, then m simply equals proportion giving birth. It is, of course, necessary to assume that only fully mature females are involved in the calculations. In reality, individuals usually do not reach full reproductive capability in a single year, so that some corrections for reproductive rates of younger animals may be needed.

7.6 Projection models

The major uncertainty in using projection models is whether or not some functional representation of density dependence needs to be incorporated in a given example. Although most ecologists largely accept density dependence as reality, there simply is not enough information available to specify how it may apply in particular circumstances. If regulation is to be considered in a projection, then a specific model has to be used. Two classes of models have been used in actual applications concerning vertebrates. The model used for large vertebrates has been described as a "generalized logistic":

$$r_t = r [1 - (N_t/K)^z] \quad (7.3)$$

where the annual rate of increase (r_t) is reduced from a maximal rate (r) as the population (N_t) increases towards an asymptotic value (K). The rate (z) controlling approach to the asymptote determines the inflection point:

$$p = (1 + z)^{-1/z} \quad (7.4)$$

When $z = 1$, this model reduces to the discrete form of the ordinary logistic growth model, hence the term "generalized logistic". The main practical application thus far has been in calculations performed for the International Whaling Commission.

The second class of models used in practice includes the stock-recruitment models of fisheries research and management. The two main forms are those of Beverton and Holt and of Ricker (Ricker (1975) gave extensive details). The Beverton and Holt curve is essentially based on the ordinary logistic growth curve. As a growth model, Ricker's model can be shown to be essentially equivalent to the Beverton and Holt curve when r is in the range typical of large vertebrates (Eberhardt 1977c). It is only when annual rates of increase approach those reported for some species of fish that the two curves need to be distinguished.

Probably the only real assurance as to whether a functional model of density dependence needs to be incorporated in a given actual projection comes from experience. There is a slowly increasing body of evidence that the inflection point for large mammals is usually well above the 50 percent level of the ordinary logistic curve (Fowler 1981). If this is generally true, then it may not be necessary to utilize a density dependence function in a projection model unless the population approaches relatively high levels. At present, past experience with the particular population likely has to be the principal source of guidance on this score. Eberhardt (1987) described growth patterns for 16 populations with a simple exponential curve, but also found it necessary to use a density dependent function at higher levels for several of these populations.

In a number of cases, projection-type models have been used for "back-calculations". Most examples result from the need to compare present population levels with likely pre-exploitation abundance (see, for example, Breiwick, Eberhardt and Braham 1984). It will then be essential to include density dependence in the model, and the inflection point chosen will have an appreciable impact on outcome of the back-calculation. Most such back-calculations have been made to comply with regulations that mandate a given population should not be reduced below its "maximum net productivity" or maximum sustainable yield level (cf. Eberhardt 1977a). However, choosing different inflection points may influence the outcome substantially.

When it comes to actually using a projection model with data, very little work has been done with fitting matrix models to large vertebrate populations. This is largely a consequence of the fact that the necessary data are usually not available. About all that can presently be done is to attempt to estimate the essential parameters for the Leslie matrix at some point in time. If population estimates for a series of years are available, it may then be feasible to attempt to see how well matrix calculations "track" the observed data. If age structure data are also available for a number of years, it should then be possible to attempt to estimate some of the parameters in the matrix by iterative fitting procedures using a chi-square criterion.

In the much simpler case where the projection model is a difference equation, it may then be feasible to estimate parameters by direct fitting procedures (some examples were given by Eberhardt 1987). If it is necessary to use equation (7.3), or some other functional representation of density dependence, then the fitting will become much more complicated.

7.7 Some historical features of population analysis

The word "population" derives from the Latin *Populus*, meaning people, while "demography" stems from the Greek *Demos*, also meaning people. These roots indicate clearly the origins of the terminology and methodology now applied to aggregations of all kinds. It is now common practice to use "population" to mean any well-defined collection of objects, both animate and inanimate. There are, however, those who interpret "demography" as meaning only the study of human populations. Because the bulk of the techniques in use in ecology stem directly from early work on human

populations, it seems pointless to use a separate term for animal and plant populations. It is obviously bad grammar to use "animal demography".

An excellent review by Cole (1954), under the title "Sketches of general and comparative demography", provides many interesting historical details, and is the basis for much of the present section. Students should also refer to the book by Allee et al. (1949) for additional perspective on origins of ecological population studies. Although over 50 years have elapsed since this book was published, it is still one of the better references, due to its detailed and thorough coverage of many features of animal ecology.

Population enumeration was no doubt practiced well before the census was developed by the Romans, and an elaborate system of population registration was in use in China before Marco Polo's time. Surprisingly enough, not until the 17th century did the "modern" nations begin complete population enumerations. Cole (1954) suggested that "Plato, and probably Solon before him, had a definite concept of an optimum population size and an understanding of factors regulating population size."

An Italian, Botero, in 1588 clearly recognized the limitations placed on population growth by environmental resources. He also preceded Malthus by some two centuries in formulating the concept of potential geometric growth of populations. Skellam (1955) noted that Linnaeus (in 1740) described potential population increase in plants by a geometric growth scheme, thus also preceding the famous 1798 essay by Malthus.

The Romans sold annuities at rates that changed with advancing age, but it was not until 1662 that the basis for an effective life table was developed by John Graunt, and then refined into something approaching modern versions by Huygens in 1669. In the 1750's Buffon enunciated a clear qualitative statement of the principles of the "balance of nature", while a century later Darwin and Wallace produced the ultimate key to evolutionary understanding. In the same period, the forerunners of modern mathematical and statistical development were at work. These included Quetelet, Gompertz, and Verhulst (in the 1820's and 1830's) followed by Galton and the rapid development of biometric methods culminating in the work of Karl Pearson and his associates in the early 20th century. Modern mathematical approaches to demography were pioneered in the early 20th century by Pearl, Lotka, and Volterra.

7.8 A classification of methods for estimating abundance

The classification used here rests on a basic dichotomy between situations where plants or animals can be readily and directly counted, and those where this is not feasible. The direct sampling methods can be further subdivided in terms of the sampling units employed in the field. Indirect sampling methods can be conveniently subdivided in terms of whether or not an individual animal is likely to be observed on more than one occasion. Single observations necessarily result when the animal is killed (e.g., catch-effort methods) and are generally expected for most of the index methods. Repeated observations on individuals are necessary for the capture-recapture methods.

The classification is given below in segments, with each unit followed by a brief discussion. References to Chapters or Sections are included, along with a few special literature citations. Additional references are included in the appropriate chapters. The classification is adapted from one given by Eberhardt (1978a). The first section deals with direct counts of individuals.

 Classification Applications References

I. DIRECT SAMPLING	Used where direct counts are feasible	Chapter 4
A. Area counts		
1. Discrete sampling units	Parasites on or in hosts, colony counts, artificial substrates, sampling catches of individual vessels, time-area counts	
a. Counts of all individuals or absence		
2. Quadrats		
a. Counts of all individuals	Counts of plants, deer drive counts, corers.	
b. Tally of presence or absence	Used in attempts to reduce sampling effort	
c. Proportion of plot occupied	Used for plants when individuals difficult to distinguish.	
7. Strip transects		
a. Counts of all individuals	Counts of plants, inanimate objects, sessile animals.	Section 5.10
b. Partial counts		
i. Visibility decreasing with distance	Animals that do not flush or are sessile, plants, inanimate objects.	
ii. Intermittently visible	Marine mammals	
B. Counts at fixed points	Counts at dams or weirs, or vantage points along streams or coastlines (usually	

migratory species).

C. Line methods

- | | | |
|--------------------------|--|-------------|
| 1. Line intercept | Plant canopies and other sizable objects | Section 5.3 |
| 2. Line transect | | |
| a. Animal flushes | Animal censuses | Section 5.5 |
| b. Searching by observer | Inanimate objects, animals that do not flush | Section 5.6 |

D. Point methods

- | | | |
|--------------------------------------|--|--------------------|
| 1. Point frames | Plants | Greig-Smith (1964) |
| 2. Distance methods | | |
| a. Radial | Plants, sessile animals | |
| b. Linear ("variable-area" plot) | An alternative arrangement more suitable for use in the field. | Section 5.4 |
| 7. Bitterlich method ("angle-count") | Used to estimate basal area in forestry | Section 5.4 |

An important distinction in "area counts" is whether or not natural groupings can be accurately distinguished. Examples of such natural sampling units include individual plants, beaver colonies, fishing vessels, and the like. When natural sampling units are not available, or cannot be precisely delineated, some sort of artificial sampling unit has to be defined. The commonest example is a sample plot or quadrat. When natural sampling units are available, two factors must be considered. One is whether or not a random sample of such units can be obtained, and the other is whether all of the items of interest on selected units can readily be enumerated. If random sampling of units is feasible, and complete counts of individual items on the units are readily obtained, then standard sampling methods are appropriate. When there are difficulties with either factor, then other techniques need to be used. These can be quite complex, as will be evident from looking through any of the sample survey texts (e.g., Cochran 1977). However, a simple approach that works well in many practical situations is to use a plot sample to estimate the number of units, and then to tally the items of interest on all or a subsample of the natural units falling in the plots (cluster sampling).

In the case of either discrete sampling units or quadrats, various efforts have been made to reduce the labor involved in tallying individual plants and animals by only recording whether there are any individuals present or not ("presence and absence" data). Sometimes this information is all that is wanted, e.g., in determining the proportion of plants or animals infested by parasites. However, when this approach is used as a shortcut for estimating a total count, it usually fails. This is because the usual underlying assumption is that the individual items (e.g., parasites) are randomly distributed to sampling units (hosts). If this is the case, then the binomial distribution (or Poisson

approximation) holds, and the proportion of "zero counts" can be used to estimate a total number or density. Almost invariably some kind of clumping or "contagion" (i.e., non-randomness) holds, and the method does not work. It may be possible to assume some kind of non-random distribution (e.g., the negative binomial) and proceed to make an estimate from the frequency of zeros postulated by that distribution. Unfortunately, this involves knowing or estimating one or more additional parameters for the assumed distribution, so that it is almost always better to resort to stratified sampling or ratio estimation.

Some plants do not have readily distinguishable individuals, so that counts are very difficult. In this situation, and when it may not be desirable to attempt to tally all of the individuals present, one may simply resort to measuring or estimating the proportion of the plot covered by vegetation. An alternative is to measure the biomass present, either by clipping and weighing material on the entire plot, or by subsampling.

Strip transects are essentially long, narrow plots, and thus can be treated by the methods already discussed. They are, however, discussed separately because of the close connections with other transect methods, and in consequence of some special problems. One is that of objects present on the transect strip, but not observed, and another is that a set of sample transects may have quite different lengths, requiring some provisions for adjustments.

Counts at fixed points have mainly been used for migratory species, but not much statistical analysis has been done on the resulting data, or in designing appropriate sampling schemes. Stratification may be the best approach, with strata being times of day, season, etc.

The "line" methods have two major categories. One depends on the interception of some sizable object by a line laid out by the observer. Data may be collected to either determine the proportion of the total area covered by the objects, or the number (density) of objects on the study area. Two different measurements are taken (length of interception, and width of object), and these permit unbiased estimation. The line transect methods depend on measurements of distances from a transect line traversed by the observer to objects of interest. When the "objects" are animals that are observed because they are startled (by the observer's approach; they "flush"), one kind of theory seems appropriate. If, on the other hand, detection depends on the observer's locating the object, a different theoretical approach may be preferred.

The "point" methods depend on measurements taken at sample points. In one version, the points are projected onto vegetation (e.g. by a set of long thin metal pins, a "point frame") to determine proportions of vegetative cover provided by various species. Several problems exist, including layering of vegetation and inefficiency in sampling, and the method is not very widely used.

The distance methods have several variants, but the best-known and most useful depend on the distance between a randomly selected sampling point and the n^{th} nearest object of interest (usually a plant). Although there has been much interest, and variety of

theoretical developments, the current picture seems to be one of use for studying pattern, and not for estimating density. The basic method depends on searching outward in a spiral from the sampling point until the n^{th} object is located. This can be rather difficult in the field, so an alternative method is worth considering. This is to use an open-ended plot that is extended until the n^{th} object is located. The underlying theory approximately that of the conventional plot approach if n is large enough.

The Bitterlich method is mainly used by foresters, and has an interesting connection to line-intercepts. The basic method depends on whether or not an "angle-gauge" appears to be narrower than the apparent width of a tree-trunk. The remaining methods are based on indirect tallies of various kinds, conveniently split into two sections based on whether or not individuals need to be observed on more than one occasion. The first segment deals with methods that require only a single observation of individuals.

 Classification Applications References

II INDIRECT SAMPLING Used where direct visual count not feasible

A. Single observation
 of individuals

1. Catch-effort methods	Mainly used with harvest data;, also electro-shocking	Ricker(1975)
a. Closed population	Population assumed unaffected by mortality, recruitment, emigration or immigration during sampling.	
i. Variable effort		
(1) Leslie method	Regression of catch per unit effort on cumulative catch.	Seber (1982:297)
(2) Ricker method	Regression of log c.p.u.e. on cumulative effort, catchability coefficient large.	Seber (1982:302)
(3) DeLury method	Regression of log c.p.u.e. on cumulative catch, catchability coefficient small.	Seber (1982:303)
ii. Constant effort	Same effort applied in each sampling	Zippin (1956,1958)
b. Open population	Situations where mortality, recruitment, emigration, immigration are likely to be significant.	

2. Indices

a. Visual	Roadside counts, aerial counts, roadkills, transects (unadjusted), census with dogs	Chapter 9 (Overton and Davis 1969:427)
-----------	---	---

b. Capture or harvest	Bag and creel census, drift-samplers nets, trawls, set-lines, traps plankton-pumping, grabs, dredges, electro-fishing, poison.	Ricker (1975)
c. Signs	Fecal counts, dens, mounds and nests, tracks, beds, roosts, scent posts, muskrat houses, beaver dams and lodges, amount of food consumed.	Chapter 9
d. Auditory		
i. Active	Echo-ranging (fish, aquatic invertebrates).	
ii. Passive	Tallies of calls or other sounds.	

The catch-effort methods almost universally depend on data obtained by exploitation of a population. It is thus not surprising that they were largely developed in fisheries management, and continue to be mainly applied in that field. An important part of the development of fisheries usage is due to W. E. Ricker, whose 1975 book should be consulted for more details. The methods largely utilize simple regressions of catch per unit effort (cpue) on either cumulative catch or cumulative effort. In exploitation, effort normally varies from day to day (or week to week, depending on how records are kept), so the variable effort models are of major importance. Seber (1982) used different names for three main equations, but many authors lump the methods as "Leslie-DeLury models". One situation where effort usually remains constant from day to day is that of "removal trapping".

The catch-effort models are very simple and easy to use when the population can be assumed to be "closed", i.e., affected only by the harvests. When the duration of the period of exploitation is not short, it becomes necessary to deal with losses from other causes (e.g., "natural mortality"), and, in some cases, with additions to the population ("recruitment"). In these situations, the models may become rather complex, and the results may be quite unsatisfactory unless additional information about the population is available. One way to provide such auxiliary information is to introduce tagged animals into the population at various times. Readers faced with this situation are advised to consult Ricker's (1975) book and more recent texts on fisheries management. Surprisingly little use has been made of catch-effort methods on game harvest data, which is unfortunate, considering that such data are quite widely available.

A major source of information about relative levels of populations has received very little quantitative and statistical treatment. This is data that can be expressed as an "index of abundance", or measure of relative abundance. The classification given above amounts to a convenient way to categorize the data by the means by which an animal's presence is observed. The actual analysis of such data depends on some sort of model, on the sampling method used, and on any auxiliary information that may be available.

Some indices can quite readily be converted to direct estimates of population density, given the appropriate conversion data are available. One example is the deer pellet-group count. Other indices may be expressed in terms of population density but are known to be biased. Conceivably these methods might be treated separately as direct estimates of abundance, but they seem most readily dealt with as indices. The best known census methods depend on repeated observations of individuals (the capture-recapture methods) and are usually divided into the simpler applications requiring the assumption of a closed population, and the more complex situations where a population is "open" to losses and gains. We first consider methods appropriate for closed populations.

Classification	Applications	References

B. Repeated observations of individuals		
1. Capture-recapture	Mobile and secretive animals	Chapter 8
a. Closed population		
i. Petersen method (Lincoln Index)	Usually only 2 sampling periods. May also be applied with stratification.	Section 8.2
(1) Basic method		
(2) Sampling with replacement	Second sample by visual observation.	
(3) Sequence of removals	Second sample taken as sequence of observations (e.g., tag recovery in commercial fishery).	Seber (1982:125)
(4) Subsampling	Second sample observed on random sample of subareas.	Seber (1982:111)
(5) Inverse sampling	Second sample size (tagged or untagged) fixed in advance.	
ii. Schnabel method	More than 2 sampling periods; tagging continues throughout sampling.	Section 8.3
(1) Basic method		
(2) Mean Petersen method	Petersen estimates from successive pairs of samples averaged; may reduce effects of departures from assumption of closed population.	
(3) Inverse and sequential methods	Fixing number of tagged or untagged to capture in advance.	
(4) Corrections for catchability	Attempts to correct for violation of assumption of equal probabilities of capture.	
(a) Frequency of capture	Number not captured estimated by fitting frequency distribution to	

methods	capture data.	
(b) Marten's model	Assumes catchability changes at constant rate.	Seber (1982:150)
(c) Tanaka's model	Regression of log(prop. marked) on log(cumulative marked)	Seber (1982:145)
(5) Multi-sample single-recapture	Each tag release followed by permanent removals (as in commercial fishery).	Seber 1982:193

As the name implies, the capture-recapture methods depend on catching an animal, marking and releasing it, and then again capturing it at a later time. When the population under study can be assumed "closed", that is, not to gain or lose members during the study period, then rather simple estimation methods can be used. Also, the marking method does not have to distinguish between individuals. It only needs to indicate that a given animal has previously been captured. The simplest approach (Petersen method) requires only an initial marking, followed by one recapture period.

In the basic method, it is assumed that a sample of the population is somehow marked, and that a random sample is later captured and examined for marks. This provides an estimate of the proportion marked in the total population, so that it is a simple matter to calculate an estimate of the total population size. It is usually assumed that the members of the second sample are all caught at nearly the same time. However, this isn't necessarily the case. It may be possible just to observe the animals on a number of occasions and to record the fractions marked. This amounts essentially to sampling with replacement, and leads to a somewhat different model. In some situations, the second sample may not be obtained from one recapture operation, but may come from a sequence of removals from the population. Data of this kind may be treated by a regression model.

When large areas are under study, it may be necessary to use subsampling, leading to various complications, for which not enough experience is yet available for definite recommendations on procedures. A final variant on the basic Petersen model depends on fixing, in advance, the number of marked or unmarked animals to be captured on the second occasion. Since this is rather difficult to do in practice, the approach is mainly of theoretical interest.

The Schnabel method was developed to deal with situations where animals are captured and released on more than two occasions, and the unmarked animals are marked as they are caught. As the number of capture sessions increases, the number (or proportion) of marked animals will increase, so that some sort of regression analysis can be used to extrapolate to total population size. It is, however, also possible to calculate Petersen estimates for each pair of successive capture occasions, and then average the resulting estimates. This has some advantages when the assumption of a closed population is doubtful. Various arrangements for fixing numbers to be examined in

advance of the samplings have been developed for the Schnabel method, but again are mainly of theoretical interest.

One of the major problems in the main capture-recapture models is that probabilities of capture are assumed to be the same for all individuals at any given time. There is a lot of practical experience to show that this is not usually true, so various attempts to make adjustments or corrections have been proposed, but unfortunately none of these modifications seems to work satisfactorily in all circumstances. Three such models are listed in the above classification.

A final variant of Schnabel-like methods is the situation where a number of releases of marked animals is followed by a single capture period. It is of main interest in connection with commercial fishing investigations. The final section of the classification deals with capture-recapture models applied to open populations.

Classification	Applications	References

b. Open populations occur during census period.	Gains and losses to population	
i. Jolly-Seber method	Stochastic model, estimates mortality and recruitment as well as population size.	Section 8.4
ii. Bailey's triple-catch method	Limited to 3 sampling periods, simplest instance of more general theory.	
iii. Fisher-Ford method	Deterministic model ("trellis" arrangement of data).	Cormack (1968:476)
iv. Manly-Parr method	Avoids assumption that all individuals have same survival rate as required in Seber-Jolly method	Seber (1982:282)
v. Regression method	Mainly used for survival estimation	Seber (1982:237)
2. Change-in-ratio method	Uses change in ratio induced by known removals. Principal use with harvests, but much wider potential scope.	
7. Bounded count method	Adjustment to maximum observed count (assumes finite probability that every member of population can be counted in a single census).	

As soon as it becomes necessary to assume a population "open" to additions and losses during the study period, the necessary models become quite complex. The main method in current favor is due to independent work by Seber and by Jolly. It requires that individual tag-releases be distinguishable, since this information is used to estimate additions and losses to the population over time. The earlier methods are mostly now of historical interest, but the Bailey "triple-catch" method is worth consideration as the simplest instance of the more general theory. The Manly-Parr method provides a way to avoid one assumption required by the Seber-Jolly approach.

Two methods of population estimation that do not fit into the above classification are the change-in-ratio method and the bounded-counts method. More experience is needed with both methods to determine their ultimate value. The change-in-ratio method depends on observing some ratio in a population, such as the sex ratio, before and after a removal that is restricted to one of the two classes making up the ratio (e.g., males). The method is conceptually very versatile, and can potentially provide various estimates other than population size, such as recruitment and survival. It also turns out to encompass a variety of other methods, and is thus worth study as a means for understanding the other methods. A practical drawback is that the method will usually be based on observations taken before and after a season of exploitation. Groupings and spatial distribution of the population are likely to change meanwhile, making for various sampling problems.

The bounded counts method depends on the assumption that it is possible, although perhaps with low probability, to see every member of a population in a given survey. In practice, the confidence limits appear to be quite large, and the method may not be very useful.

8. CAPTURE-RECAPTURE METHODS.

8.1 Introduction

This chapter deals with some methods for estimating the absolute abundance of an animal population, using observations of marked individuals. There are two broad uses of marking for population studies. One use is in studying only the marked population with little attention paid to unmarked individuals. Such studies may be concerned with temporal movements (either local or migratory), delineation of geographic range ("home range" of individuals, range of particular population or sub-group of a population), or life history aspects (growth rates, survival rates, age specific reproductive rates, and so on). The other use is that of concern here, in which the primary interest is in the change in observed proportion tagged (which may initially be zero) as tagging progresses. The basic principle is the same in the change-in-ratio method. An important distinction is that the more complicated capture-recapture methods do make use of information as to the identity of single individuals. A distinction is made between "single-recapture" and "multiple-recapture" methods. In the early studies, the basic assumption for virtually all of the methods was that each and every individual in the population has the same probability of capture in any given sampling. Under such an assumption, any individual behaves like every other individual, and information that an individual has been caught once before or many times is largely irrelevant. One might equally well replace a marked individual with one from another population insofar as the theory of the method is concerned.

The assumption that individuals all behave alike insofar as capture is concerned is not very acceptable to most experienced field workers. In most cases, it is clear that the assumption is not realistic, and the issue is whether, in a given situation, the resulting bias can be tolerated. Various aspects of this problem will turn up in what follows. It is worth mentioning here, though, that one important means of testing the assumption of equal probability of capture does depend on the history of capture of individuals -- if individuals do not behave identically, then their past history gives information on that fact. In some models such information is used to produce an improved estimate of population size.

In describing the various methods, perhaps the most important aspect is that of whether the population is "open" or "closed" to those factors that may produce changes in the size of the population during the course of a capture-recapture study. Individuals may move into and out of the study areas, some may die, and others may be born or otherwise "recruited" to the population (in entomological studies, transfers between instars, pupation, etc., are additional such factors). A "closed" population with constant probability of capture permits very simple analyses, mostly based on the binomial distribution. Although such populations may not exist in practice, such a model provides a useful starting place, and may at times be adequate.

"Open" populations produce many more complications, particularly if they are small or if small changes are important. Then the models need to take into account chance effects, that is, stochastic (as opposed to deterministic) models need to be employed. Only quite recently have fully stochastic models been developed; fortunately their application in practice is less difficult than many of the previous methods. We are thus currently at a stage where only one general method may need to be considered in many practical situations. However, since that model also depends on the same unrealistic assumption of equal probability of capture, further

developments are needed, and the practicing ecologist will have to spend a great deal of effort in checking and cross-checking his estimates.

In some cases, the simpler methods may be adequate, or constitute about all that can be done under the circumstances of the study. Occasionally they may shed some light on particular aspects of a problem, or weakness in a necessary assumption that may not be apparent in the analysis of the more sophisticated form. Students will also need an understanding of the essentials of the various methods to understand and appreciate much of what they will find in the literature on a particular species. Thus a number of methods will be described here.

8.2 Petersen's method (Lincoln Index)

While the first recorded use of this method has been ascribed to Laplace in the 16th century (Cormack 1968), there are two commonly cited origins for its use in fish and wildlife work. Most fisheries workers know the technique as "Petersen's method" due to a suggestion by C.G.J. Petersen in 1896. Wildlife workers tend to refer to "Lincoln's Index" due to its use by F.C.Lincoln in efforts to estimate North American waterfowl abundance in the 1930's. A third early use was by C.H.N.Jackson (1933) in his studies of tsetse flies in Africa.

The method requires only two census periods, one involving the initial marking of M individuals, of which m are recovered in the n animals caught on the second occasion. If the population is closed (i.e., there are no gains or losses due to immigration, emigration, mortality, etc.), then it can be intuitively supposed that the fraction marked in the population (M/N) may be estimated by the proportion of marked animals (m/n) found in the second sample; that is:

$$m/n = M/N \quad \text{and} \quad \hat{N} = Mn/m \quad (8.1)$$

The relevant probability distribution is the hypergeometric distribution.

The assumptions necessary to the method can readily be understood by reference to the model resulting in a hypergeometric distribution. One description is via an "urn" model. Suppose we have a vessel of some sort (an "urn") containing N objects, M of which bear distinguishing marks (tags). If the objects are thoroughly mixed, one is removed and recorded (but not returned), then the remaining objects thoroughly mixed again, another removed and recorded, the objects mixed, another removed, and so on until n have been removed (m of which are marked), then the hypergeometric distribution serves to describe the behavior of the random variable m , the number of marked individuals recovered in a sample of n . In practice, n is also a random variable, but the usual approach is to consider the results to be "conditional" on the n actually observed, or to suppose that we sample on the second occasion until exactly n individuals are examined for marks.

One advantage of the urn model is that it helps clarify the position as regards random sampling. In the urn model there is no requirement that the M marks be put on according to any special scheme; the only requirement is thorough mixing before each draw (equivalent to random selection of the n animals taken on the second occasion). For all practical purposes, one thorough mixing is enough, so that one can infer the essential assumption to be one of either a random marking or a random recovery. Both are not required, as has been stated in the literature. In fact, it can be shown that the only essential feature is that the methods of capturing individuals need only be such that the individual probabilities of capture on the first occasion

are independent of those existing on the second (Junge 1963). Thus one might attempt to put tags on by, say, trapping and use another method (e.g., hunting) for recoveries. A difficulty with such an approach is in assuring that the two methods do in fact result in independence of the two sets of probabilities. This cannot be ascertained from two sampling periods alone. Another, less crucial, limitation is that the variance formulas (given below) do not apply unless capture probabilities are equal over all members of the population on the recapture occasion.

A summary of the assumptions is as follows:

- (1) The marks (tags) are not lost and are always identified on recapture.
- (2) The population is closed (but this assumption can be modified).
- (3) Every individual has the same probability of capture (at recovery time).

Assumption (2) can be modified in two ways:

(2a) There are losses of tagged and untagged animals which occur at the same rate, but there are no additions to the population. This does not change the proportion tagged and the estimate of population size remains valid but applies only to the population at the time of first sampling (tagging).

(2b) There are gains to the population between initial tagging and recovery of tags, but no losses, and the probability of capture is the same for all individuals during the recovery period. If there are no losses, then at the time of recovery, there are still M tagged individuals in the population, and the proportion tagged (m/n) estimates what fraction of the current population carries tags, so that the estimate of eq.8.1 applies to the population at the time of recovery. Note that (2a) and (2b) thus estimate the population size at tagging and at recovery.

As we have already noted, assumption (3) can be modified as:

(3a) Marking at random.

(3b) Independent probabilities of capture at both marking and recapture (which necessarily includes (3a) as a special case).

These several modifications depart from the conditions necessary for the hypergeometric distribution to hold, and thus prevent strict applications of the relevant variance formulas. One simple way to obtain a useful estimate of the variance is to randomly subdivide the number of animals initially marked (M) into several subgroups, and to estimate the population size separately for each such subgroup. These independent estimates then provide the data for calculating a valid variance. How many subgroups to use depends on the number marked initially (M) and the fraction recovered, since there are obvious drawbacks in having any of the subgroups result in no recoveries of marked individuals. Presumably one might tolerate one such group, using the modified estimation formulas given below.

From a formal statistical point of view, the estimate of eq.(8.1) has the drawback of having an "infinite bias". This results because there is always a finite probability that $m=0$ (i.e., no marked animals are caught on the second occasion). Chapman(1951) proposed an adjusted equation to circumvent this difficulty:

$$\hat{N}_C = \frac{(M+1)(n+1)}{(m+1)} - 1 \quad (8.2)$$

which has variance estimate:

$$v^* = \frac{(M+1)(n+1)(M-m)(n-m)}{(m+1)^2(m+2)} \quad (8.3)$$

When N is large, the hypergeometric distribution (sampling without replacement) is very closely approximated by a binomial distribution (sampling with replacement), and when P is sufficiently small, a Poisson distribution also provides an excellent approximation. On the other hand, when P is not very small, the normal distribution may provide an adequate model. Various rules have been suggested as to when to apply the several approximations in practice. Chapman(1948) used m/n as a guide to magnitude of P , and gave the following criteria:

$N < 500$	$m/n \leq .10$	Poisson
	$m/n > .10$	binomial
$500 < N < 1000$	$m/n \leq .075$	Poisson
	$m/n > .075$	normal
$1000 < N$	$m/n \leq .05$	Poisson
	$m/n > .05$	normal

However, other authors used less stringent rules. One of the best ways to gain insight into the differences due to various approximations is to intercompare sets of tables of the distributions for several examples.

The various approximations are particularly convenient in calculating a confidence interval around an estimate from the Petersen method. DeLury (1951) noted that, under the binomial assumption, $P=M/N$, so that the expected value of the random variable (m) here representing the number of successes is:

$$E(m) = \frac{nM}{N} \quad (8.4)$$

with binomial variance:

$$V(m) = \frac{nM}{N} [1-(M/N)] \quad (8.5)$$

so that if we substitute m/n as an estimate of P , an estimate of the variance of m is just $m[1-(m/n)]$. Going one step further, and assuming m is approximately normally distributed, approximate 95 percent confidence limits for m are:

$$m \pm 2 \left[m \left(1 - \frac{m}{n} \right) \right]^{1/2}$$

and DeLury inserted these values in the Petersen estimate (eq.8.1) to provide confidence limits on N , i.e.,

$$\text{Upper limit} = \frac{nM}{\{m - 2[m(1 - m/n)]^{1/2}\}} \quad (8.6)$$

$$\text{Lower limit} = \frac{nM}{\{m + 2[m(1 - m/n)]^{1/2}\}}$$

An alternative way to proceed is to write the estimate (eq.8.1) as $N = M/p$, where $p = m/n$, and to use tables or graphs to find confidence limits on P in order to calculate upper and lower limits for N from those for p .

Still another approach (Leslie 1952) can be described by changing eqs. 8.4 and 8.5 from those representing the random variable m , to those for a random variable multiplied by a constant:

$$E(aX) = aE(X) \quad \text{and} \quad V(aX) = a^2V(X)$$

where a is a constant. If we use $a = \frac{1}{nM}$, then:

$$E\left(\frac{m}{nM}\right) = 1/N \quad (8.7)$$

$$V\left(\frac{m}{nM}\right) = \frac{1}{nNM} \left[1 - \frac{M}{N}\right]$$

but since N is unknown, we estimate the variance by replacing $1/N$ by m/nM (eq.8.7 justifies this), and obtain:

$$s^2\left(\frac{m}{nM}\right) = \left[\frac{m}{(nM)^2}\right] \left[1 - \frac{m}{n}\right] \quad (8.8)$$

whereupon, assuming $1/N$ to be approximately normally distributed, one can obtain approximate 95% confidence limits on $1/N$ as:

$$\frac{1}{N} \pm 2 s\left(\frac{m}{nM}\right) \quad (8.9)$$

The main advantage here is that estimators of N are quite skewed as can be seen from sampling experiments, or by considering $p (=m/n)$ to be normally distributed and reflecting what the distribution of $1/p$ will look like. It turns out that $1/N$ is much more symmetrically distributed, hence confidence limits expressed as in eq.(8.9) are presumably less biased than those previously described here.

Assuming a binomial distribution (rather than the hypergeometric) leads to a slightly different correction for bias in the estimation equations. Bailey (1952) suggested:

$$\hat{N}_B = \frac{M(n+1)}{(m+1)} \quad (8.10)$$

with variance estimate:

$$v_1 = \frac{M^2(n+1)(n-m)}{(m+1)^2(m+2)} \quad (8.11)$$

The difference between eq.(8.2) and eq.(8.10) is clearly very small. Bailey (1952) showed that N as estimated from eq.(8.1) tends to overestimate, having a bias "of order $1/m$ ", while eq.(8.10) has a bias of order e^{-m} .

Cormack (1968:460) and Seber (1982) provide convenient summaries of a number of schemes to avoid biased estimates through "inverse" sampling. These schemes require sampling on a second occasion to continue until some pre-

determined event happens, e.g., until exactly m marked animals are caught. In many field situations such schemes are very difficult to carry out, and, as previously indicated here, the crucial source of bias is that having to do with unequal probabilities of capture, for which satisfactory corrections are presently difficult. In most practical situations, the investigator should use eqs. (8.2) or (8.10). If the numbers of recaptures (m) is small enough to make the theoretical biases in estimation important, it will also be true that the estimates will be highly variable, and thus will provide very little information on the population under study in any case.

Example 8.1 Petersen method

Nixon et al. (1967) trapped and marked squirrels (*Sciurus*) in Ohio in 1962. In their first day of trapping, 22 individuals (M) were caught, while on the second day, 13 were caught (n). Seven of these were marked (m), having also been caught on the first day. From eq. (8.1), we have:

$$\hat{N} = \frac{22(13)}{7} = 40.9, \text{ while Chapman's equation (8.2)}$$

gives:

$$\hat{N}_c = \frac{23(14)}{8} - 1 = 39.2, \text{ with variance estimate:}$$

$$v(\hat{N}_c) = \frac{23(14)(15)(6)}{64(9)} = 50.31$$

DeLury's estimated confidence limits are (eq. 8.6) calculated from:

$$\frac{13(22)}{7 \pm 2[7(1 - \frac{7}{13})]^{1/2}}$$

which gives: $28.0 \leq N \leq 84.0$. Leslie's approach (eq. 8.8 and 8.9) gives:

$$s^2\left(\frac{nm}{M}\right) = \frac{7}{[13(22)]^2} \left(1 - \frac{7}{13}\right) = 0.0000395$$

and $0.0119 \leq 1/N \leq 0.0370$ which is useful if one has some interest in the reciprocal of population size. Inverting gives essentially the same result as DeLury's approach. Bailey's method (eq. 8.10) gives:

$$\hat{N}_B = \frac{22(14)}{8} = 38.5 \text{ with variance estimate (eq. 8.11):}$$

$$s^2(\hat{N}_B) = \frac{(22)^2 14(6)}{64(9)} = 70.6.$$

8.3 The Schnabel method

We now consider a situation wherein sampling is not restricted to two occasions, and all of the unmarked animals caught in successive samples are marked and returned to the population. In most applications, the marks only serve to indicate that the animal has been caught previously and do not identify individuals. The

method was first proposed by Schnabel (1938) and further studied by Schumacher and Eschmeyer (1943) who provided a variance estimate and an alternative estimation formula (also derived by Hayne 1949). The basic assumptions are those previously given, i.e., (1) marks are not lost nor missed, (2) the population is closed, and (3) constant capture and recapture probabilities.

A very convenient way to visualize the process and to derive the various equations is that of DeLury (1958). An initial sample serves to introduce some marked individuals into the population, and then k further samples are taken (giving $k+1$ sampling periods in all). Estimates are obtainable for population size in each of the subsequent sampling occasions but not for the first (although the assumptions do, of course, imply that the population is of constant size throughout the study). It is worth noting that there is no requirement that the first set of marks be put on at random (i.e., that all individuals have the same probability of capture). Thus if it is feasible to mark a substantial number of animals by some inexpensive but obviously biased method, the investigator might profitably do so, and then revert to some more expensive means of capture that is more in line with the assumptions for the k subsequent recapture periods.

The notation used here is as follows:

M_i = number of marked individuals in the population just before the i^{th} sample is taken; $i = 0, 1, 2, \dots, k$ so that M_0 is the number marked on the first occasion.

n_i = number of individuals caught on i^{th} sampling (since, in most studies, $n_0 = M_1$, we will be concerned here with n_1, n_2, \dots, n_k).

m_i = number of tagged individuals caught in the i^{th} sample; $m_0 = 0$, and we consider m_1, m_2, \dots, m_k .

In any given sampling (after the initial marking), m_i/n_i gives an estimate of the proportion marked in the population. If sampling is random with respect to whether or not the animal is marked (i.e., a constant probability of capture holds), then we have a binomial-type situation applying at the time of sampling, and we can write (with $p = \frac{m_i}{n_i}$):

$$E\left[\frac{m_i}{n_i}\right] = \frac{M_i}{N_i} \quad (8.12)$$

$$V\left[\frac{m_i}{n_i}\right] = \frac{M_i}{N n_i} \left[1 - \frac{M_i}{N}\right] \quad (8.13)$$

Since N is unknown, it is necessary to use m_i/n_i to estimate the variance, as is usual in dealing with samples from a binomial distribution. The data from a capture-recapture study can be conveniently plotted with m_i/n_i on the vertical axis and M_i along the horizontal, and ideally should constitute a series of points (m_i/n_i and M_i) scattered about a straight line through the origin. Solutions to the problem of estimating N depend on the choice of methods for fitting a regression line to the data. That is, eq.(8.12) can be represented as a straight line through the origin with slope

equal to $1/N$; if one writes $y_i = m_i/n_i$ and $x_i = M_i$, then the line $y_i = bx_i$ is equivalent to eq. (8.12) with $b = 1/N$.

Since the "y" values (m_i/n_i) are subject to sampling (chance) errors, a weighted regression scheme is indicated. One choice is to use reciprocals of the variance estimates [eq.(8.13)] as weights (thus the more precise data have greater weights). This procedure leads to Schnabel's original formula, which has to be solved by iterative (trial-and-error) methods; that is one finds a value of N most nearly satisfying:

$$\sum m_i = \sum \frac{(n_i - m_i)M_i}{\hat{N} - m_i} \quad (8.14)$$

for which Schnabel gave an approximate solution as:

$$\hat{N} = \frac{\sum n_i M_i}{\sum m_i} \quad (8.15)$$

However, DeLury pointed out that one of the common features of actual application of the method is that of a tendency for tagged individuals to be grouped or clustered in the habitat, which makes eq.8.13 a poor variance estimate (it underestimates). He therefore proposed weighting by the "sample size" (n_i) at each point in time. This gives the simpler Schumacher-Eschmeyer formulation:

$$\hat{N} = \frac{\sum n_i M_i^2}{\sum m_i M_i} \quad (8.16)$$

A variance for the estimate (8.16) is calculated in the same manner as for weighted regression equations. This gives:

$$s^2 = \left\{ \sum \left(\frac{m_i^2}{n_i} \right) - \frac{(\sum m_i M_i)^2}{\sum n_i M_i^2} \right\} / (k-2) \quad (8.17)$$

and confidence limits are calculated from:

$$C.L. = \frac{\sum n_i M_i^2}{\sum m_i M_i \pm t_\alpha [s^2 \sum n_i M_i^2]^{1/2}} \quad (8.18)$$

where the \pm sign determines lower and upper limits respectively, and t_α refers to the value obtained from t-tables for selected α and $k-2$ degrees of freedom. An important point here is that a small number of sampling times will result in fairly large values of t_α and hence wider confidence limits than might be obtained with more days of tagging.

Example 8.2 The Schnabel method

The trapping mentioned above (Example 8.1) was continued for a total of 11 days. The data are given below, along with the calculations for eq.(8.16).

Day	Number caught (n_i)	Recaptures (m_i)	Tagged in population (M_i)	$\Sigma m_i M_i$	$\Sigma n_i M_i^2$	\hat{N}
1	22	0	0	0		
2	13	7	22	154	6,292	40.9
3	15	10	28	434	18,052	41.6
4	10	5	33	599	28,942	48.3
5	6	5	38	789	37,606	47.7
6	5	3	39	906	45,211	49.9
7	15	10	41	1,316	70,426	53.5
8	11	6	46	1,592	93,702	58.9
9	18	8	51	2,000	140,520	70.3
10	8	7	61	2,427	170,288	70.2
11	16	10	62	3,047	231,792	76.1

Using DeLury's regression approach to the data, we let $y_i = \frac{m_i}{n_i}$, $x_i = M_i$, and $\beta = \frac{1}{N}$. Thus eq. (8.12) becomes $E(y_i) = \beta x_i$, and a weighted equation using sample sizes (n_i) as weights is:

$$\hat{\beta} = \frac{\Sigma w_i x_i y_i}{\Sigma w_i x_i^2} = \frac{\Sigma m_i M_i}{\Sigma n_i M_i^2} \quad \text{and this estimates the reciprocal of } N, \text{ hence}$$

eq.(8.16).

Calculating a variance estimate (eq. 8.17). we get:

$s^2 = \frac{1}{9} [44.772 - \frac{(3,047)^2}{231792}] = 0.5243$, and 95% confidence limits for \hat{N} , from eq. (8.18) are:

$$\frac{231792}{3047 + 2.26[(0.524)(231792)]^{1/2}} = \frac{231792}{3047 \pm 788.59} \text{ or } 60.4 \leq \hat{N} \leq 102.6$$

Since the last day's trapping turned up 6 unmarked squirrels, there were at least 68 ($M_{11} = 62$ and 6 unmarked) in the population, so the lower limit should be 68. In the next hunting season, 41 squirrels were shot, of which 25 were marked. Using eq.(8.2), we have $N_c = \frac{69(42)}{26} - 1 = 110.5$ which is appreciably above the upper limit. The authors felt that probabilities of capture were not constant, with some individuals being more likely to be recaptured.

Example 8.3 Estimation from frequency of capture

In a situation like that of Example 8.1 and 8.2, where it appears that the population is being underestimated, the best cure no doubt is to identify the

faulty assumption and do something about it. This is not always as easy as it sounds. One way is to mark by one technique and recover by another, e.g., to tag by trapping and recover tags in hunting. However, this approach doesn't necessarily cure the problem (see, for example, Eberhardt et al. (1963:43-47), in which a particular model, the geometric distribution was postulated for recaptures). There is, of course, no assurance that this model should hold widely. Seber (1982:Ch.4) summarizes the theory and gives some other models that might be used. Eberhardt (1969) found that the geometric distribution did seem to fit a wide range of recapture data (40 sets on 10 species). An example of application of the method is available in the paper by Edwards and Eberhardt (1967). A series of taggings were carried out on a population of cottontail rabbits confined to a 40 acre pen in Ohio, in the fall of 1961. Data for a Schnabel census are set forth below.

<u>Date</u>	No. of captures n_j	Recaptures m_j	Tagged in Population M_j
Oct. 24	9	0	0
25	8	2	9
26	9	6	15
27	14	3	18
28	8	4	29
29	5	4	33
30	16	8	34
31	7	4	42
Nov. 1	9	3	45
2	3	2	51
3	8	7	52
4	14	5	53
5	2	1	62
6	5	0	63
7	11	5	68
8	0	0	74
9	5	5	74
10	9	7	74
Totals	<u>142</u>	<u>66</u>	<u>76</u>

Students should carry out the Schnabel calculations in order to gain familiarity with the method. Edwards and Eberhardt (1967:Table 3), using eq. 8.16, obtained a population estimate of 97 animals. The actual population in the pen was 135 rabbits, previously caught by drive-netting (to avoid previous experience with box-traps, which were used for the experimental study of capture-recapture methods), and introduced on October 19, and 29, 1961.

Frequencies of capture were:

<u>Number of times caught</u>	<u>Number of rabbits</u>	<u>Number of captures</u>
1	43	43
2	16	32
3	8	24
4	6	24
5	0	0
6	2	12
7	1	7
	76	142

The underlying model for the frequency distribution (geometric distribution) is a very simple one:

$$f(x) = pq^x \quad (x=0,1,2,\dots)$$

where $q=1-p$, and p is the probability that the animal will not be caught at all, i.e., $f(0) = p$. Strictly speaking, the geometric distribution applies to a conceptually infinite series of trials, and can be at best an approximation to reality. Seber (1982:Ch. 4) can be consulted for various other theoretical difficulties and for the nature of the approximations on which the method rests. In the present instance, there were 18 trapping days and the maximum number of time any individual was caught was 7. In some situations, when the maximum number of captures approaches the number of trapping occasions, an adjustment for truncation may be needed (the number of trapping occasions sets an upper limit on the possible number of recaptures). Seber (1982:172-174) gives a method for doing this and uses the data of Example 8.2 above to illustrate the method.

The essentials for estimation by the frequency of capture method are as follows:

$$\hat{q} = \frac{s-r}{s-1} \quad \text{and} \quad \hat{N} = \frac{r(s-1)}{s-r}$$

where r is the number of individuals that are caught s times. Referring to the data above, it may be seen that $r = 76$ and $s = 142$. Estimates thus are:

$$\hat{q} = \frac{142-76}{141} = 0.468 \quad \text{and} \quad \hat{N} = \frac{76(141)}{142-76} = 162.4$$

In this instance, the method thus overestimates the known population. The estimates above can be used to set up a goodness of fit test by calculating expected numbers as:

$$E(n_x) = Npq^x \quad x=1,2,\dots$$

and introducing estimates of \hat{N} and \hat{q} . This yields the following results:

<u>Number of times caught</u>	<u>Original number of rabbits</u>	<u>Calculated number</u>
1	43	40.42
2	16	18.92
3	8	8.86
4	6	4.15
5	0	1.94
6	2	0.91
7	1	0.42
	76	75.62

These are obtained from $Npq = 162.4(0.5319)(0.4681) = 40.43$ for the first entry, $Npq^2 = 162.4(0.5319)(0.4681)^2 = 18.92$, and so on (multiplying each successive entry by 0.4681). It can be seen that the data are fitted reasonably well. However, students should do a chi-square calculation to check this. In the present example, 135 rabbits were introduced into the pen, so we have $135 - 76 = 59$ in the not-caught (0) category. The expected number is $162.4(0.5319) = 86$, which is substantially larger.

Example 8.4 Mean Petersen method

The Schnabel method depends on the population being closed (i.e., the same population of N individuals is present throughout the study). If this assumption is doubtful or disproven, then it is necessary to use a more complex method in which rates of loss (and/or gain) to the population are estimated. Before doing so, it may be worthwhile to consider a very simple approach, in which Petersen estimates are formed from successive entries in the table of data. That is, referring to the data of Example 8.2, the first 2 days can be used to obtain a Petersen estimate, then the results from day 2 and day 3 can be used, and so on. As noted in Sec. 8.2, the assumption of a closed population can be relaxed somewhat for a Petersen estimate (assumptions 2a and 2b), so that the sequence of Petersen estimates may be used to look for evidence of a trend in the population. If both gain and losses are taking place, the method isn't, strictly speaking, acceptable. However, if day to day changes aren't large, the overall average may be useful. This leads to the "mean Petersen" estimate proposed by Chapman (1952; see also Seber 1982:138).

Estimates are formed according to eq.8.2 and averaged:

$$\hat{N} = \frac{1}{k-1} \sum_{i=2}^k N_i \quad (\text{only } k-1 \text{ estimates can be obtained from } k$$

periods).

Variances can be estimated by averaging the estimates of eq. 8.3 as:

$$v_1(N) = \frac{1}{(k-1)^2} \sum_{i=1}^k v(N_i)$$

or as the variance of the individual estimates:

$$v_2(N) = \frac{1}{(k-1)(k-2)} \sum_{i=2}^k [N_i - \text{ave}(N_i)]^2$$

Students should perform the calculations using the data of Example 8.3.

8.4 Methods for "open" populations

All populations are subject to change, so the methods described thus far are mostly useful under circumstances such that any change in population size is likely to be of minor importance. As we noted earlier, if there are only gains or losses, the Petersen method may give a valid estimate for one of the two sampling occasions. In general, however, one needs to have a method capable of taking into account temporal changes in populations.

Early workers largely dealt with open populations by assuming constant rates of gain or loss. When populations are large, such deterministic models can be quite satisfactory. However, even when a large population is being studied, it usually turns out that some aspects of the study will depend on small numbers, and thus introduce a stochastic element into the analysis. Consequently, a number of models have been developed that have both deterministic and stochastic elements. Some of these models are very complex and require cumbersome or intricate calculations. Versions of a fully stochastic model for open populations were published by G.M.Jolly (1965) and G.A.F.Seber (1965). Cormack (1968) suggested that, inasmuch as virtually identical results were obtained independently by Jolly and Seber, the technique should be called the Jolly-Seber method.

Some of the earliest efforts to deal with open populations arose from the pioneer studies of C.H.N.Jackson on tsetse fly populations in Africa (Jackson 1937,1939,1940,1948). He used two rather different schemes, one (the "positive method") depending on a single release of a large number of individuals followed by a series of samplings in which marked and unmarked individuals were tallied, but no further marking was done (however, marked individuals were released again after capture). In the second, "negative" method, marking was accomplished on a series of occasions but recaptures were tallied only in one final intensive sampling. The negative method uses the greater reduction in returns from the earlier releases (as compared to those from later releases) to estimate survival rates, which are in turn used to estimate the number of marked animals alive in the population at the time of the final large scale recapture sampling. An estimate of population size at the final sampling can thus be obtained from the Petersen formula, but M is now estimated rather than known exactly. Because any immigrants are reflected in the final sampling, it is not necessary to make special provision for measuring immigration (of course the rate of immigration is not estimated).

On the other hand, the positive method may be expected to measure dilution by immigrants since all of the marking is done in the initial survey. Thus the decrease in proportion marked in successive surveys should reflect the effect of immigration (or other sources of unmarked animals). Combining the two methods gives the necessary ingredients for a complete analysis, and this is what the more recently developed methods are designed to accomplish. Although direct use of the Jackson methods is not now recommended, it may happen that one of the two schemes may be useful in special circumstances -- for example, in Jackson's work unskilled assistants

were sometimes used to carry out marking on a broad scale. Bailey (1951,1952) gave improved estimates for Jackson's methods, while Chapman and Robson (1960) described methods to improve on his survival estimates.

A method ascribed to Fisher and Ford (1947) is mainly of historical interest by virtue of its use of a "trellis" diagram to classify recaptures on each day by the dates of release. The method thus uses data on all previous recaptures of individuals. No variance estimate was given. A detailed study by Leslie and Chitty (Leslie and Chitty 1951, Leslie 1952, and Leslie, Chitty, and Chitty 1953) developed a rather extensive approach to dealing with open populations. They assumed that, with small samples, observational data as to various classes (e.g., date last caught) can be represented as a multinomial distribution, and thereby produced a series of estimating equations. One problem is that the solutions are very difficult to obtain if there are a number of sampling periods.

It is intuitively evident that estimation of gains and losses from a population will require a minimum of three sampling periods. Thus "Bailey's triple catch" method (Bailey 1951, 1952) is of interest both as an illustration and a prospective method for either pilot studies or rapid estimates. The various items of data are as follows:

Period(i)	Time	Total captured	Total tagged and released	Marked individuals caught later
0	0		s_0	
1	t_1	n_1	s_1	m_{01}
2	$t_1 + t_2$	n_2		m_{012}, m_{02}, m_{12}

In the above table, m_{01} are those caught in the first time period and recaptured in the second. Some of these appear again in the third period and are labelled m_{012} . The estimates are:

$$N_1 = \frac{s_1 n_1 (m_{02} + m_{012})}{m_{01} m_{12}} \quad (8.19)$$

$$\hat{\lambda} = \exp(\hat{\beta} t_2) = \frac{m_{01} n_2}{n_1 (m_{02} + m_{012})} \quad (8.20)$$

$$\hat{\mu} = \exp(-\hat{\alpha} t_1) = \frac{s_1 (m_{02} + m_{012})}{s_0 m_{12}} \quad (8.21)$$

Variance estimates are:

$$v(\hat{N}_1) = \hat{N}_1^2 \left[\frac{1}{m_{01}} + \frac{1}{m_{12}} + \frac{1}{m_{02} + m_{012}} - \frac{1}{n_1} \right] \quad (8.22)$$

$$v(\hat{\lambda}) = \hat{\lambda}^2 \left[\frac{1}{m_{01}} + \frac{1}{m_{02} + m_{012}} - \frac{1}{n_1} - \frac{1}{n_2} \right] \quad (8.23)$$

$$v(\hat{\mu}) = \hat{\mu}^2 \left[\frac{1}{m_{12}} + \frac{1}{m_{02} + m_{012}} \right] \quad (8.24)$$

Note that $\exp(\beta_{t_2})$ estimates the gains to the population in the time interval (t_2) between second and third captures (β is an instantaneous rate, while $\exp(-\alpha_{t_1})$ measures losses in the first time interval (t_1 between initial marking and the second sample (first recaptures). For the method to be strictly valid, it has to be assumed that rates of loss and gain are constant during the study period, so the important estimates are β and α . With this arrangement, one avoids the necessity for having $t_1 = t_2$.

The fully stochastic (Seber-Jolly) models use some additional notation and, in common with many earlier models, require knowledge of the identity of individual animals, or at least the occasions on which individuals are marked, so that in many cases identification of individuals is practically essential. Additional symbols used are as follows:

s_i = marked animals released on the i^{th} occasion (the s_i may be equal to the n_i if all unmarked animals are marked and no individuals are killed in handling or otherwise removed from consideration).

r_i = the number of the s_i that are again caught before the study is concluded.

z_i = number of individuals in the population that have been marked before the i^{th} period and are caught again after the i^{th} period but not during the i^{th} period (this then is a measure of the marked animals known to be present during the i^{th} period but not caught then).

The first estimate required is that of the number of marked individuals (M_i) alive at the i^{th} period:

$$\hat{M}_i = \frac{z_i s_i}{r_i} + m_i \quad (i=1,2,\dots,k-1) \quad (8.25)$$

There are again $k+1$ marking periods, the first (denoted by a subscript of zero) and the last (k) for which there is not sufficient data to estimate M_i . The basis for eq.(8.24) can be seen intuitively by considering the fraction $z_i/(M_i - m_i)$ -- this is the proportion of the marked animals alive at time i that are not caught then but are subsequently caught. Furthermore, out of the s_i released on the i^{th} occasion, r_i are caught later. If the animals behave alike (the key assumption of equality of capture probabilities, again), then clearly these two fractions should measure the same quantity; hence equating them gives:

$$\frac{z_i}{M_i - m_i} = \frac{r_i}{s_i}$$

and rearranging yields the estimate of M_i given in eq. 8.26. Perhaps it should be mentioned that m_i represents number of tagged animals in the catch (n_i) at the i^{th} period, as it has in previous models. Also, we retain the assumptions that tags are neither lost nor misread, as well as that of equality of probability of capture among individuals on each occasion (however, this probability can change between occasions).

Population size on the i^{th} occasion is simply estimated by the Petersen method:

$$\hat{N}_i = \frac{n_i \hat{M}_i}{m_i} \quad (i = 1, 2, \dots, k-1) \quad (8.26)$$

Again, estimates for the first and last periods are not available.

Survival between sampling occasions (ϕ_i) is estimated very simply from the data on M_i :

$$\hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i + s_i - m_i} \quad (i=0, 1, \dots, k-2) \quad (8.27)$$

The denominator is comprised of the \hat{M}_i animals estimated to be alive at the i^{th} trapping plus any newly marked animals actually released at that time ($s_i - m_i$).

The number of animals coming into the population is estimated as:

$$\hat{\beta}_i = \hat{N}_{i+1} - \hat{\phi}_i (\hat{N}_i - n_i + s_i) \quad (i=1, 2, \dots, k-2) \quad (8.28)$$

and the logic of the estimate is evident from its structure. There are N_i animals in the population at the i^{th} sampling of which $n_i - s_i$ are removed (i.e., the n_i caught minus any removals; often $n_i = s_i$ and none are removed by the experimenter). A fraction ϕ_i of these survive to the next period, so the equation estimates the number of "recruits" still alive at the $i+1^{\text{st}}$ sampling.

A fifth estimate, that of the probability of capture at the i^{th} sampling is often useful:

$$p_i = \frac{n_i}{\hat{N}_i} \quad (8.29)$$

Again the logic is straightforward, and an equivalent estimate is $p_i = m_i / M_i$.

The variances of the several estimates are complicated, reflecting the complexity of the underlying theoretical development. Seber (1982:Chapter 5) gave a full treatment. A briefer version with simpler equations appears in the monograph by Pollock et al. (1990). They use corrections for small sample biases of the kind used in eq.(8.2). Most users will no doubt depend on a computer program to estimate variances. A number of programs are available, and the Wildlife Society maintains a Web Page with access to a variety of programs. The address is:

<http://fwie.fw.vt.edu/wsb/>

This page contains a link to the Colorado State University Department of Wildlife and Fisheries and Colorado Cooperative Unit Web Page which provides access to program MARK, currently one of the major programs for capture-recapture and survival estimation. The Appendix to the present course contains a program (JSMP) to do the main calculations and to bootstrap results.

Example 8.5 the Jolly-Seber method

The main difficulty in applying this method lies in understanding and making an accurate tally for the basic tables from which the estimates are made. These procedures can best be understood by starting with Table 8.1, which is a tabulation of the history of capture for 56 Weddell seals. There were 5 sampling or census occasions. On the first ($i=0$), 28 tags were put out. On the second ($i=1$), 12 new animals were tagged, 11 on the third, and 5 on the fourth. None were tagged on the fifth census (but this quantity is not relevant to the estimates, anyhow)

The entries in each census column denote the history of a given animal. Thus, tag number 1 was caught initially, but never seen again, while numbers 11 and 12 showed up on every occasion. Number 30 was not tagged until the second census ($i=1$) and was not seen on the fourth ($i=3$), but showed up again on the fifth visit to the study areas.

The basic table of summary data is Table 8.2, in which the entries are the m_{hi} , which is defined as the number caught in the i^{th} sample that were last captured in the h^{th} sample. The top two rows contain basic data from the actual census trip. The first item is the n_i , total number of animals examined on each census. The second item contains the number of tagged animals that were released at that census. In many studies, all of the untagged animals would be tagged and released, so that s_i , the number of tagged animals released into the populations at each census would be equal to n_i . In the present example, this was not possible, so that of the 63 (n_0) animals observed in the first census, only 28 (s_0) were actually tagged.

The m_{hi} entries in the body of the table start with m_{01} , which is the number of the 28 animals tagged on the first visit ($i=0$) that were found again on the second visit. The next entry in that row ($m_{02}=2$), are those of the 28 tagged animals that did not show up until the third census ($i=2$). The final entry ($m_{04}=1$) is tag number 22, who was observed only in the first and last censuses. All of this data comes from the first 28 tag numbers.

The next row of the table of m_{hi} includes tagged animals last seen on the second census ($h=1$) and then observed on the third, fourth or fifth census. This now brings in the data from tag numbers 29-40, as they were tagged in the second census. Totals of the table row entries constitute the r_h , the number of animals last captured on the h^{th} sample and then reobserved at some time in the future. The final column includes those animals initially tagged but not observed again (i.e., tag numbers 1,2,5,6, etc.). This number is not used in the calculations, but serves as a check, since r_h plus this number equals the number of tagged animals released in the appropriate census (e.g. the 28 of s_0 equals $17 + 11$).

The only really practical way to be sure of one's understanding of this procedure is to actually reconstruct the other table entries from Table 8.1.

A second table is usually prepared as a means of calculating the z_i . This is shown as Table 8.3, which shows the c_{hi} , those individuals caught in the i^{th} sample that were last caught in or before the h^{th} sample. Thus inspection of Table 8.1 shows that $c_{02}=2$, i.e., tag number 3 and 8 meet this criterion. We also have, as a further example, that $c_{04}=1$, since there is only one

individual (tag 22) caught in the first and last censuses only. The sum of the row entries give z_1 , z_2 , and z_3 . These totals should be checked by scanning the tables with the appropriate definition in mind. For example, z_2 is the number of animals caught both before and after the 3rd census ($i=2$). Hence, tag numbers 9,10,13,16,17,22,...,40 qualify (11 in all). One could, of course, just make such a tally directly and not bother with Table 8.3, but it is best to have the cross-check resulting from making the table and then a direct count. Logically, entries c_{01} , c_{12} , c_{23} , and c_{34} should be in the table, but these are just m_0 , m_1 , and m_3 and do not correspond with the definition of z_1 , i.e., caught before and after but not in the i^{th} sample. Once the necessary ingredients are in hand, calculation of the estimates is straightforward from the definitions given in equations 8.25 to 8.29.

Table 8.1. History of capture for Jolly-Seber census.

Tag number	Census number				
	0	1	2	3	4
<u>Tags put out at first census</u>					
1	1	-	-	-	-
2	1	-	-	-	-
3	1	-	1	1	1
4	1	1	-	-	-
5	1	-	-	-	-
6	1	-	-	-	-
7	1	1	-	-	-
8	1	-	1	1	1
9	1	-	-	1	1
10	1	-	-	1	1
11	1	1	1	1	1
12	1	1	1	1	1
13	1	1	-	-	1
14	1	1	1	1	1
15	1	1	1	-	1
16	1	1	-	-	1
17	1	1	-	1	-
18	1	1	1	-	-
19	1	-	-	-	-
20	1	-	-	-	-
21	1	1	1	-	1
22	1	-	-	-	1
23	1	-	-	-	-
24	1	-	-	-	-
25	1	-	-	-	-
26	1	-	-	-	-
27	1	-	-	-	-
28	1	1	1	1	-
<hr/>					
<u>Tags put out at second census</u>					
29	-	1	1	1	1
30	-	1	1	-	1
31	-	1	1	-	-
32	-	1	-	-	-
33	-	1	-	-	-
34	-	1	-	-	1

35	-	1	-	-	-
36	-	1	-	1	1
37	-	1	-	1	-
38	-	1	-	1	1
39	-	1	1	1	1
40	-	1	-	-	1

Tags put out at third census

41	-	-	1	-	-
42	-	-	1	-	-
43	-	-	1	-	-
44	-	-	1	-	-
45	-	-	1	1	1
46	-	-	1	-	-
47	-	-	1	-	-
48	-	-	1	1	-
49	-	-	1	-	-
50	-	-	1	-	-
51	-	-	1	1	1

Tags put out at fourth census

52	-	-	-	1	1
53	-	-	-	1	1
54	-	-	-	1	-
55	-	-	-	1	-
56	-	-	-	1	-

Table 8.2. Tabulation of the m_{hi} , the number caught on the i^{th} sample last captured on the h^{th} sample.

i	1	2	3	4	5		
n_i	63	43	41	42	58		number not
s_i	28	24	24	22	23	r_h	seen again
0	-	12	2	2	1	17	11
1	-	-	11	4	4	19	5
2	-	-	-	11	3	14	10
3	-	-	-	-	15	15	7
m_i	0	12	13	17	23		

Table 8.3. Tabulation of c_{hi} , the number caught in the i^{th} sample last caught in or before the h^{th} sample.

h	i					Total
	1	2	3	4	5	
	-	-	2	2	1	5 = z_1
1	-	-	-	6	5	11 = z_2
2	-	-	-	-	8	8 = z_3
3	-	-	-	-	-	

A program in the Appendix (JSMP) calculates these results, and is illustrated with the above data.

8.5 The Manly-Parr method.

The Jolly-Seber method has largely become the standard method for dealing with open populations. A method devised by Manly and Parr (1968) gives very much the same results, but has not been widely used. With the availability of computer programs for the Jolly-Seber method, the Manly-Parr approach may not receive much further attention. It does have the advantage of simplicity and is worth remembering inasmuch as estimates can be obtained with an ordinary calculator. One can thus explore a set of data with without needing reference or access to a computer. Tabulate the data in a matrix of zeros and ones with dates of observation as columns and records of individuals as rows. For every column (except the first and last) identify those individuals known to be in the population on that date by finding those that were seen before and after that date. Label this group as C_i where $i = 2, 3, \dots, k-1$ (or $1, 2, 3, \dots, k-1$ if you designate the first occasion as 0 as done in the Jolly-Seber notation). Now count the number of individuals that were actually observed on that occasion, and designate them as c_i . Then:

$$\hat{p}_i = \frac{k_i}{C_i} \quad (i = 2, 3, \dots, s-1) \quad (8.30)$$

estimates the probability of capture on the i th occasion, and the population present on that date is estimated as:

$$\hat{N}_i = \frac{C_i}{\hat{p}_i} \quad (i = 2, 3, \dots, s-1) \quad (8.31)$$

The following figure shows Jolly-Seber and Manly-Parr estimates for a number of sets of data from a capture-recapture study of Weddell seals in Antarctica, and shows that the two methods gave virtually identical results.

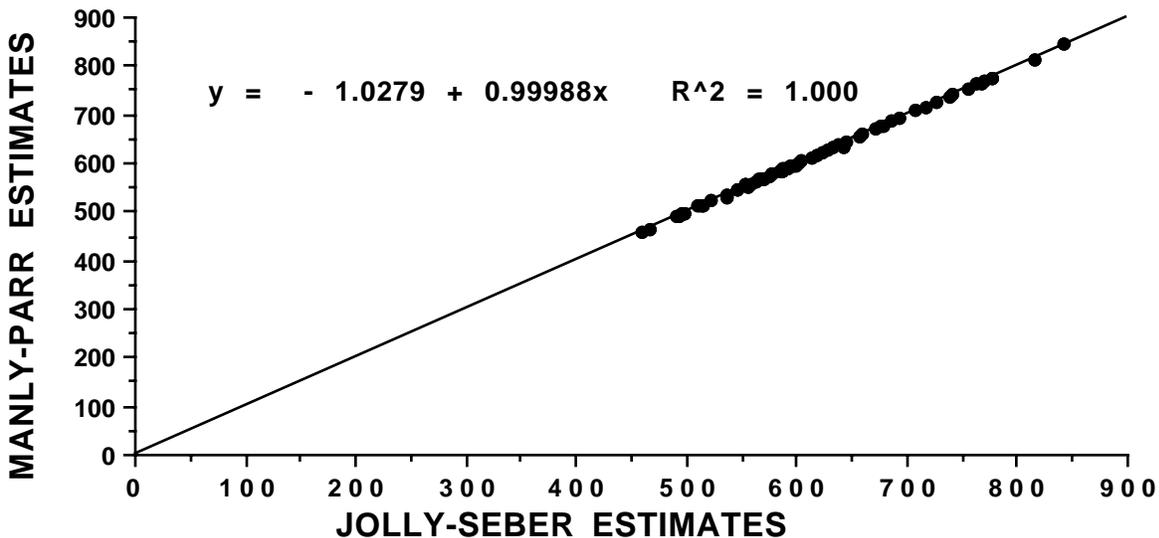


Fig. 8.1. Comparison of Manly-Parr and Jolly-Seber estimates for Weddell seal data.

A segment of data from that study appears below to illustrate the method.

DATE TAG NO.	1	2	3	4	5	6
727	1	0	1	0	0	0
748	1	0	1	1	1	0
818	1	1	1	1	1	1
1030	0	0	1	1	1	0
1040	1	1	1	1	1	1
1274	1	0	0	1	1	0
1286	1	1	1	0	1	1
1288	1	1	1	1	1	0
1445	0	0	1	0	1	0
1541	1	1	1	0	0	0
1577	1	1	0	0	0	0
1590	1	1	1	1	0	0
1594	1	1	1	1	0	1
1620	1	1	1	0	0	0
1893	1	1	1	1	1	0
1901	1	1	0	0	1	1
2018	1	1	1	1	1	0
2097	1	1	1	1	1	1
2285	1	1	1	1	1	1
2593	1	1	1	1	0	0
2619	1	1	1	1	0	0
2673	0	1	1	1	1	1
2700	0	0	1	1	1	1
2708	0	1	0	1	1	1
2717	1	1	1	1	1	1
2807	0	0	0	1	1	1
2930	1	1	1	1	1	0
2945	1	1	1	1	1	1
3010	1	1	1	1	1	1
3454	1	1	1	1	1	1
3511	1	1	1	1	1	0
3585	1	1	1	1	0	0
3685	1	1	1	1	1	1
3714	0	0	0	1	1	1
3923	1	1	1	1	1	1
3949	0	0	0	0	1	1
3999	1	1	0	1	1	1
4071	1	1	1	1	1	1
4190	0	1	1	1	1	0
4220	1	1	1	1	1	1
4224	1	1	1	1	1	1
C_i		31	31	30	22	
c_i		28	27	27	21	
p_i		0.903	0.871	0.90	0.954	

Because this is only a segment of the data, the results should not be taken as actual estimates of seal numbers.

The reason that the Jolly-Seber and Manly-Parr methods give essentially the same results, as suggested in Fig. 8.1, is that the equations for the proportion marked in the population actually converge to give identical values as the number of observations (i) increases. As discussed in Example 8.5 above, in some cases all unmarked individuals cannot be tagged at the time of capture, but instead “batch” tagging may be employed and an individual is treated as though it were newly tagged at the time of first capture. With this change, the main results and calculations are essentially the same, whether all newly caught individuals are tagged at capture, or only those carrying a tag from independent captures are considered as newly tagged. Convergence of the two methods can be shown most conveniently by starting with the estimate of the proportion tagged given below eq. (8.29):

$$p_i = \frac{m_i}{M_i} \quad (i = 2, 3, \dots, s-1) \quad (8.32)$$

introducing the estimate of the number marked in the population given by eq.(8.25), and rearranging to give:

$$p_i = \frac{1}{1 + \frac{R_i z_i}{m_i r_i}} = \frac{1}{1 + \frac{R_i (C_i - k_i)}{m_i r_i}} \quad (8.33)$$

This uses the evident fact that $z_i = C_i - k_i$. As i increases, the number of marked animals released (R_i) eventually becomes equal to the number of marked animals encountered in the i th sample (m_i) because there are no more initially marked animals that have not already been sighted for the first time. Similarly, k_i becomes equal to r_i , giving:

$$p_i = \frac{1}{1 + \frac{C_i - k_i}{k_i}} = \frac{k_i}{C_i} = \hat{p}_i$$

This then ultimately results in identical population estimates by the two methods for “batch” marking.

In the case of “batch” marking, the total population size is estimated by dividing the total number of animals seen (marked and unmarked) by the estimates of proportion seen [Eqs. (8.29) and (8.30)]. Comparisons of the two methods thus depend on eqs.(8.29) and (8.30), as the total number seen will be the same in both cases.

When marking of previously unmarked individuals continues throughout the study (“continued” marking), the Jolly-Seber method estimates total population size as Seber (1982, eq. (5.8,p.200):

$$N_i = \frac{M_i n_i}{m_i} \quad (i = 2, 3, \dots, s-1)$$

When there are no losses at capture, the number released is equal to the number seen, so:

$$N_i = R_i \left[\frac{R_i z_i}{m_i r_i} + 1 \right] \quad (8.34)$$

For the Manly-Parr method with continued marking:

$$\mathcal{N}_i = \frac{R_i C_i}{k_i} \quad (8.35)$$

Again, as R_i approaches m_i and k_i approaches r_i with $z_i = C_i - k_i$ we have equivalence of the two methods.

Example 8.6. Manly-Parr method.

Example 8.5 illustrated the Jolly-Seber method and the same data set is used here to demonstrate calculations for the Manly-Parr method. C_i and k_i are calculated just as illustrated above (page 8.21), with the data of Table 8.1 rearranged below in the same manner as on page 8.21.

1	1	0	0	0	0
2	1	0	0	0	0
3	1	0	1	1	1
4	1	1	0	0	0
5	1	0	0	0	0
6	1	0	0	0	0
7	1	1	0	0	0
8	1	0	1	1	1
9	1	0	0	1	1
10	1	0	0	1	1
11	1	1	1	1	1
12	1	1	1	1	1
13	1	1	0	0	1
14	1	1	1	1	1
15	1	1	1	0	1
16	1	1	0	0	1
17	1	1	0	1	0
18	1	1	1	0	0
19	1	0	0	0	0
20	1	0	0	0	0
21	1	1	1	0	1
22	1	0	0	0	1
23	1	0	0	0	0
24	1	0	0	0	0
25	1	0	0	0	0
26	1	0	0	0	0
27	1	0	0	0	0
28	1	1	1	1	0
29	0	1	1	1	1
30	0	1	1	0	1
31	0	1	1	0	0
32	0	1	0	0	0
33	0	1	0	0	0
34	0	1	0	0	1

35	0	1	0	0	0
36	0	1	0	1	1
37	0	1	0	1	0
38	0	1	0	1	1
39	0	1	1	1	1
40	0	1	0	0	1
41	0	0	1	0	0
42	0	0	1	0	0
43	0	0	1	0	0
44	0	0	1	0	0
45	0	0	1	1	1
46	0	0	1	0	0
47	0	0	1	0	0
48	0	0	1	1	0
49	0	0	1	0	0
50	0	0	1	0	0
51	0	0	1	1	1
52	0	0	0	1	1
53	0	0	0	1	1
54	0	0	0	1	0
55	0	0	0	1	0
56	0	0	0	1	0

Calculations for the above table:

C		15	22	21	
k		10	11	13	
n	63	43	41	42	58
p-tilde		0.667	0.500	0.619	
N-hat		64.5	82.0	67.8	

Here, p-tilde is calculated with eq.(8.30) and N-tilde uses eq. (8.31). A program (JSMP) to do the calculations is in the Appendix. It can be used for both the Jolly-Seber and Manly-Parr methods. A sequence longer than that of this example is required before the two sets of estimates converge. Even with this short sequence, the estimates are in fair agreement. Bootstrapping can be accomplished with the program, and the results for 1,000 bootstraps appear in the following table:

		95% CONFIDENCE INTERVALS		
JOLLY SEBER				
	i	2	3	4
LOWER		48.2	63.5	51.5
UPPER		107.5	214.9	114.4
MANLY-PARR				
LOWER		47.8	58.6	51.9
UPPER		103.2	143.5	105.0
ORIGINAL DATA COMPARED TO BOOTSTRAP MEANS				
	i	2	3	4
JOLLY-SEBER				
N-hat		65.6	100.5	71.0
BOOTSTRAP MEANS		68.2	109.9	74.0
MANLY PARR				
N-tilde		64.5	82.0	67.8
BOOTSTRAP MEANS		67.0	86.9	70.8

The limited amount of data results in wide confidence limits, with only the upper estimates for the 3rd period being much different for the two methods, likely as a consequence of the different estimates for that period from the two methods.

For an example based on a much larger sample, some data on Hawaiian monk seals used by Eberhardt et al. (1999) have been processed with the same program. The results from the "Original data" worksheet follow:

	A	B	C	D	E	F	G	H	I	J	K	L
1	NROWS	C	k	R	m	r	n	M-hat	p-hat	p-tilde	N-hat	
2		225	70	35	81	35	81	83	70.00	0.5	0.5	166
3	NCOLS	116	45	74	45	74	74	116.00	0.3879	0.3879	190.8	
4		20	145	56	79	56	78	81	146.14	0.3832	0.3862	211.4
5			167	69	90	69	90	91	167.00	0.4132	0.4132	220.2
6			188	63	80	63	80	80	188.00	0.3351	0.3351	238.7
7			203	58	68	60	66	69	209.39	0.2865	0.2857	240.8
8			207	77	84	81	80	84	217.50	0.3724	0.372	225.6
9			210	67	67	67	67	67	210.00	0.319	0.319	210
10			206	76	86	80	82	86	216.34	0.3698	0.3689	232.6
11			203	81	92	90	83	92	225.23	0.3996	0.399	230.2
12			204	87	88	88	87	89	206.34	0.4265	0.4265	208.7
13			194	84	94	94	84	94	217.10	0.433	0.433	217.1
14			184	74	84	84	74	85	208.86	0.4022	0.4022	211.4
15			170	72	86	86	72	86	203.06	0.4235	0.4235	203.1
16			157	58	71	71	58	71	192.19	0.3694	0.3694	192.2
17			145	67	79	79	67	79	170.97	0.4621	0.4621	171
18			134	71	82	82	71	82	154.76	0.5299	0.5299	154.8
19			98	61	97	97.00	61	97	155.84	0.6224	0.6224	155.8

Here we see that the estimates p-hat (Jolly-Seber) and p-tilde (Manly-Parr) are virtually identical. Bootstrapping (B = 1,000) gave the results of Fig. 8.2.

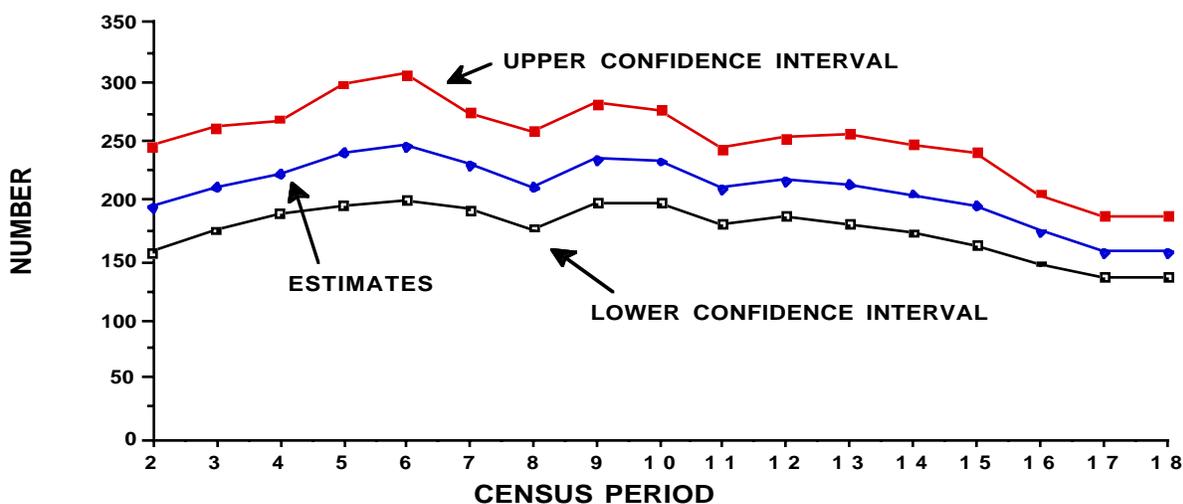


Figure 8.2. Bootstrap confidence limits for census data on Hawaiian monk seals collected on Laysan Island.

The "estimates" in Fig. 8.2 are actually the bootstrap means, but the actual estimates were virtually the same as the bootstrap means, being within 1 or 2 individuals. The original data for this site and for another (Lisianski Island) are on sheets attached to the program (JSMP) described and attached to the Appendix.

8.6. Tests of assumptions

As has been emphasized in previous sections, the crucial assumption in capture-recapture studies is that of equal probabilities of capture among the individuals comprising the population. Two major categories of failure of this essential assumption are:

- (1) The probability of capture inherently varies among individuals.
- (2) The probability of capture changes with exposure to the capture method.

The first class can be subdivided into two obvious groups, one being based on the frequent observation of behavioral differences between sex and age classes. This is a prospect most investigators will have in mind in designing a study, and one which can normally be dealt with by doing the necessary calculations for such groups separately. No doubt there will be other, more subtle, differences among individuals that will result in non-uniformity of response to the capture method, but, as with most features of free-living populations, such "second-order" differences can be neglected for many purposes.

The second sub-division of the first category constitutes the prospect of non-uniform application of the capture method. One of the most likely prospects is the occupancy of a "home-range" or "territory" by many terrestrial species. Unless the capture methods can be uniformly (or randomly) applied to the area under study on each sampling occasion, it is very likely that individuals will have rather different probabilities of capture. Live-trapping provides one example. If traps fall well within a heavily-used part of an individual's home range, clearly that individual can be expected to have a higher probability of capture than would be the case if the trap falls outside the periphery of the usual home range. One obvious precaution is to use a high density of traps relative to home range size; another is to move traps about during the study. Some work has been done on apparent effects of trap density, but very few efforts have been made to study the effect of shifting trapping patterns during the course of a study.

The second category is well-known, usually being labelled "trap-shyness" or "trap-proneness". Certainly there is little doubt that individuals of some species become very skilled at avoiding traps. Trap-proneness needs to be viewed with a little more caution. This is because a uniform probability of capture may nonetheless result in what appears to be an excessive number of captures for some individuals. More explicitly, with probability of capture P and n sampling periods, the probability that a given individual is caught x times follows the binomial distribution, which in turn, if P is small and n large, may be closely approximated by a Poisson distribution. Samples from a Poisson may well give the impression that some individuals are caught unduly often, that is, randomness usually doesn't look "random".

An immediate test to use on recapture data is to compare the frequencies with which individuals are captured with those expected under the hypothesis of a constant probability of capture: either a binomial or the corresponding Poisson

approximation. The simplest such test is the chi-square goodness of fit test. If there is evidence of appreciable gains to or losses from the population during the study then it is not appropriate to apply the test to all of the data. One must instead limit the test to a group of individuals known to be in the population during the period used. This means taking a set of individuals caught early in the study, and again before its conclusion. One thus has a group of individuals caught two or more times and known to be alive in some fixed time interval. Ideally, one would deal with a set of animals all caught on one of the first sampling periods, and then all caught again in the same period near conclusion of the study. In practice it may be necessary to group adjacent periods to obtain enough individuals to make the test worthwhile. An alternative test, suggested by Leslie(1958), is to apply the "binomial dispersion" test. This test compares the observed variability in frequencies of capture with that expected on theoretical grounds if the binomial distribution applies. Leslie(1958) suggested that at least 20 individuals should be available for the test, with 3 or more intervening recapture periods. Very likely the binomial dispersion test is to be preferred over the chi-square goodness of fit test.

Before applying any test to detect a violation of the underlying assumptions one needs to have a rather explicit notion as to what is being tested for, and what alternatives exist. Thus the above test assumes a constant probability of capture throughout the study, and may give an indication as to prospects for inherent differences in "catchability". Those problems associated with "trap-shyness" and its converse may be mainly associated with the first capture (or first few captures) and thus may not be detected in these tests. Also, if there are individuals in the population with essentially zero probability of capture, then no test based on recapture data can detect that problem (unless the true population size is known or otherwise estimated; even so very substantial numbers of recaptures may be required, see Cormack(1966)).

Various tests for the second category of violation of the basic assumption of constant probability of capture are given by Seber(1962,1965), Darroch(1958,1959), Leslie (1952), and Leslie, Chitty, and Chitty(1953). Little use seems to be made of these tests in practice, perhaps as a consequence of the complexity involved in their derivation and description. It is also unfortunately true that none of the tests are very "sensitive", i.e., they do not detect anything but extreme departures from equality of capture probabilities. The program "MARK" referenced in Section 8.4 above provides a wider range of tests, along with a criterion (AIC) for choosing among candidate models.

8.7 Exercises

8.7.1 Plot $y(i)$ and $x(i)$ of Example 8.2 and comment on validity of the underlying model.

8.7.2 Carry out the calculations for a Schnabel estimate on the data of Example 8.3. Plot $y(i)$ and $x(i)$ and comment on validity of the underlying model.

8.7.3 Do a chi-square calculation on the observed and expected frequencies of capture calculate in Example 8.3. The actual population in the pen was 135 rabbits giving $135-76=59$ in the not-captured category. Use chi-square to check this against the expected number and comment on your results.

8.7.4 Mean Petersen method.

Calculate the mean Petersen estimate described in Example 8.4 on the data of Example 8.3. Plot the data and compare with a plot of the Schnabel estimates obtained in Exercise 8.7.2. Discuss your results.

8.7.5 Testing for constant probability of capture

In Example 8.2 (Schnabel), frequencies of capture were:

<u>Number of times caught</u>	<u>Number of squirrels</u>
1	33
2	16
3	10
4	4
5	2
6	3
	<hr style="width: 20%; margin: 0 auto;"/> 68

if the probability of capture is constant, these frequencies should be approximated by a Poisson distribution. Use a chi-square test to check the goodness of fit.

8.7.6 Interpenetrating sampling

When animals are tagged in groups and tend to stay that way (i.e., clumped) until the recapture period, then it is essential that recaptures yield a random sample of the population, or at least that the probability of recapture not depend on whether or not a given individual is tagged. Sometimes this can be accomplished by taking a random sample of locations for recapture. If the recaptures come in groups, then eq.8.3, or any equation based on the assumption of random sampling of individuals, is not realistic. For an example of the bias that may result, students should refer to Example 4.8 and compare the variances obtained there with interpenetrating sampling with what would be obtained by combining all of the recaptures and using equation 8.3.

8.7.7 Double-tagging

Tag loss can be a major source of bias in many circumstances. One way to improve the situation is to apply two tags to each individual. When this is done, then the probability of loss can be estimated from:

$$\hat{p} = \frac{m_s}{m_s + 2m_d} \quad \text{where } m_s \text{ and } m_d \text{ refer respectively to the number of}$$

recaptures carrying only one tag (m_s) and retaining both tags (m_d). A variance estimate is:

$$v(\hat{p}) = \hat{p}^2 (1 - \hat{p})^2 \left[\frac{1}{m_s} + \frac{1}{m_d} \right]$$

and the appropriate Petersen estimate then becomes that of eq.8.1:

$$\hat{N}_d = \frac{Mn}{\tilde{m}} \quad \text{where } \tilde{m} = \frac{(m_s + 2m_d)^2}{4m_d}$$

with approximate variance:

$$v(\hat{N}_d) = \frac{\hat{N}_d^2}{\tilde{m}} \frac{\hat{p}}{(1-\hat{p})^2} + \frac{\hat{N}_d^2}{\tilde{m}}$$

Data for calculations are taken from a fur seal double-tagging study (Abegglen et al. 1958), in which 34,923 male fur seal pups were single-tagged and 5,000 double-tagged on St. Paul Island off Alaska. In 1961, 48,458 three year old males were harvested, yielding the following returns:

2,098 originally single-tagged
258 double-tagged (m_d)
140 single-tagged (m_s)

Students should estimate: (1) \hat{p} , (2) $v(\hat{p})$, (3) \tilde{m} , (4) $v(\hat{N}_d)$, and, (5) calculate an adjusted estimate from the single-tagging using \hat{p} , i.e.:

$$\hat{N}_{adj} = \frac{Mn}{\tilde{m}} (1 - \hat{p})$$

with variance:

$$v(\hat{N}_{adj}) = (1 - \hat{p})^2 v(\hat{N}_d) + \hat{N}_{adj}^2 v(\hat{p})$$

8.7.8 Models for double-tagging

The underlying probability model for double-tagging is quite simple, and students should work out the basis for the above equations. Note that the probability of loss of each of the two tags is assumed to be the same. If the tags are of different kinds (or location, etc.) this assumption should be checked (with chi-square) and separate corrections may be required.

8.7.9 Survival estimation in a three-point census

The simplest census method for an open population requires observations on 3 occasions (an initial marking, a recapture period in which any unmarked are marked and a final capture period). Survival from tagging to the first recapture time is simply estimated as:

$$\hat{\Phi}_{01} = \frac{M_1}{M_0} \frac{m_{02}}{(m_{12} + 1)} \quad \text{with } v(\hat{\Phi}_{01}) = \hat{\Phi}_{01}^2 \left(\frac{1}{m_{02}} + \frac{1}{m_{12}} \right)$$

Data for southern hemisphere fin whales (17th report of International Whaling Commission) are:

Season	Number			$\hat{\Phi}$
	marked*	m_{02}	m_{12}	
1953-1954	118	14	42	0.64
1954-1955	231	40	48	0.77
1955-1956	217	35	12	1.60
1956-1957	129	9	36	0.28
1957-1958	151	16	14	0.74
1958-1959	105	7	4	0.75
1959-1960	56	3	14	

* Recoveries in the same season as the marking are not utilized.

Students should calculate $\hat{\Phi}_{01}$ and the associated variance estimates, which might be compared with a variance calculated from the 7 estimates (which are not, however, independent).

9.0 INDICES OF ABUNDANCE

9.1 Introduction

In theory, ecologists and wildlife managers depend on a range of sophisticated methods for assessing population abundance. In practice, it is often necessary to rely on some sort of index to abundance, supplemented perhaps by occasional use of the technically more satisfactory methods on a few sample areas. Almost always the limitation is simply one of costs. If one must deal with large areas it simply is too expensive to use the better methods on a regular basis. Dice (1941:402) expressed the general idea very well:

"The difficulty of obtaining accurate counts of the number of individual mammals present on a given area has led to attempts to develop indices of abundance for the species concerned. Such indices may or may not be convertible into terms of population density. For many practical uses, however, it is sufficient to know the relative abundance of a particular species in different areas or at different times without having an exact count of the population".

Most of the experience in the 60 years since Dice's statement was published tends to show that exact counts are often not feasible. Estimates of some kind are about all that can be managed, and these are usually difficult to achieve.

One of the risks in using an index is that it may not accurately reflect actual population trends. White-tailed deer (*Odocoileus virginianus*) pellet group counts provide one of the best examples of a useful index of abundance, yet Ryel (1971) found that drive counts on a fenced Michigan area yielded an inverse correlation between counts and index values over 11 years of data. He pointed out various reasons for failure of the pellet group count method used in this instance, and gave a good discussion of ways and means for maintaining quality and integrity of the method.

Three approaches to using indices can be considered: (1) Direct conversion to a census method. The pellet group counts provide one of the better known examples, inasmuch as the counts, under proper circumstances, can be converted directly to an estimate of average numbers of deer present on an area. (2) Calibration through ratio and regression methods, including double sampling. A simple linear relationship between an index and an actual estimate of abundance is used to convert the index to an actual estimate of abundance. (3) Calculation of an improved index or a prediction equation. At times, supplemental information may be used to strengthen an index without converting it to a direct estimate. These aspects are summarized in Fig.9.1.

Several sources of variability need to be considered in appraising an index method. One is stochastic in nature, arising from the variability of the chance fluctuations in the births and deaths that result in change in population size. These changes occur even in the presence of constant birth and death rates, diminishing in importance as the population increases in size. Theoretically, such effects can be neglected if fairly large populations are under study. However, we usually can only study part of large

populations, either through sampling the large population or by counts on relatively small sample areas. Stochastic effects may then become important. A second source of variability arises when birth and death rates are influenced by environmental fluctuations, such as an unusual cold spell or other environmental change during the reproductive season. Sometimes such changes are large enough to be labelled "catastrophic" and the results then are usually dramatic enough to attract attention to specific causes. Less dramatic changes may be difficult to detect. A third source of variability is that engendered by the sampling or observational process through which data on the population are obtained. Trying to sort out these several sources of fluctuations in numbers may be very difficult and deserves more attention than it usually gets.

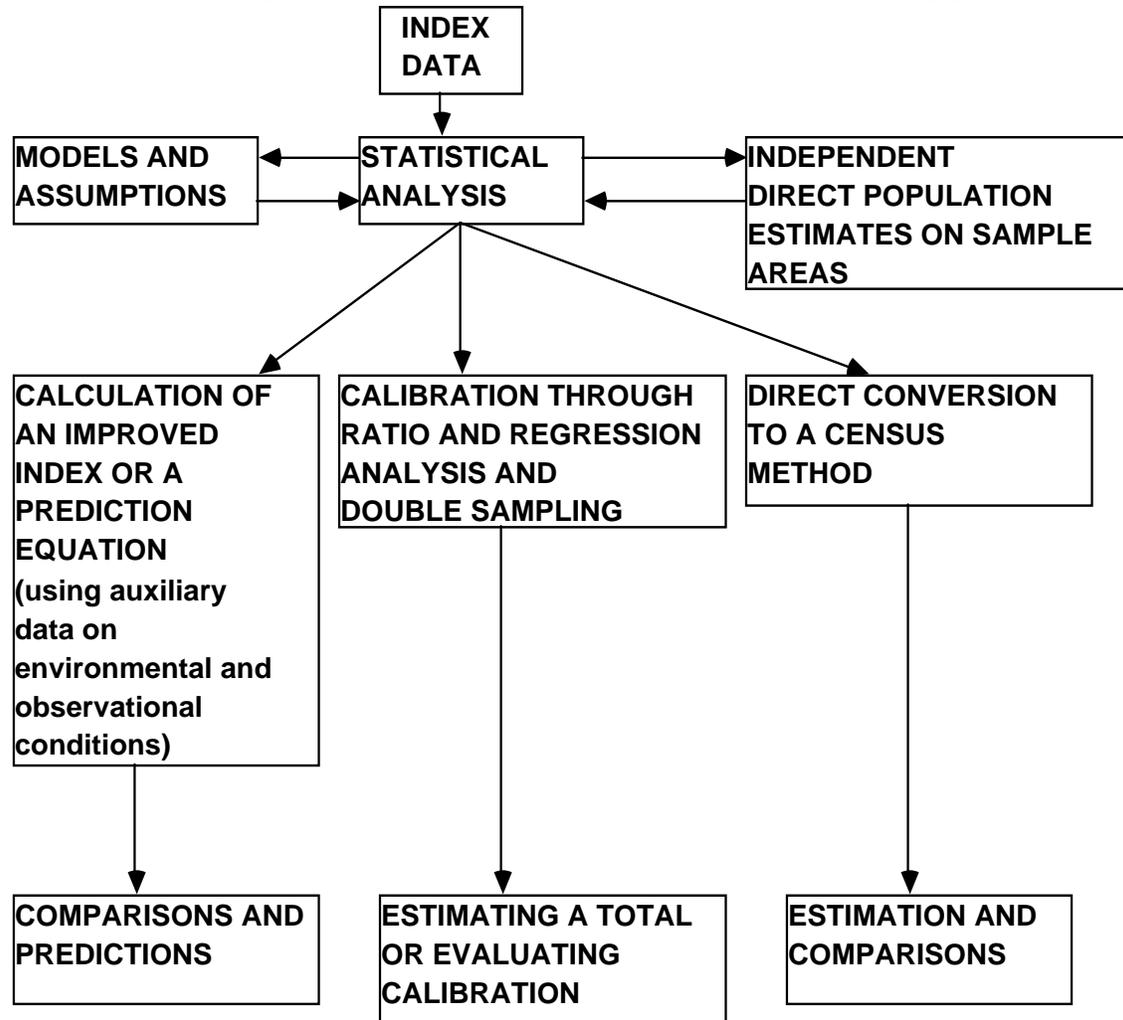


Fig. 9.1 Aspects of the analysis of potential index data.

Indices have received very little attention. The large and extensive reference on estimating animal abundance by Seber (1982) contains only a few pages on indices. This is not Seber's fault, but results simply because there is very little literature on the subject, whereas the stochastic models implicit in tag-recovery and survival data have received much attention in both theory and practice. The presentation here will thus start out by examining a number of sets of data on population trends, and then consider some specific techniques.

Example 9.1 The pellet-group count

One of the better-known wildlife management indices is based on the enumeration of accumulated fecal pellets of ungulates (sometimes politely known as "sign", although the term may also include other kinds of evidence of an animal's presence). Under favorable conditions the method may provide direct estimates of abundance. Such conditions usually depend on a substantial leaf-layer deposited in the fall of the year (or other methods of separating "old" and "new" pellet-groups) plus the over-winter accumulation of groups. Even though many ungulates appear to produce nearly 13 pellet-groups per day (with surprisingly small variance), it is nonetheless true that even an assumption of random distribution of groups still results in the need for rather large numbers of plots being searched to give useful precision in estimation. Thus long accumulation periods are essential. Some workers have resorted to removing pellets from plots (or marking those initially present) in order to be assured of an accurate starting date. Unfortunately, this doubles the fieldwork required.

Although there is a variety of possible sources of error (a major one is simply failure to find all of the groups on a plot), some experience with pellet-group counts on areas where population density could be determined quite accurately by other methods has shown reasonable accuracy in estimating the actual number of deer present. For immediate purposes here, though, we will consider the method as an index. The model is simply:

$$E(x_i) = \beta D_i$$

or that the expected number of pellet-groups per unit area (x_i) is directly proportional to the density (D_i) of animals present on the i^{th} area. The proportionality constant (β) depends on the defecation rate (essentially 13 groups per day), length of accumulation period, and units of measurement of the plots. One further complication is that, if the accumulation period is long (and includes winter conditions and perhaps a hunting season), then there may be an appreciable mortality over the period represented by the counts, so that the estimate obtained is really for an average density.

If independent direct estimates of deer density are available, calibration may be attempted by using ratio or regression methods to convert pellet-group counts to density estimates, i.e., we use a set of direct population estimates (y_i) and a set of pellet-group counts (x_i) on the same areas and turn the above equation around to as to have:

$$D_i = E(Y_i) = (1/\beta) E(X_i)$$

so that an estimate of $1/\beta$ is obtained from the comparisons. However, as remarked above, the pellet-group count can also be converted directly. If the mean number of pellet-groups on $1/50^{\text{th}}$ acre plots on the i^{th} area is x_i , then a direct estimate is given by (Eberhardt and Van Etten 1956):

$$D_i = 50(640)x_i/12.7(\text{days since leaf-fall})$$

where D_i denotes the estimated number of deer per square mile and $1/\beta$ is calculated from the several quantities on the right-hand side of the equation. Hence, pellet-group counts can be used in all 3 of the ways suggested in Fig. 9.1; directly, by calibration, or by direct conversion. The latter two methods should give the same value for $1/\beta$, within sampling error.

Pellet-group counts represent an excellent example, perhaps the best known, of an index method susceptible to exact treatment by statistical methods. This is because there is an exactly definable population (the total number of pellet-groups) available on a discrete area. In most cases, there will be a considerable advantage to be gained by using stratified sampling methods, and the costs of travel versus those associated with actual counting on plots are such as to dictate the use of a cluster of plots at each location (the individual plots cannot be very large due to the effects of plot size on counting errors).

Some experience has indicated that the negative binomial distribution provides a good fit to the observed frequency of pellet-groups per plot, and this finding may be useful in the efficient design of new surveys.

Example 9.2 Conversion factors for pellet-group counts

Part of the Michigan experience with pellet-group counts as a census method includes counts on two fenced areas in which deer numbers are supposedly known with some accuracy. Data from one of these areas (Cusino Enclosure) is as follows (Ryel 1971:124):

<u>Year</u>	<u>Mean number of groups per plot</u>	<u>Known number of deer per square mile</u>
1953	9.403	28.8
1954	2.252	25.0
1955	1.778	28.1
1956	2.246	29.3
1958	<u>2.943</u>	<u>15.5</u>
	12.622	120.7

Using a ratio estimator, we obtain: $\hat{b}_1 = \frac{\sum y_i}{\sum x_i} = 12.622/120.7 = 0.105$.

Data for a direct estimate can be used to calculate: $\hat{b}_2 = [211(19.56)]/[50(640)] = 0.089$. Here, 211 is the average number of days since leaf-fall, and 19.56 is an average number of pellet-groups per deer, adjusted for the sex and age composition of the known number of deer on the area. The direct conversion factor is the reciprocal of b_2 or 11.2, which is somewhat higher than the reciprocal of \hat{b}_2 which is 9.5. One might thus expect to overestimate true deer density by using pellet-group counts, if \hat{b}_1 and \hat{b}_2 are indeed as different as suggested here. There are various other problems in comparisons on this and another such area (George Reserve) and these were described in detail by Ryel (1971).

Example 9.3 Roadside counts

Roadside surveys have been widely used in assessing the numbers of many species, and a useful model for such a survey is that used for pellet-group counts. However, β is not so readily defined in this case, being affected importantly by the behavior of the species, cover (habitat) conditions, along with the weather and various other factors such as time of day (many species are most active in early morning and in the evening). Quite a lot of effort has gone into attempts to standardize roadside surveys by taking counts during specified times and weather conditions. Very likely such standardization techniques are suitable for controlling most factors other than habitat differences. So far, little has been done to try to take into account the effect on visibility generated by different intensities of vegetative cover. One obvious prospect is to attempt to record distances from the observer for each individual animal seen, in the manner of line transects, but a suitable model is needed for use of such a correction.

The net effect of the several uncertainties about roadside and other visual counts (such as aerial surveys) is to make it likely that such methods may be reasonably satisfactory for comparisons from year to year on the same routes, but rather less useful for comparisons of routes in different cover types. Conversion will need to be accomplished by use of ratio or regression methods and independent direct estimates of density on a sample of areas, as there is presently no way to write an equation like that used for pellet-group counts for roadside surveys.

Well-known examples of calling counts are counts of crowing pheasants, or of the drumming of ruffed grouse, cooing of doves, etc. The usual technique is to make counts for a fixed period of time (typically 2 to 5 minutes) at each of a number of stations or "stops". Normally the counts are made along roads as a consequence of the need to cover sizable areas. As with roadside counts, time of day, weather, and seasonal effects are important. Often a degree of standardization is achieved by making frequent counts on a single route at different times of the season. If the population on that route is assumed to remain constant, curves of calling intensity against time of day (and time of season) may serve as a reference standard for adjusting the other counts.

Auditory counts bring in the hearing acuity of the individual observer as an important additional variable. A useful model may then be:

$$E(x_{ij}) = \beta r_i^2 D_j$$

where r_i represents the radius within which a call may be heard by an individual observer. Presumably such a radius may also depend on cover and weather conditions (plus interference from other sounds -- traffic noise being usually the main offender), so there may be a "regional effect" as well as an "observer effect" on the recorded counts.

The following table gives results of hunting success and calling counts for Gambel quail in Arizona.

Hunting success and calling counts for Gambel quail in Arizona (Smith and Gallizioli 1965).

Year	Oracle Junction		Pinnacle Peak		Cave Creek	
	Quail	Call				
	per trip	count	y	x	y	x
	y	x				
1958	3.81	61	3.53	83	--	--
1959	2.70	24	1.37	10	--	--
1960	6.40	103	3.74	94	2.96	72
1961	2.57	25	1.20	22	0.64	8
1962	6.09	75	2.83	59	2.55	64
1963	4.84	62	1.70	25	1.82	36
1964	2.91	41	1.60	15	1.38	26
Totals	29.32	391	15.97	308	9.35	206
$\hat{b}_1 = \Sigma Y / \Sigma X =$						
		.075		.052		.045
$\hat{b}_2 = \Sigma YX / \Sigma X^2 =$						
		.070		.044		.043
Regression slopes						
		.054		.030		.034
Reg. intercepts						
		1.200		.94		.45

Example 9.4 A pocket-gopher example.

Reid et al.(1966) gave an interesting example in which an index (mounds and earth plugs) of pocket-gopher (*Thomomys talpoides*) abundance is compared with actual abundance, as established by trapping-out gophers on sizable plots. They concluded that the relationship between number of gophers present and the index was curvilinear, but did so by plotting the number of gophers per unit area against the number of signs. Such a plot does indeed suggest a nonlinear relationship. However, there were 2 areas (Black and Grand Mesas) studied in 3 years (1962 to 1964).

If we assume the trap-outs to measure absolute abundance, without sampling error, and adopt the model of simple proportionality, $E(x_i) = \beta D_i$, then the appropriate plot is of signs (Y) against gophers per acre (X). Also, if the areas and years are plotted separately (or distinguished by individual symbols), then it seems that the apparent curvilinearity may really be due to differences in β between years and/or areas. It is further evident that variability increases with increasing density of gophers. Hence, we adopt the model: $x_{ij} = b_{ij} D_{ij} e_{ij}$, where x_{ij} denotes number of signs in the i^{th} year and j^{th} area, D_{ij} is actual density of gophers, b_{ij} is the proportionality constant between gophers and sign, and e_{ij} is a proportional "chance" error. Since D_{ij} is assumed to be measured without error (or with negligible error), we can arrange the equation in terms of signs per gopher, i.e., consider:

$$y_{ij} = \frac{x_{ij}}{D_{ij}} = b_{ij} e_{ij}$$

and take logarithms:

$$\log y_{ij} = \log x_{ij} - \log D_{ij} = \log b_{ij} + \log e_{ij}$$

Then the data can be subjected to a simple one-way analysis of variance. The results indicate that there are significant differences and, using a multiple-comparison test (Scheffe 1959) shows that Grand Mesa in 1963 is significantly different from its 1964 value and from Black Mesa in 1962. We are thus fairly confident that the main factor in the apparent curvilinear relationship is really a difference between areas and years, but could not, of course, exclude a behavioral difference as a possibility at low densities. The conclusion here, as with the quail example above, is that it will very likely be necessary to include both spatial and temporal data in any initial efforts to calibrate an index.

Abundance of pocket-gopher sign (mounds and earth plugs on one-acre plots) expressed as logarithm (base e) of signs per gopher (Reid et al. 1966).

Black Mesa			Grand Mesa	
1962	1963	1964	1963	1964
1.91	2.76	2.01	2.03	1.82
2.32	2.23	2.02	1.99	2.06
2.05	2.16	2.18	2.17	1.41
2.02	1.92	2.54	2.24	1.69
1.86	2.36	1.55	2.07	1.54
1.31	2.48	1.40	2.88	1.76
1.58	2.65	1.61	2.18	2.27
1.80	1.88	1.83	2.63	2.12
2.05	1.93	1.43	2.42	2.09
2.04	2.14	2.24	2.54	2.04
1.70		2.20		
1.74				
2.07				
\bar{x} = 1.882	2.253	1.910	2.316	1.880
n_j = 13	10	11	10	10
s^2 = 0.0670	.0947	.1395	.0865	.0780

9.2 Trends in abundance

For at least the larger vertebrates, there are few occasions when data on the full course of the growth of a population are available. Usually only a relatively short segment of the record of population size is at hand. A conceptual model of the overall possible course of events is nonetheless useful for interpreting these shorter segments of data. Ecology textbooks describe long-term population growth by the logistic model. Further details of this model appear in the chapter on population models, but the general shape is illustrated by the trend of an elephant seal population (Fig. 9.2). The logistic curve is characterized by rapid initial population growth that slows down over time, with

the curve ultimately approaching an asymptotic level (often denoted as K). The approach to an asymptotic value may be erratic with large year-to-year fluctuations. Slowing-down of the growth rate usually is associated with resource limitations of one sort or another, often food or space. Such restrictions make populations highly susceptible to year-to-year weather fluctuations.

The logistic curve assumes a constant decline through time of the rate of increase. Although data are limited, evidence for large mammals (Eberhardt 1977b, Fowler 1981) suggest that a different model may be appropriate, with the rate of population growth virtually constant over much of the range, and then slowing down sharply as the asymptotic value (K) is approached. A simple exponential growth curve may then be adequate to describe the initial stages of growth. The essential feature for present purposes, however, is the overall sigmoid shape of the curve -- concave-upwards in the early stages, and concave-downwards as growth slows down.

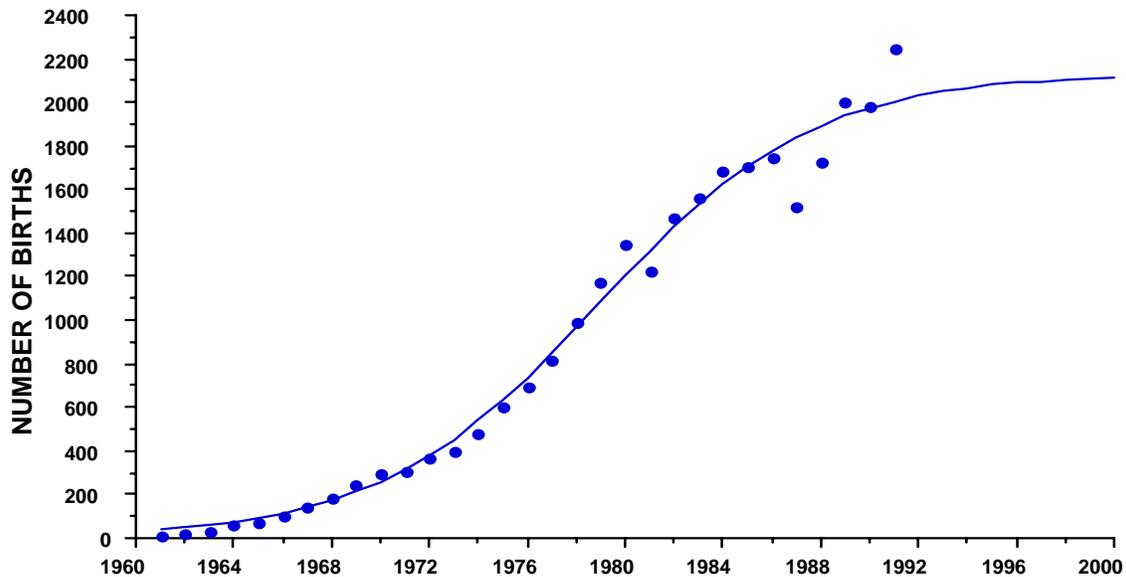


Fig. 9.2. Growth of an elephant seal population (Stewart et al. 1994) with a logistic curve fitted by non-linear least-squares. Data are number of births tallied on colonies on Ano Nuevo Island and the adjacent mainland.

If one accepts generality of the sigmoid curve illustrated in Fig. 9.2, then the analysis of a shorter series of observations will obviously be influenced by the position of the data segment on the overall curve. A basic criterion is simply whether a curve drawn through the data segment is concave-upwards (initial growth stages) or concave-downwards (approaching the asymptotic level). If the early growth segment essentially follows an exponential curve, then plotting logarithms of the counts against time should yield a straight line. An example appears in Fig. 9.3.

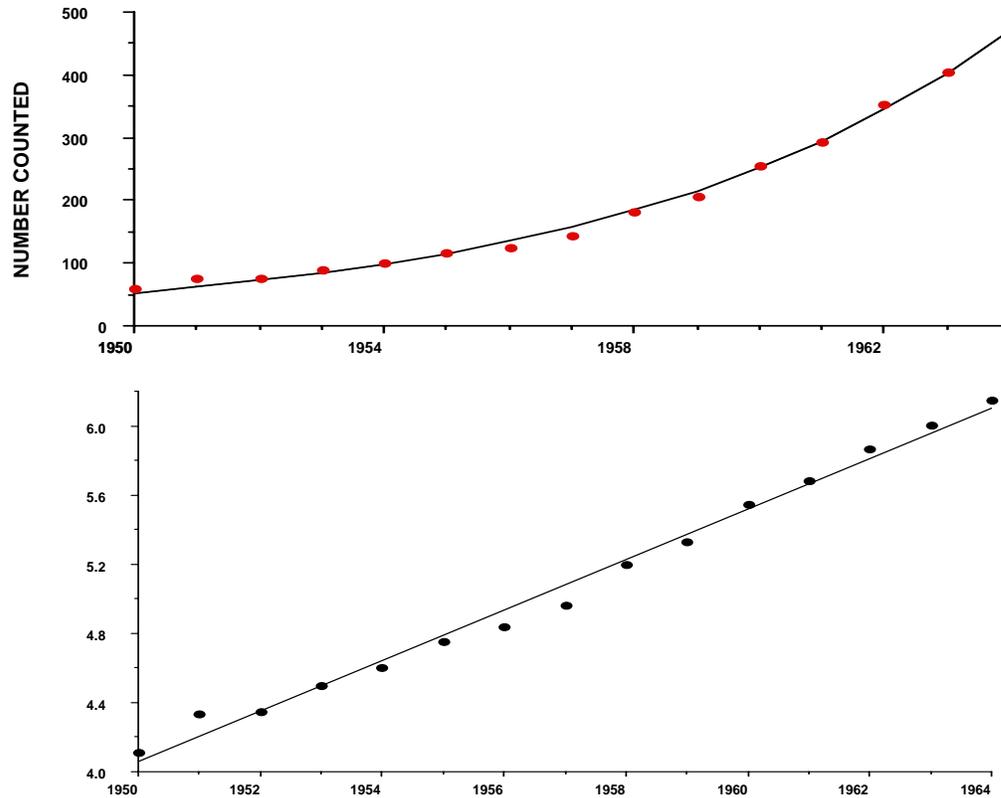


Fig. 9.3. Growth curve for muskox (upper panel) with an exponential curve fitted to numbers counted and a straight line fitted to logarithms of number counted below Data from Spencer and Lensink (1970).

Many sets of population trend data are much more variable than that of Fig. 9.3, and we thus need to consider statistical tests. A simple test for curvilinearity appears in Section 1.7. Various monte carlo simulations were conducted by Eberhardt (1992) in an attempt to determine whether a segment of population trend data could be demonstrated to be above or below the inflection point on a curve like that of Fig. 9.2 (the inflection point divides a growth curve into 2 segments; the slope of the curve increases up to the inflection point and decreases beyond it). A two-stage test was developed. The observed, untransformed, numbers of the trend index are checked for curvilinearity in the first stage. If this test is non-significant, then the data are log-transformed and the test again applied. The first stage examines data of the form of the left side of Fig. 9.3. The basis for the second-stage test is the fact that the exponential curve becomes a straight line under log-transformation (lower graph of Fig. 9.3), so should not show a significant departure from linearity in the test. However, the right-hand limb of Fig. 9.2 remains curved under the log-transform.

9.3 Testing significance of trend lines

In exhibiting curves like that of Fig. 9.3, the exponential curve was used. Growth of large vertebrates may better be described by a closely related curve, the geometric (see

Fig. 1.10 and the accompanying discussion). Either curve can be represented by the model of eq.(9.1) by letting $\lambda = e^{rt}$ or $\lambda = (1+r)^t$, as discussed in Section 1.10, so we use the following simple model for a population growing at a constant rate.

$$N_t = N_0 \lambda^t \quad (9.1)$$

where N_t represents abundance at time t , and λ is a measure of the rate of change of the population. Taking natural logarithms converts this model to a linear equation, so we expect the trend to be a straight line if the index used is the logarithm of observed numbers:

$$\log_e N_t = \log_e N_0 + t \log_e \lambda \quad (9.2)$$

We thus tend to plot logarithms of observations of abundance, and to do various kinds of linear regression analyses in exploring the data. In the present section, we will examine a number of sets of data on actual populations and try to infer something about trends from simple regressions.

To study trend, we fit eq.(9.2) by linear regression methods, rewriting it as $y = a + bx$, where $y = \log_e N_t$, $x = t$, and $b = \log_e \lambda$. For purposes of calculation, t will be used here as the sequence of years, 1,2,3,... We are mainly concerned with the slope (b) and variance about regression, so that the observed sequence of dates might serve equally well, i.e., we could use 1967, 1968, 1969, etc. in regression calculations. However, this should not be done in practice due to the prospect of introducing "roundoff" errors in the regression calculations when the x -values are sizable numbers (i.e, one should use the sequence 1,2,3.. for calculations and later plot data against the actual years, 1967, 1968, 1969). A key measure of variability is the "regression mean square", or variance about regression, often written as MS_{reg} . It can be calculated by fitting the linear regression and calculating:

$$MS_{reg} = \{\sum [y_t - (a + bt)]^2\}/(n - 2) \quad (9.3)$$

In most instances, it will be desirable to check to see whether there is evidence that the logarithms of the index data do not appear to change linearly with time, suggesting a constant rate of change in the population is not tenable. An easy way to do this is to fit a curve to the data, and test for a significant change in the variability around the fitted line. The simplest such curve is the "quadratic" or second degree polynomial. We now fit the function:

$$Y = a + bt + bt^2 \quad (9.4)$$

and calculate a variance as in eq.(3):

$$MS_{quad} = \{\sum [y_t - (a + b_1t + b_2t^2)]^2\}/(n - 3) \quad (9.5)$$

In this instance, we use $n-3$ "degrees of freedom", because another constant (b_2) has been estimated from the data. Actual fitting of eq. (9.4) is readily done by multiple regression. Many computer programs are available to do the fitting, and most of the spreadsheet programs will provide such fits. An F-test of significance can be calculated from the F-ratio:

$$F = \frac{SSLIN - SSQUAD}{MS_{quad}} \quad (9.6)$$

where $SSLIN = (n-2)MS_{reg}$ and $SSQUAD = (n-3)MS_{quad}$, and significance of the F-ratio is checked in tables of the F-distribution with 1 and $n-3$ degrees of freedom. These procedures are described in more detail in Chapter 1.

The test of eq.(9.6) can be applied with other alternative curves. For example, a third degree polynomial can also be fitted to the data by multiple regression. The equation is:

$$y_t = a + b_1t + b_2t^2 + b_3t^3 \quad (9.7)$$

Because this is a more flexible curve than eq.(9.4) it will often appear to give a better fit. However, once a departure from linearity is established, one really needs more information than can be derived from the curve alone in order to assess the situation. Estimates from eq.(9.3) for a number of species appear in Fig. 9.4. Cases where eq.(9.6) gave statistically significant results are denoted by an asterisk (0.05 level of significance) or two asterisks (0.01 level of significance). A very high regression mean (0.250) square for elephant seals results from the nonlinearity evident in Fig. 9.2, and that value is not shown in Fig. 9.4 because of the clear evidence that a different curve is appropriate.

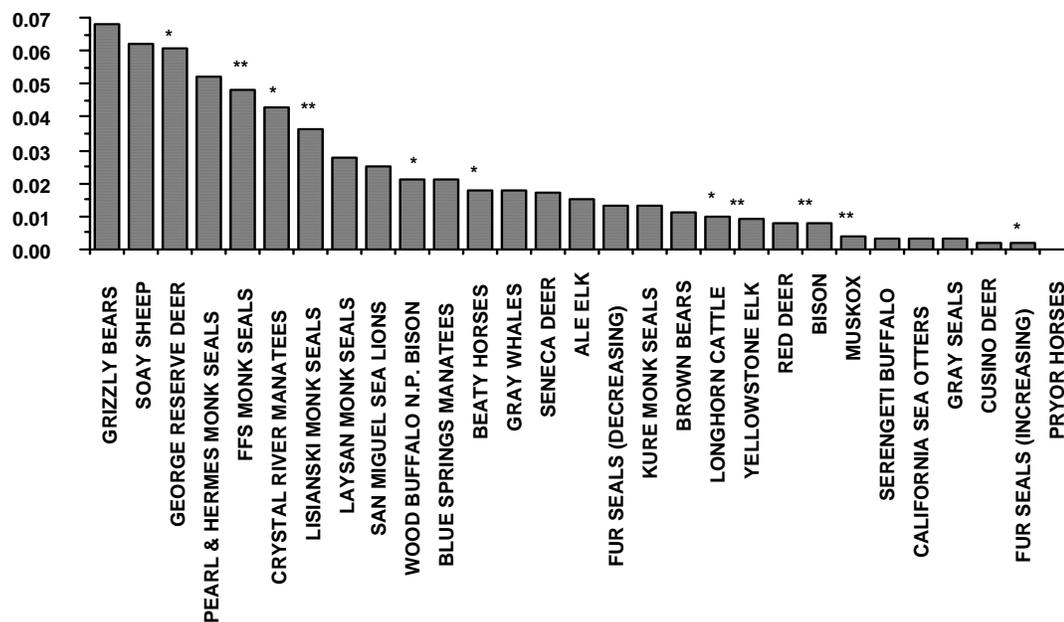


Fig. 9.4. Regression mean square values computed from eq.(9.3) for various species. A single asterisk indicates significance at the 0.05 level for the test of eq. (9.6) and a double asterisk indicates significance at the 0.01 level. Sources and scientific names appear in Table 9.1.

9.4 Assessment of trend data

Fig. 9.4 shows a wide range of regression mean squares, and indicates that the evidence of curvilinearity does not seem to be associated with variability. Some of the likely sources of differences among species can be identified as follows. The mean square for grizzly bears has been assumed to result from the effect of environmental conditions on counts. In wet years, bears can find sufficient food without spending much time in the open and thus are very difficult to find. In especially dry years, they forage widely, and counts tend to be higher in those years. The data used for Fig. 9.4 come from the study reported by Knight et al. (1995). Further analysis of the data yields an improved index and thus a smaller mean square. The variability in the Soay sheep regression evidently results from environmental conditions. Boyd (1974) described the pattern of a build-up for several years, followed by heavy mortality under stress from weather and nutritional conditions. Only 7 years of data are available for the George Reserve deer data set, so it is mainly of interest for the overall observed high rate of increase, and the curvature does not seem to have a recognizable pattern.

Table 9.1 Sources and nature of data and scientific names for data used in Fig. 9.4

NAME	NATURE OF COUNT	SCIENTIFIC NAME	SOURCE
ELEPHANT SEALS	Births	<i>Mirounga angustirostris</i>	Stewart et al. 1994
GRIZZLY BEARS	Females with cubs	<i>Ursus horribilis arctos</i>	Knight et al. 1995
SOAY SHEEP	Total count	<i>Ovis sp.</i>	Boyd 1974
GEORGE RESERVE DEER	Total	<i>Odocoileus virginianus</i>	McCullough 1983
MONK SEALS	Beach counts	<i>Monachus schauinslandi</i>	Gilmartin and Eberhardt 1995
MANATEES	Totals	<i>Trichechus manatus</i>	Eberhardt and O'Shea 1995
SEA LIONS	Pup counts	<i>Zalophus californicus</i>	DeMaster et al. 1982
WOOD BUFFALO N.P. BISON	Total population estimate	<i>Bison bison</i>	Carbyn et al. 1993
FERAL HORSES	Total counts	<i>Equus caballus</i>	Eberhardt et al. 1982
GRAY WHALES	Population estimates	<i>Eschrichtius robustus</i>	Breiwick 1994
SENECA DEER	Reconstructed pop.	<i>Odocoileus virginianus</i>	Hesselton et al. 1965
ALE ELK	Total counts	<i>Cervus elaphus</i>	Eberhardt et al. 1996
FUR SEALS (DEC.)	Total counts	<i>Callorhinus ursinus</i>	
BROWN BEARS	Spawning stream counts	<i>Ursus horribilis</i>	R. A. Sellers, pers. comm.
LONGHORN CATTLE	Total counts	<i>Bos bos</i>	Fredin 1984
YELLOWSTONE ELK	Aerial counts	<i>Cervus elaphus</i>	Houston 1982
RED DEER	Total counts	<i>Cervus elaphus</i>	Clutton-Brock et al. 1982
BISON	Total counts	<i>Bison bison</i>	Fredin 1984
MUSKOX	Total counts	<i>Ovibos moschatus</i>	Spencer and Lensink 1970
SERENGETI BUFFALO	Population estimates	<i>Syncerus caffer</i>	Sinclair 1977
CALIF. SEA OTTERS	Total counts	<i>Enhydra lutris</i>	
GRAY SEALS	Births	<i>Halichoerus grypus</i>	Bonner 1975
CUSINO DEER	Total	<i>Odocoileus virginianus</i>	Ozaga and Verme 1982
FUR SEALS (Inc.)	Total counts	<i>Callorhinus ursinus</i>	Kenyon et al. 1954
PRYOR HORSES	Total counts	<i>Equus caballus</i>	Garrott and Taylor 1990

The monk seal counts at Pearl and Hermes Reef are highly variable because the seals occupy a number of small islets spread over a sizable area there and thus are difficult to reach and to count. Curvilinearity in the French Frigate Shoals data likely result from incomplete tallies in the early years. Evidence of curvilinearity in the Lisianski monk seal data (Fig. 9.5) seems to be characteristic of many of the declining populations, including bison in Wood Buffalo National Park, and the decreasing fur seal population. In many such populations, the causes of decline are unknown or imperfectly understood, and likely vary over time, whereas an increasing population often is doing so in consequence of ample food and other resources, and thus is likely to exhibit a smooth pattern.

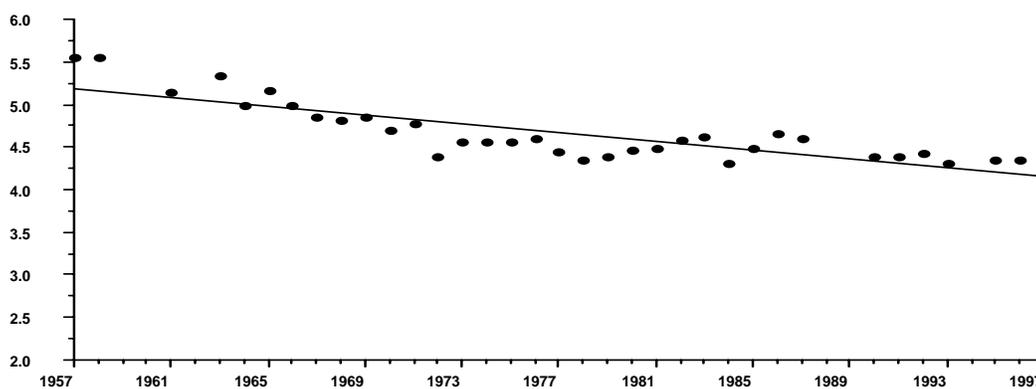


Fig. 9.5. Logarithms of beach counts of monk seals on Lisianski Island.

Somewhat erratic growth patterns may, however, also be evident in increasing populations, as is the case with manatees (Fig. 9.6). Very likely some of the fluctuations result from conditions under which the counts were made, inasmuch as the counts may be made by divers and at times in turbid water (Crystal River).

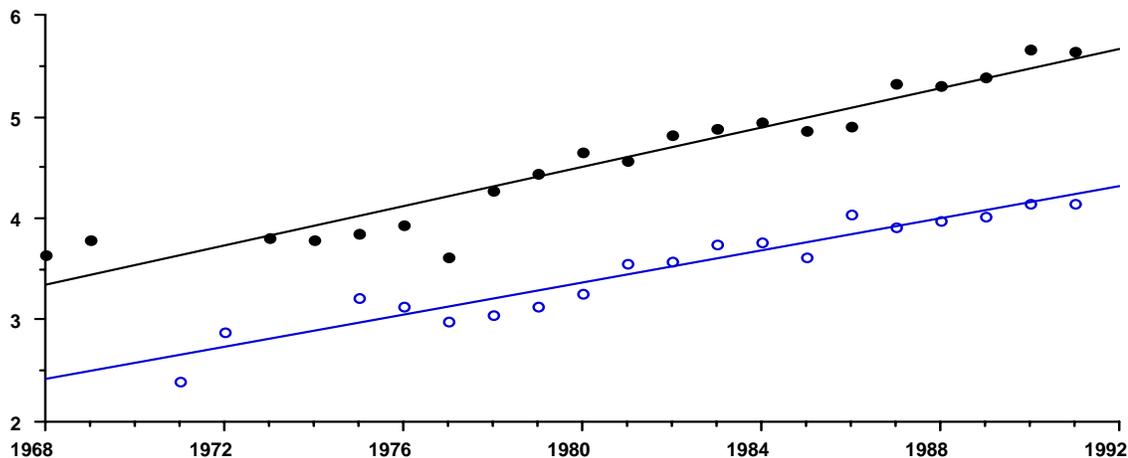


Fig. 9.6. Logarithms of counts of manatees at Crystal River (upper data set) and Blue Springs (lower data points).

It seems quite possible that data on the right side of Fig. 9.4 may approach the circumstance where the fluctuations may largely be associated with the stochastic behavior of the underlying birth and death processes. In several instances (muskox and bison, fur seals (increasing), and Seneca deer), there is a suggestion of an oscillatory pattern in the deviations from linearity. Such oscillations may be a consequence of a changing age structure (Keyfitz 1968 gave the theoretical basis) and are suggested by simulations of longer sequences of observations (Eberhardt 1981). Detection of such an oscillatory pattern is aided by plotting deviations from fitted curves as in Fig. 9.7, which shows the pattern for the muskox data of Fig. 9.3.

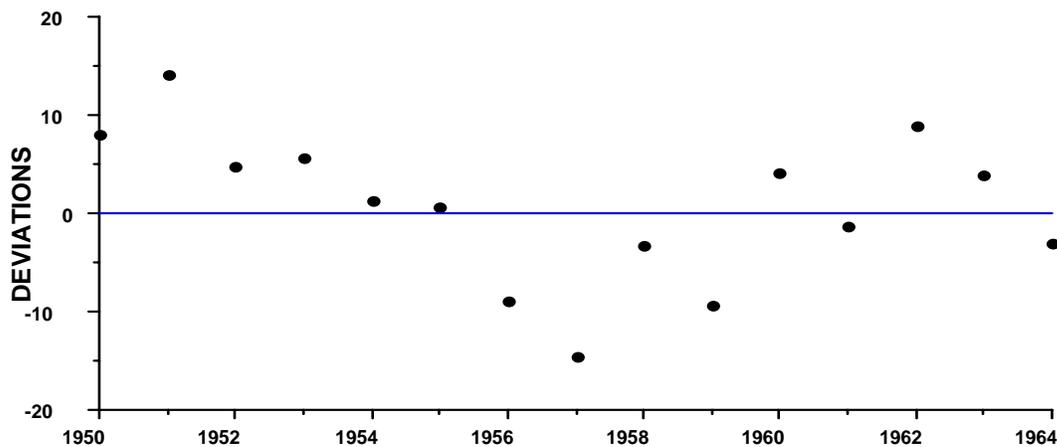


Fig. 9.7. Deviations from the exponential curve fitted to muskox data in Fig. 9.3.

9.5 A test for significant deviations from regression using replicate points.

The test for significant deviations from linearity used in Section 9.3 depends on fitting a curve and testing to see whether the improvement in fit might simply be due to chance. In some cases, replicate counts may be available, so that one can use the variability within years to test significance of deviations from linearity. The test was introduced in Section 1.6 using data on monk seals. The advantage here is that we do not need to specify an alternative model like the quadratic or cubic (which may very well be the wrong model). Some counts of brown bears at spawning streams provide another example (Fig. 9.8). In this case, the test consists of making the usual analysis of variance to test for significance of the linear regression (Table 9.2), and then using the pooled variance of individual observations within years to estimate "pure error" (Draper and Smith, 1998). The data for calculation of pooled error appear in Table 9.3. A sum of squares of deviations from the mean is calculated for the data in each year where there are two or more observations and these values are summed to give an overall sum of squares, which is subtracted from the "residual" sum of squares in Table 9.2 to yield the "lack of fit" sum of squares (i.e., the variability not accounted for by "pure error"). The degrees of freedom used to calculate pure error (32) is similarly subtracted from the degrees of freedom for residual error to get the degrees of freedom used to calculate a mean square for "lack of fit". An F-ratio as shown in Table 9.2 would be used to test significance of the lack of fit, but the F-test for regression (1.14) is not significant, so there is no real point in going on to test lack of fit, other than to illustrate the method.

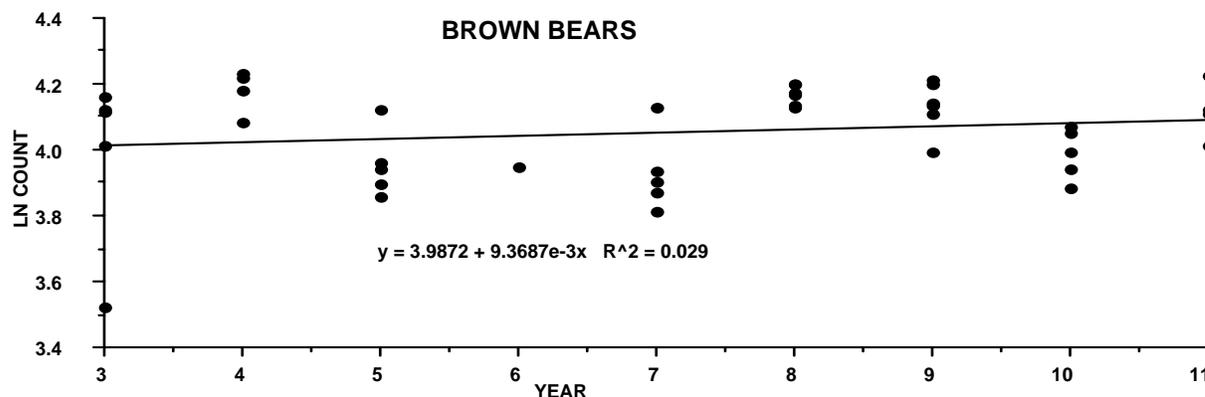


Fig. 9.8 Logarithms of counts of brown bears on salmon spawning streams.

Table 9.2 Test of significance for deviations from regression

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	1	0.02415	0.02415	1.14655
Residual	39	0.82140	0.02106	
Total	40	0.84555		
Lack of fit	32	.4753	0.01485	
Pure error	7	0.3461	0.04944	

Another example of variability of individual counts is provided by some gray whale data (Breiwick 1994). In this case, replicate counts over time could not be made, as the whales are counted as they migrate past a shoreline counting station. However, estimates of sampling error are available (Table 9.4) and can be used to assess the mean square error obtained in log-linear regression analysis of the data. Because the standard errors of individual estimates were calculated in terms of number of whales, a transformation needs to be used to change to variability on the logarithmic scale. Using the delta method (Section 3.9), we find:

$$V(\ln x) = V(x)\left(\frac{1}{x}\right)^2 = \text{Coef. var.}^2(x)$$

Plotting the coefficients of variation against the estimates shows little evidence of correlation, so the average coefficient may be a reasonable estimate of overall variability. The squared value (0.0027) is substantially smaller than the mean square about regression (0.0170) so it seems quite evident that there is a significant departure from linearity, even though the quadratic and cubic mean squares are little different from the mean square from linear regression (Table 9.4). Inspection of the data (Fig. 9.9) suggests the possibility that the population might have reached an equilibrium level. More recent counts reinforce this prospect.

Table 9.3 Data for computation of "pure error" for brown bear counts

Year	Bears/hour	Logarithm bears/hr	Sum of squares	Deg. of freedom
3	33.85	3.5219		
3	64.04	4.1595		
3	61.88	4.1252		
3	61.2	4.1141		
3	55.24	4.0117	0.2819	4
4	68.7	4.2297		
4	59.3	4.0826		
4	67.9	4.2180		
4	65.3	4.1790	0.0134	3
5	49.4	3.9000		
5	51.4	3.9396		
5	61.6	4.1207		
5	47.4	3.8586		
5	52.45	3.9599	0.0400	4
6	51.88	3.9489		
7	45.14	3.8098		
7	62	4.1271		
7	48.13	3.8739		
7	49.58	3.9036		
7	51.21	3.9359	0.0572	4
8	62.06	4.1281		
8	66.59	4.1986		
8	62.32	4.1323		
8	66.88	4.2029		
8	65.03	4.1748		
8	64.58	4.1679	0.0051	5

9	54.17	3.9921		
9	67.49	4.2120		
9	66.67	4.1998		
9	62.8	4.1400		
9	61	4.1109		
9	62.42	4.1339	0.0311	5
10	48.68	3.8853		
10	51.47	3.9410		
10	58.51	4.0692		
10	57.65	4.0544		
10	54.08	3.9905	0.0238	4
11	61.12	4.1128		
11	55.15	4.0101		
11	68.29	4.2238		
11	61.52	4.1194	0.0229	3
Sums	2386.0800	166.2194	0.4753	32

Table 9.4. Estimate of total numbers of gray whales along with standard errors of estimate and coefficients of variation.

YR	ESTIMATE	S.E.	COEF. VAR.	LN ESTIMATE
1	12921	964	0.075	9.4666
2	12070	594	0.049	9.3985
3	12597	640	0.051	9.4412
4	10707	487	0.045	9.2787
5	9760	524	0.054	9.1860
6	15099	688	0.046	9.6224
7	14696	731	0.050	9.5953
8	12955	659	0.051	9.4692
9	14520	796	0.055	9.5833
10	15304	669	0.044	9.6359
11	16879	1095	0.065	9.7338
12	13104	629	0.048	9.4807
13	16364	832	0.051	9.7028
18	21443	1182	0.055	9.9732
19	20113	927	0.046	9.9091
21	20869	913	0.044	9.9460
26	17674	1029	0.058	9.7798
27	23109	1262	0.055	10.0480

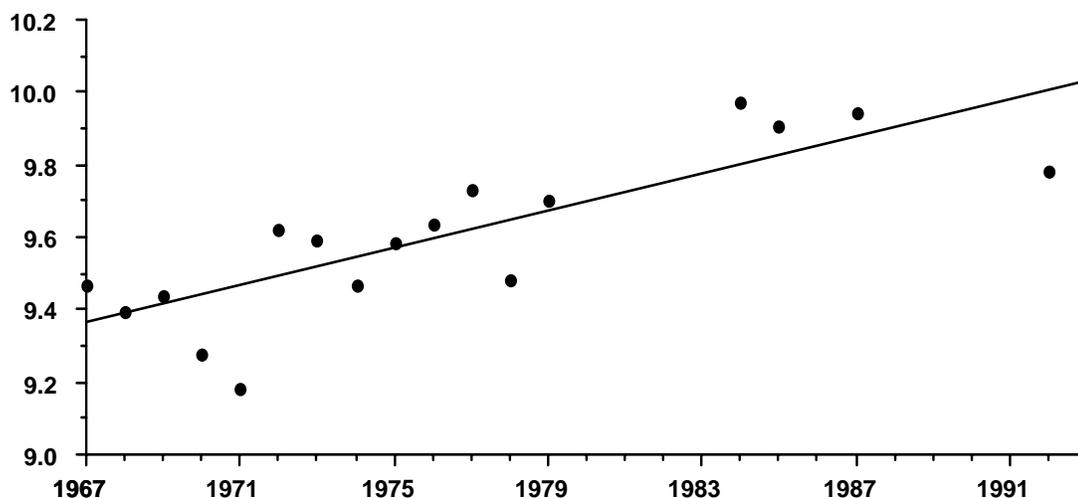


Fig. 9.9 Logarithms of counts of gray whales. Data from Breiwick (1994).

9.6 Studying linearity with the lowess method

The test for deviations from linearity based on fitting a quadratic curve to the data uses the difference between sum of squared deviations about the loglinear regression line and that about the quadratic curve, and tests this difference against the regression mean square about the quadratic, using an F-test. Such a test can also be used with the cubic equation (or any suitable alternative), with the only change being that of reducing the degrees of freedom to adjust for the extra parameters fitted. An alternative is to consider curves fit with a "locally weighted regression" technique introduced in Section 1.12. Weighted linear regressions are fit at each point on the graph (e.g., if the data span 30 years, then such regressions are fit at each of the 30 years) by selecting data points in the immediate neighborhood of each point on the x-abcissa. The number of points in each such neighborhood is usually taken to be about 30% of the total number of observations. Weights diminish by a cubic function, so points very near to the selected point get by far the most weight. The individual fitted regression line determines only the y-value for the selected abscissal value. In effect, the technique behaves much like a moving average, but has various advantages.

In principle, one might extend the test of linearity to the lowess technique, but the fitting procedure is such that there is no way to calculate an effective number of degrees of freedom. However, the method can be bootstrapped (Efron and Tibishirani 1993, Chapter 7) and may thus be used to study more complex departures from linearity.

9.7 Confidence limits from bootstrapping

The usual linear regression model assumes a normal distribution of deviations about the regression line as a basis for calculating confidence limits and tests of significance. The logarithmic transformation does seem to result in a symmetric distribution of deviations from regression, but it may be worthwhile to use bootstrapping to obtain confidence limits that do not depend on the assumption of normality.

Efron and Tibishirani (1993) describe two approaches to bootstrapping in regression. One depends on fitting the usual regression model to calculate individual deviations from regression, and then taking random samples of these deviations with replacement (same sample size as the original data) and adding this set of deviations to the y-values predicted by the original regression equation to generate a new set of "bootstrap data". A regression line is fit to this data and the parameters recorded. Doing this, say, 1,000 times provides the basis for calculating confidence limits on the parameters. For 95% limits we use the 25th and 975th ordered value of a parameter. The original set of x-values is used throughout. This approach thus apparently makes it feasible to work with small data sets. A disadvantage is that the method assumes that the linear model is exactly true. Also, monte carlo simulations of the type described in Section 3.8 indicate that "coverage" of the method based on bootstrapping deviations is appreciably less than the expected 95%, being approximately 85-90% for simulations using normal errors and $n = 10$. Until more is known about bootstrapping in regression with small samples, the best course appears to be to use ordinary regression for confidence limits on samples below $n = 10$.

An alternative "bootstrapping pairs" takes n random samples with replacement of the x,y pairs and computes new regressions. The disadvantage is that, in small data sets, the sample observations may pile up on a few x -values, yielding strange results. The advantage is that the method does not assume the linear model holds exactly. The two methods were tested on a number of data sets from increasing and decreasing populations, and compared with confidence limits obtained from the usual regression equation. The approach using deviations gave confidence intervals (Fig. 9.10) that were much narrower than the usual 95% regression limits on slope of the loglinear regression line. Most of the samples had $n \geq 10$, but two feral horse populations had only 5 and 8 observations, respectively (Table 9.5).

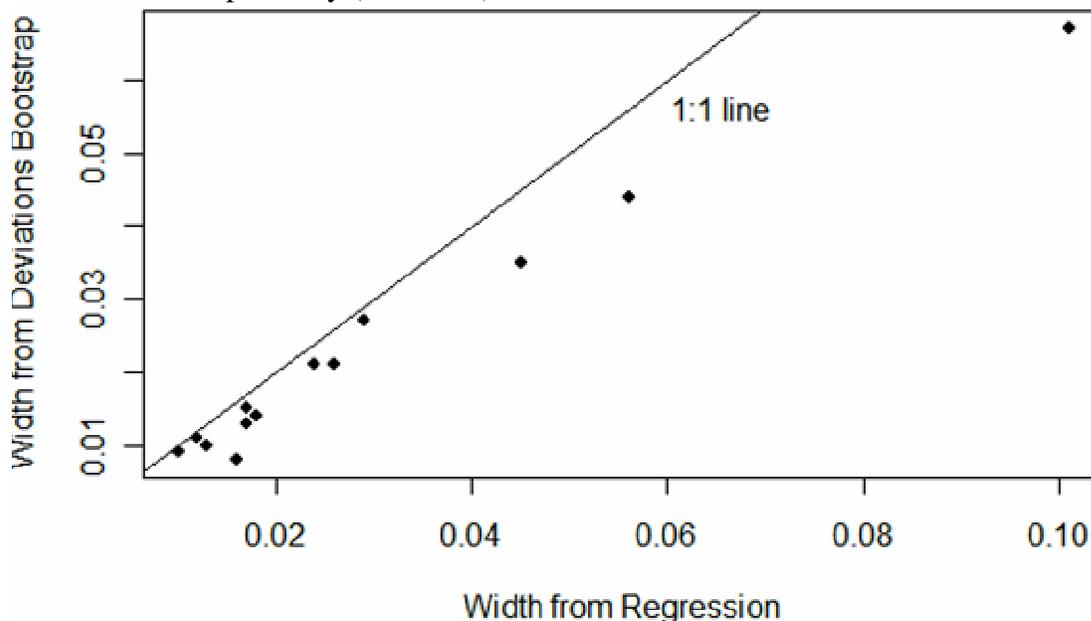


Fig. 9.10 Comparison of 95% confidence intervals on the regression slope (b) calculated by ordinary regression methods to those obtained from bootstrapping deviations.

Loglinear regressions of population growth data are not likely to conform to the assumption of normal distribution of deviations needed for setting confidence intervals by the usual regression approach. Consequently, when sample sizes are 10 or larger one should consider bootstrapping pairs for confidence intervals. Thus far we have mainly used the simple “percentile” intervals. Efron and Tibishirani (1993) recommend a better method, the “BCa” method of Section 3.6. The method was tested on 14 data sets, with smallish samples ($n = 11$ to $n = 35$), and gave confidence interval widths that were essentially identical to those produced by the percentile method. Using the improved BCa method as standard, it turns out that the regression intervals are appreciably narrower than those from bootstrapping pairs (Fig. 9.11).

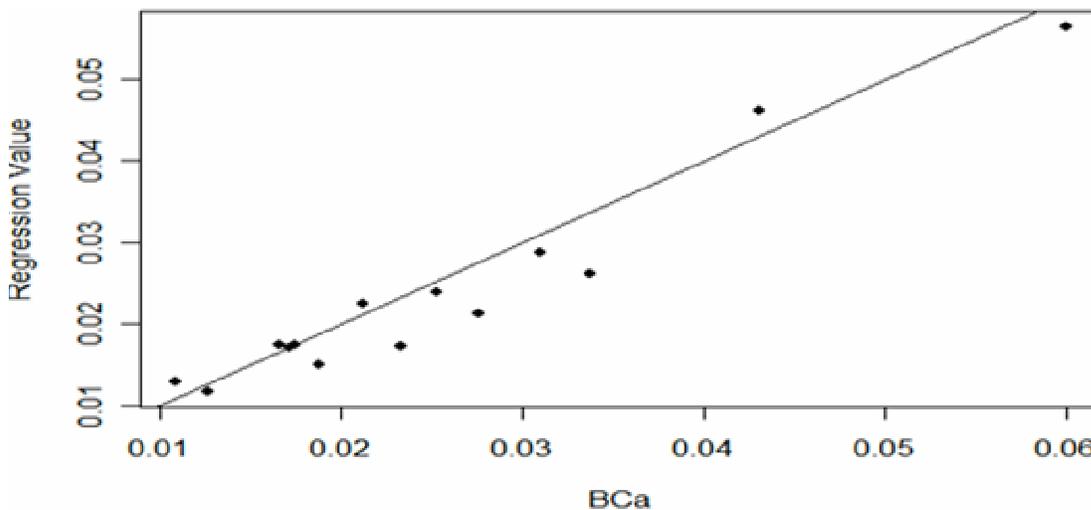


Fig. 9.11 Comparison of 95% confidence interval widths on the regression slope (b) calculated by ordinary regression methods to those obtained from bootstrapping pairs with the BCa corrections of Section 3.6.

When sample sizes are, say, 10-15 or larger, a worthwhile approach is to compute confidence limits both by bootstrapping pairs, and by the usual approach (readily available in EXCEL). If they agree, then there should be little reason for concern. Both of the bootstrap methods support the notion that regression estimates of population rate of change are unbiased. Efron and Tibishirani (1993) indicate that close agreement between the mean of bootstrapped data and the value obtained from the original approach indicates an unbiased estimate. Table 9.5 shows that the 3 methods give essentially the same values for rate of increase.

Table 9.5 Slope estimates from log-linear regression compared to those from bootstrap samples using deviations and those using x,y pairs. The close agreement indicates unbiased estimation (Efron and Tibishirani 1993).

SPECIES	REGR SLOPE	DEV MEAN	PAIRS MEAN
OTTERS	0.0521	0.0523	0.0524
GRAY WHALES	0.0257	0.0259	0.0259
CRYSTAL R	0.0968	0.0971	0.0975
BLUE SPGS	0.0793	0.0791	0.0793
SENECA DEER	0.4115	0.4151	0.4138
MUSKOX	0.1463	0.1461	0.1465
GRAY SEALS	0.0741	0.0742	0.0743
BISON	0.2068	0.2068	0.2064
FUR SEALS	0.0824	0.0825	0.0833
BEATY HORSES	0.2447	0.2438	0.2438
LISIANSKI	-0.0260	-0.0260	-0.0257
PRYOR HORSES	0.1854	0.1855	0.1854
BROWN BEARS	0.0167	0.0165	0.0159
FUR SEAL DEC	-0.0263	-0.0264	-0.0261

9.8 Alternative estimates of rate of population change

In Chapter 12 two very simple models (Section 12.4) are proposed for evaluating populations from which known numbers of individuals are removed at various times. The log-linear approach of the previous sections (eq. 9.2) is not useful in such circumstances, so it is worthwhile to consider alternative ways to establish rates of change. These make use of ratios of successive observations. We first consider using the methods for data without removals, and thus can compare them with the regression approach. The basic idea comes from the simple relationship (eq. 9.1):

$$N_t = \lambda N_{t-1}$$

where λ is the multiplier needed to project a population at time $t-1$ to time t (we assume an interval of one year, with observations taken at the same time each year). If we want to estimate λ from a sequence of years, we can consider 3 ratio estimates (Eberhardt 1987) where x_i represents population size in one year and y_i the size in the next year:

The mean of individual ratios $\hat{\lambda} = \frac{\sum(y_i/x_i)}{n}$ (9.7)

Ratio of sums $\hat{\lambda} = \frac{\sum y_i}{\sum x_i}$ (9.8)

Regression through the origin $\hat{\lambda} = \frac{\sum y_i x_i}{\sum x_i^2}$ (9.9)

Because the individual observations other than the first and last appear twice in a sequence of years, one can get spurious results in statistical analysis of such data (Eberhardt 1970). We thus resort to bootstrapping here. Using the data previously used to evaluate the regression approach (Table 9.5), 1,000 bootstraps were used with the estimates of eqs. (9.7) through (9.9). Bootstrap bias calculations were made, along with

comparisons with the rate of change from the regression method (using $\lambda = e^b$, where b is slope of the loglinear regression). All 4 sets of estimates were highly correlated (Table 9.6).

Table 9.6. Correlations between four methods of estimating rate of increase.

	LOGLINEAR	MEAN	SUMS	REGR
LOGLINEAR	1			
MEAN	0.992	1		
SUMS	0.983	0.986	1	
REGR	0.957	0.961	0.990	1

The mean of the individual ratios had the smallest relative bias, and deviated the least from the loglinear regression estimates. Relative bias was calculated as:

$$\text{Bias} = \frac{\lambda_{\text{orig}} - \lambda_{\text{boot}}}{\lambda_{\text{orig}}} \quad (9.10)$$

where λ_{boot} is the mean of 1,000 bootstrap estimates and λ_{orig} is the estimate from log-linear regression. The average of absolute values for relative bias was 0.020 for the method of means [eq.(9.7)], 0.107 for method of sums [eq.(9.8)], and 0.130 for the regression method [eq.(9.9)].

Confidence intervals (95%) were also calculated from bootstrap results, but were much wider than the regression estimates and poorly correlated with those estimates (Fig. 9.12). Excluding the Beaty horse data ($n = 8$), the ratio of confidence intervals (Fig. 9.12) was about 3.5 to 1, i.e., confidence intervals from the mean of individual ratios [eq.(9.7)] were nearly 4 times wider than those from log-linear regressions. This raises an issue that needs further attention. Why should the bootstrapping confidence intervals be so much wider than those from linear regression? The answer appears to lie in the fact that bootstrapping is based on random samples with replacement of *individual* observations. In this case, we have to deal with successive pairs of observations in order to get the ratios, so the pairs were bootstrapped, and this gives unsatisfactory results, as shown in Fig. 9.12.

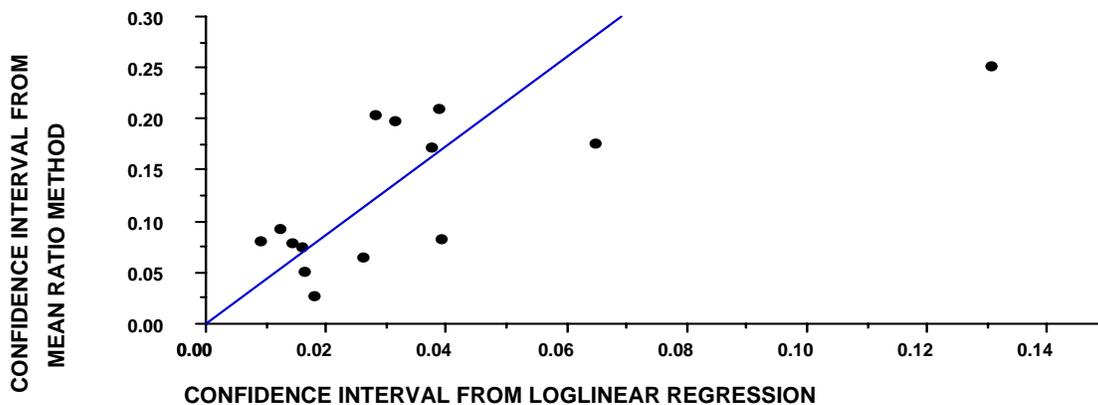


Fig. 9.12. Relationship between confidence interval widths (95%) calculated by log-linear regression and those from a ratio method [Eq. (9.7)].

Example 9.5 A census using ratio estimation

An example of ratio estimation in an actual census is provided by studies of the abundance of the sea otter (*Enhydra lutris*) along the California coast. Sea otters then occupied about 150 miles of the California coastline, and are almost always found swimming or resting (in kelp beds, usually) just off the shoreline. A number of aerial censuses were conducted by employees of the California Department of Fish and Game. Since only a fraction of the otters present are observed in aerial counts, an ingenious use of auxiliary counts (devised by D.J. Miller) was used to correct the aerial counts by what amounts to ratio estimation. On a sample of shoreline areas, observers on vantage points (well above sea level, where possible) made counts on well-defined areas. These were then used to correct aerial counts of the same areas.

In later development of the method, aerial photos were made in advance of the census to map kelp beds. Using the maps, shoreline ("ground truth") counters located animals in a well-defined and readily visible area, and made a record of those animals present in the sample areas at the time the aerial observers passed the counting site. The aerial observers also plotted all otters seen on identical maps. As soon as a day's flight was completed, the aerial and ground observers went over the maps together to establish which animals were seen by both air and ground observers, and which animals were not seen from the air. Aerial counts were made over the entire coastline occupied by otters, so the data can be used in a ratio estimate (Sec. 4.12).

The estimated total otter population is:

$$\hat{Y}_R = \frac{\sum y_i}{\sum x_i} X_T$$

where Y_R is the estimate of total otter population, and X_T is the total aerial count. The ground counts (y_i) are summed over the sample (n) of ground-truth areas, and divided by the sum of the aerial counts (x_i) on the same areas. If it is assumed that the ground counts are made without error (an assumption that needs further checking), then the data conform to the standard conditions for ratio estimation with the exception that X_T (the total aerial count) is subject to sampling error. Calculations are otherwise straightforward, and as given in Ch.4. Table 9.5 gives data from the 1974 survey. A total of 897 otters were counted from the air, so we estimate the total population from the equation given above as:

$$\hat{Y}_R = \frac{332}{202} 897 = 1474$$

and find the squared coefficient of variation to be:

$$CV^2(\hat{Y}_R) = \frac{(1-f)}{n} [c_{yy} + c_{xx} - 2c_{yx}] = \frac{1}{31} [1.966 + 1.216 - 2(1.4229)] = 0.0109$$

where we take $f = 0$, since an accurate value is not available. However, an appreciable segment of the coast was included in the ground counts, so that use of f would be appropriate here, if it could be calculated.

Confidence limits are readily calculated by computing a standard error as $(.0109)^{1/2} 1474 = 159.89$, and using $Z_{05} = 1.96$ giving:

$$\hat{Y}_R = \pm Z_{05} s(Y_T) = 1474 \pm 1.96(159.89) \text{ or,} \\ 1172 < Y_R < 1776 \text{ sea otters.}$$

Aerial survey (x_i) and ground counts (y_i) of sea otters along the California coast.

	<u>June 25</u>		<u>June 26</u>		<u>June 27</u>	
	\underline{x}_i	\underline{y}_i	\underline{x}_i	\underline{y}_i	\underline{x}_i	\underline{y}_i
	2	4	1	6	1	1
	1	1	9	10	6	14
	0	4	0	19	5	7
	0	6	13	19	1	6
	37	50	12	8	4	5
	6	9	5	8	2	8
	9	10	6	6	2	1
	3	3	5	9	37	47
	1	1	0	5		
	8	13	1	1		
	0	8	11	14		
			14	29		
	<hr/>					
	67	109	77	134	58	89

9.9 Criteria for regression fits

Testing for curvilinearity in regression lines was discussed in some earlier sections (Sec. 9.3, Sec. 9.5). When one fits a regression line like those summarized in Fig. 9.4, it is important to have other ways to evaluate the fit. Regression mean squares provide a useful measure of variability about the line, and can be supplemented by the widely-used R^2 criterion. For a simple regression line with one independent variable, the square root of this quantity gives the well-known correlation coefficient, and R^2 is perhaps better known when used in the multiple regression analyses that will be described in a later section. However, the simple linear regressions serve to illustrate the interplay between regression mean square and R^2 . The R^2 criterion is:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (9.11)$$

where y_i is the independent variable, \bar{y} is the mean of the y_i and \hat{y} is the "predicted" value from the regression equation. Often, R^2 is described as the proportion of the variance in the independent variable "accounted for" by the fitted regression line. The two expressions in eq.(9.11) are connected by the basic identity in the analysis of variance in regression, i.e.,

$$\sum(y_i - \bar{y})^2 = \sum(\hat{y}_i - \bar{y})^2 + \sum(y_i - \hat{y}_i)^2$$

Total S.S. = regression S.S. + Residual S.S.

In eq.(9.11), the left-hand form is that commonly used (Draper and Smith 1998:138). The equivalent right-hand expression was recommended by Anderson-Sprecher (1994) because it provides a convenient interpretation of R^2 written as:

$$R^2 = 1 - \frac{\text{RSS}(\text{full})}{\text{RSS}(\text{reduced})}$$

where $\text{RSS}(\text{full})$ denotes the regression sum of squares for the full model and $\text{RSS}(\text{reduced})$ can be interpreted as the sum of squares for the model reduced to its minimal form, i.e., the expected value of y is β_0 , estimated by \bar{y} . One advantage of this expression is that it emphasizes that model comparisons using R^2 should be made with nested models, i.e., a series of regression equations with two or more independent variables so that the number of parameters $p = 2, 3, 4, \dots$. Another advantage is that

$$1 - R^2 = \frac{\text{RSS}(\text{full})}{\text{RSS}(\text{reduced})}$$

states the fraction of variability not accounted for by regression. Because at least stochastic fluctuations are always present in trend data, this expression serves as a reminder that R^2 cannot become unity.

When R^2 is large, it is evident that the regression line does a good job of predicting the counts. This does not necessarily demonstrate validity of the index as some extraneous factor may be exerting a major influence on the counts. Usually, however, a high R^2 is reassuring. A key element in demonstrating validity of an index is an independent estimate of the trend. Such an estimate may be available from reproductive and survival data. Thus if two sources, trend index and reproductive and survival data produce much the same estimate of λ , that result is particularly reassuring.

Confidence intervals on parameters and on predictions from regression lines depend on the assumption of a normal distribution of deviations from the regression line with constant variance about regression. When population data are being considered, this may be an uncertain assumption. It is thus desirable to resort to the bootstrapping technique as a check on confidence intervals generated by regression theory, as discussed in Section 9.7. If bootstrapping is used, it furnishes an estimate of bias for the estimator of concern from the equation (Section 3.2):

$$\hat{\text{bias}}_B = \hat{\theta}^*(\bullet) - t(\hat{F}) \quad (9.12)$$

Here, $\hat{\theta}^*(\bullet)$ is the mean of the bootstrap estimates of the parameter of interest (often λ) and $t(\hat{F})$ denotes the same parameter estimated from the original data. Results from this criterion may thus yield an indication of problems with the trend index.

How should R^2 and the regression mean square be used? An interpretation of R^2 was given earlier as:

$$R^2 = 1 - \frac{\text{RSS}(\text{full})}{\text{RSS}(\text{reduced})}$$

Here RSS denotes a regression sum of squares. The numerator pertains to the fitted model, while the denominator consists of the variance about the mean of the observations

(thus "reduced" to a minimum). Usually the sample size is large enough so that the ratio amounts to comparing the variance of the fitted model to that in the data (i.e., the degrees of freedom are not different enough to matter much). Hence it is clear that R^2 and the regression mean square are closely related. The relationship can be examined by plotting (Fig. 9.13) R^2 and regression mean squares for the data of Fig. 9.4.

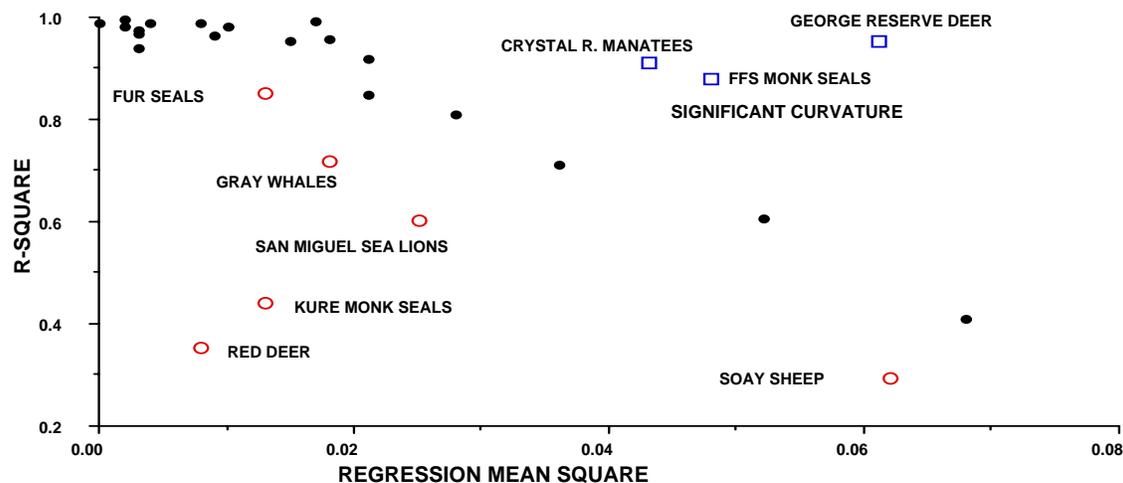


Fig. 9.13. Relationship between R^2 and regression mean square for data used in Fig. 9.4. Various sets of points described in text.

The solid points show what seems likely to be the expected relationship between R^2 and the regression mean square. The cluster at the right top contains those data sets where there is a significant curvature which inflates the regression mean square. The points represented by open circles represent cases where there appears to be a pattern in the deviations which is quite dramatic in some instances. Plotting deviations from regression, as recommended earlier (see Fig. 9.7 and Fig. 9.14), will usually make the non-randomness of the deviations evident. All of the cases shown here are based on simple log-linear regressions. Very likely a detailed investigation of the underlying circumstances might turn up significant auxiliary variables. Considering the solid points at the left of the figure, it appears that most of the variability is accounted for by the simple regression. R^2 thus appears most informative in this situation.

Evaluating the pattern in deviations from regression can be aided by the Durbin-Watson test (Draper and Smith 1998). This simple test depends on the fact that the squared difference between successive deviations will approximate the variance of the deviations if the pattern of deviations is random. The test is:

$$d = \frac{\sum (e_u - e_{u-1})^2}{\sum e_u^2} \quad (9.13)$$

where the summation in the numerator runs from $u = 2$ to n , and that in the denominator runs from 1 to n . It can be shown that the ratio, d , has an expected value of 2 under a random pattern. Draper and Smith (1998) note that $0 \leq d \leq 4$, and give tables of significant deviations from the expected value of 2 for different sample sizes. Various

statistical packages produce the Durbin-Watson test on residuals, but the test is easy to compute and thus worth calculating directly once one has the residuals from regression (readily available with the spreadsheet regression calculations). Three of the sets of data in Fig. 9.13 show significance at the 1% level from the Durbin-Watson test. The pattern of deviations for these three species appears in Fig. 9.14, where the correlation of successive observations is quite evident.

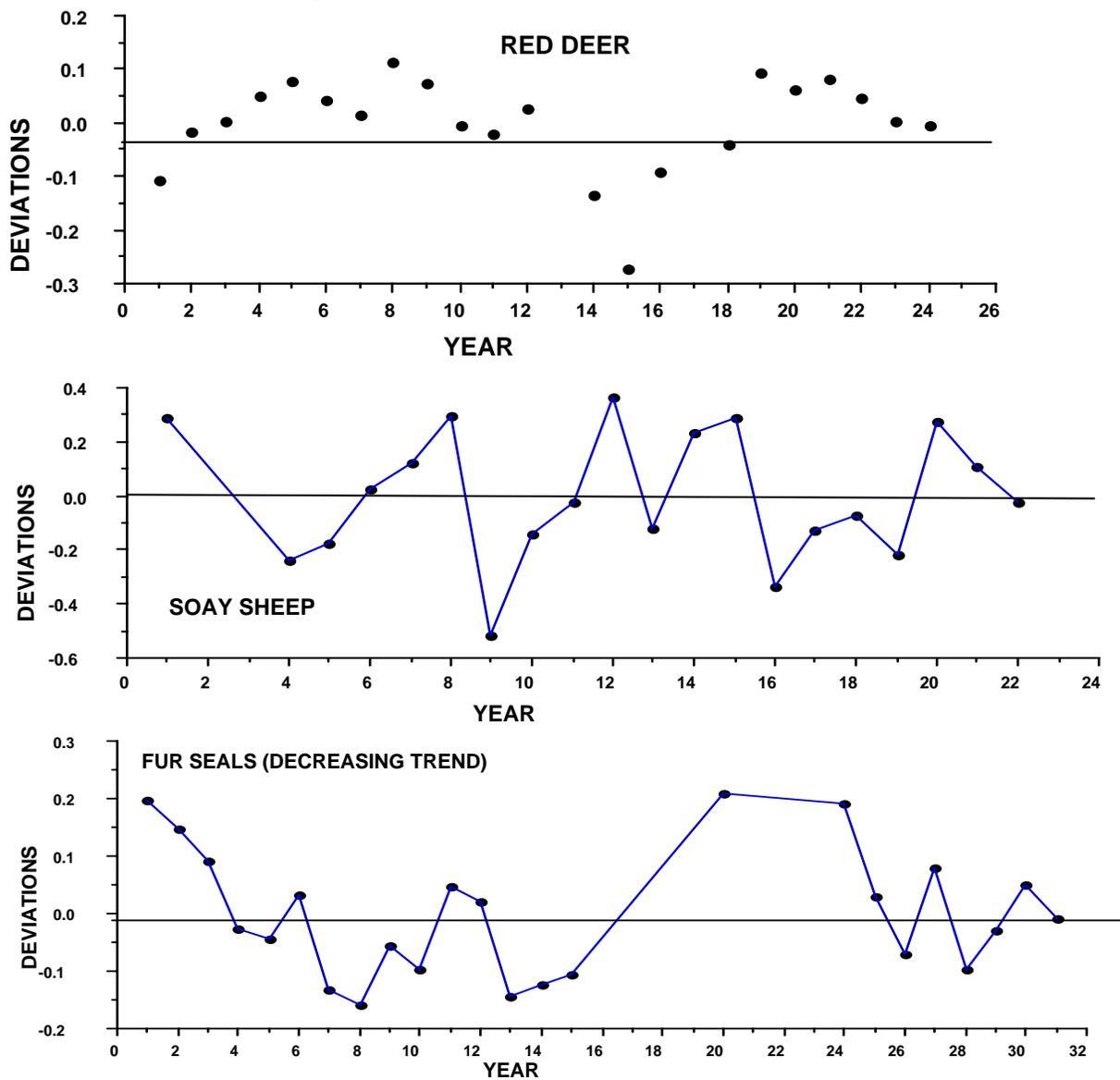


Fig. 9.14 Pattern of deviations from loglinear regression for 3 species. All 3 sets of data are significant at the 1% level with the Durbin-Watson test for serial correlation.

The results above thus suggest several steps in appraising loglinear regressions of trend data:

- 1) Test the data for curvilinearity.
- 2) Compute the regression mean square and R^2 .

3) Examine the pattern of deviations from regression over time, and calculate the Durbin-Watson test.

It may be helpful to compare the results with the data of Fig. 9.13. Most of the data points there conform to the above steps, but not all. Three of the open circles (Kure monk seals, San Miguel sea lions and gray whales) do not fit in, but plotting the residuals suggests that the data are quite erratic, and that there may be some other factors involved that need to be further investigated. Significant curvilinearity (3 points on the upper right of Fig. 9.13, Crystal River manatees, French Frigate monk seals, and George Reserve deer) also indicates a need to look further at the data, inasmuch as the curvilinearity may well indicate a significant change in trend. Table 9.7 gives the data on species shown in Fig. 9.13. The very high R^2 values shown by a number of the data sets suggest that the variability is mainly a function of stochasticity. Bison, musk oxen, Cusino deer and Pryor wild horse populations were essentially counted in their entirety, while the Seneca deer population was reconstructed from removals that were known almost completely.

Table 9.7 Data on loglinear regressions used in Fig. 9.13. Data ordered by regression mean squares. Population sizes are rough estimates in a number of cases.

SPECIES	NUMBERS	OBSNS	SLOPE	Lambda	MSreg	R-sq
GRIZZLY BEARS	400	18	0.039	1.040	0.068	.409
SOAY SHEEP	1000	20	0.026	1.026	0.062	.295
GEORGE RESERVE DEER	70	7	0.488	1.629	0.061	.956
PEARL & HERMES MONK SEALS	150	16	0.050	1.051	0.052	.606
FFS MONK SEALS	600	22	0.062	1.064	0.048	.882
CRYSTAL RIVER MANATEES	150	21	0.097	1.102	0.043	.912
LISIANSKI MONK SEALS	300	35	-0.029	0.971	0.036	.714
LAYSAN MONK SEALS	300	34	-0.034	0.967	0.028	.812
SAN MIGUEL SEA LIONS	8000	15	0.064	1.066	0.025	.603
WOOD BUFFALO N.P. BISON	5000	19	-0.052	0.949	0.021	.849
BLUE SPRINGS MANATEES	40	19	0.079	1.082	0.021	.92
BEATY HORSES	400	8	0.245	1.277	0.018	.959
GRAY WHALES	15000	18	0.026	1.026	0.018	.718
SENECA DEER	400	11	0.411	1.508	0.017	.992
ALE ELK	70	11	0.165	1.179	0.015	.956
FUR SEALS (DECREASING)	180000	24	-0.026	0.974	0.013	.851
KURE MONK SEALS	100	12	0.026	1.027	0.013	.443
LONGHORN CATTLE	20	7	0.302	1.353	0.010	.981
YELLOWSTONE ELK	7000	8	0.191	1.210	0.009	.964
RED DEER	1600	22	-0.010	0.990	0.008	.354
BISON	160	14	0.207	1.230	0.008	.99
MUSKOK	120	15	0.146	1.157	0.004	.991
SERENGETI BUFFALO	50000	11	0.064	1.066	0.003	.969
CALIFORNIA SEA OTTERS	1600	13	0.052	1.054	0.003	.94
GRAY SEALS	1200	16	0.069	1.072	0.003	.975
CUSINO DEER	80	5	0.370	1.448	0.002	.995
FUR SEALS (INCREASING)	130000	12	0.082	1.085	0.002	.984
PRYOR HORSES	120	5	0.185	1.204	0.000	.99

9.10 Using auxiliary variables with trend data

In some cases, the use of auxiliary variables may serve to reduce the regression mean square, i.e., we add independent variables other than time and use a multiple

regression equation. One such model was used to study trends in manatee numbers by Garrott et al. (1994,1995). They represented the expected number of manatees counted at any given site under average conditions at time t as $M(t)$, and assumed $M(t)$ remains constant over the annual counting period (given average conditions) and denoted it as M_i for the i th year. M_i is assumed to be proportional to the true population level. $C[t,X(t)]$ then represents the expected number of manatees counted at time t , given conditions $X(t)$, where $X(t)$ is a vector of covariate values prevailing at time t , leading to the model:

$$C[t,X(t)] = M_i R[t,X(t)] \quad (9.14)$$

where the function $R(\)$ is a rate function that takes the value unity when conditions are average. Under good counting conditions $R(\) > 1$ and under poor conditions $R(\) < 1$. If the rate function is assumed to be of the form

$$R[t,X(t)] = \exp[x(t)'\beta]$$

where $x(t)$ is the vector of covariates and β is a vector of regression parameters, then taking logarithms (base e) gives:

$$\ln C[t,X(t)] = \ln M_i + x(t)'\beta$$

One can then use multiple regression to study the effect of various covariates (auxiliary variables). The general formulation is widely used in survival studies as "Cox's proportional hazards model" (Cox 1972) and has also been used to take auxiliary variables into account in population estimation procedures. The main interest in eq.(9.14) is for studying population trend. In order to do so, M_i needs to be expressed as a function of time, usually as $M_i = N_0\lambda^t$ so that the final equation becomes:

$$\ln C(t) = \ln N_0 + t \ln \lambda + x(t)'\beta \quad (9.15)$$

Compare this with eq.(9.1) and (9.2). The underlying model is thus assumed to have the form:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots + \beta_{p-1}x_{p-1} \quad (9.16)$$

where $y = \ln C(t)$, $x_1 = t$, $\beta_0 = \ln N_0$, $\beta_1 = \ln \lambda$, $x_2 \dots x_{p-1}$ are the auxiliary variables, and there are p parameters to fit with multiple regression.

The use of regression mean square, R^2 and a bias criterion were discussed in Sec. 9.9. Two further criteria have been used for evaluation of models, Mallow's C_p for multiple regression models (Draper and Smith 1981) and Akaike's Information Criterion (AIC) for models where likelihood ratio tests are appropriate (Lebreton et al. 1992, Burnham and Anderson 1996). The C_p statistic is calculated as (Draper and Smith 1998:332):

$$C_p = \text{RSS}_p/s^2 - (n-2p) \quad (9.17)$$

where RSS_p is the residual sum of squares in a multiple regression model based on p parameters, n is the number of observations and s^2 is the residual mean square from the equation with the largest number of parameters in the set of equations evaluated. The method thus depends on having a range of auxiliary variables available for study and essentially assumes that this set of variables includes those involved in the "true" underlying regression model. Draper and Smith (1998:331) suggested plotting residual mean squares against the number of parameters (p) for a sequence of regression models as a way to estimate an asymptotic value that may approach the "true" value σ^2 , which is then used as s^2 in eq.(9.17). They noted that large samples and a sizable number of candidate variables should be available for this approach to be valid.

Burnham and Anderson (1996) gave the Akaike (AIC) criterion as: $AIC = -2(\log\text{-likelihood}) + 2p$ where p is again the number of parameters and the log-likelihood ratio is calculated from the maximum likelihood estimates of parameters in two candidate models. In theory, the method requires that a "global" set of models be identified and that this set contains the "true" model according to Burnham and Anderson (1996). A series of applications of AIC for survival analysis was provided by Lebreton et al.(1992) and Burnham and Anderson (1996) provided a further example. Because regression models are used here for analysis of indices, the C_p statistic seems useful. Burnham and Anderson (1996) discuss the analogous features of AIC and C_p . Much more detail is available in the book by Burnham and Anderson (1998) where it is claimed that the set of models considered does not need to include the "true" model.

Example 9.6 Trend indices with auxiliary variables.

Two examples of trend indices of the use of auxiliary data illustrate the approach of the previous section. One uses the data on manatees studied by Garrott et al. (1994, 1995). The other considers data on the Yellowstone grizzly bear population. Background data for both species appear in the Case Histories.

Garrott et al. (1994, 1995) evaluated a sizable number of potential auxiliary variables, but it appears that year and DD10 (cumulative heating days summed for 10 days previous to the aerial counts of manatees in warm-water refugia) may serve as well as larger sets of temperature variables (Eberhardt, Garrott and Becker 1999). The model assumed for the study was the "proportional hazards" model of eq.(9.15), fitted by multiple regression [eq.(9.16)]. R^2 was about 0.60 for several versions of the overall multiple regression analyses. A difficulty with the results is that the estimated rate of growth of the manatee population exceeded that believed likely (Eberhardt, Garrott and Becker 1999), and estimated from reproductive and survival data (Eberhardt and O'Shea 1995). An alternative approach used regression of repeated counts within years on DD10 for a covariance adjustment (Snedecor and Cochran 1967), as detailed by (Eberhardt, Garrott and Becker 1999).

The alternative approach suggested that the population remained relatively constant over recent years, in accord with the conclusion of Eberhardt and O'Shea (1995). Plotting deviations from the multiple regression fit (Fig. 9.15) suggests that some factor not accounted for in the model may have influenced the trend. As a check on the use of ordinary multiple regression, a nonparametric bootstrapping study was conducted. There were 103 data points in the manatee data set. These were randomly sampled by taking $n = 103$ random samples with replacement

and calculating a multiple regression on each such sample. Doing this 2,000 times yielded estimates and percentile confidence limits essentially the same as those given by ordinary multiple regression. The usual regression program gave an estimate of 0.0958 for the regression slope with a 95% confidence interval of 0.074 to 0.118 while bootstrapping produced a mean estimate of 0.0953 with 95% confidence interval of 0.076 to 0.115. The bias estimate of eq.(9.12) is thus $b_B = 0.953 - 0.0958 = -0.0005$, so there is no evidence from the bootstrapping about problems with the regression approach.

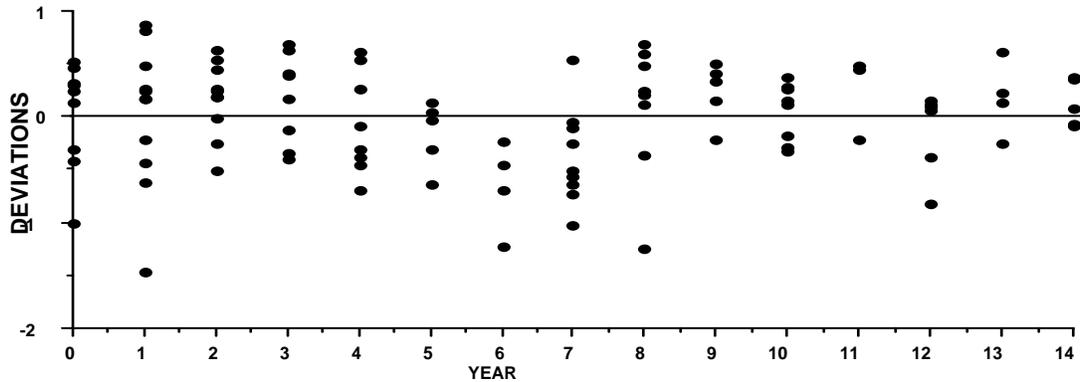


Fig. 9.15 Deviations from a multiple regression model fitted to manatee data.

The Yellowstone grizzly bear data yielded higher values of R^2 and used 3 auxiliary variables: year, frequency of sighting of given family groups, and April snow depths in the previous year. The index variable was counts of "distinct families", i.e., of females with cubs-of-the-year seen in the summer (Knight, Blanchard and Eberhardt 1995). The model thus was:

$$\ln(\text{count}) = \beta_0 + \beta_1(\text{year}) + \beta_2(\text{frequency}) + \beta_3(\text{snowpack}) \quad (9.18)$$

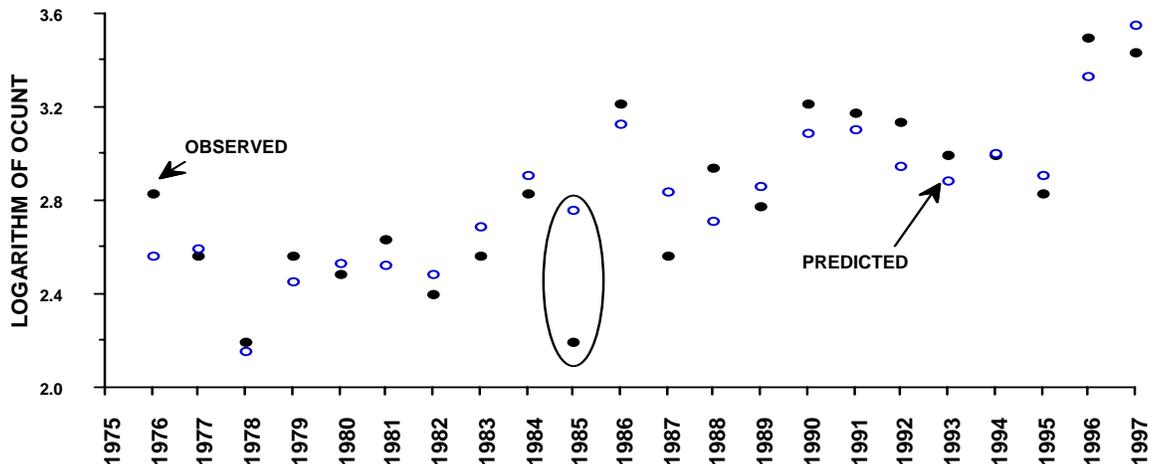


Fig. 9.16. Observed (solid points) and predicted values (open circles) of logarithms of counts of "distinct families" of grizzly bears in Yellowstone National Park and environs. Aberrant 1985 value is circled.

This model yielded an R^2 of 0.75 using data from 1976 to 1997. Comparison of values calculated from the regression equation and

observed counts (Fig. 9.16) suggests that the observed value in 1985 was somehow aberrant. Without this value, $R^2 = 0.85$. A variety of additional variables were examined in multiple regressions, including squared terms for year and frequency of sighting as used in the manatee studies of Garrott et al. (1994, 1995). None of these additional variables appeared to provide useful fits. Mallow's C_p [eq.(9.17)] was calculated (Table 9.8) by plotting s^2 against number of variables included in the regression as recommended by Draper and Smith (1998:331). This suggested a value for s^2 (0.03) from the trend of calculated regression mean squares. Table 9.8 also includes $1 - R^2$ as a measure of the proportion of variance not accounted for by the regression lines. The three measures show essentially the same trend with the number of parameters estimated (p). Another reassuring aspect of the revised index is that there is now little evidence of curvilinearity in the residuals (Fig.9.17) in contrast with an earlier index calculation using only year and frequency of capture (Eberhardt, Garrott and Becker 1999). There were no significant correlations between the independent variables.

Table 9.8. Variation in three measures of regression model adequacy with increasing number of parameters (p) included in the model. Data for a Yellowstone grizzly bear trend index fitted to models of the general form of eq.(9.18).

p	s^2	$1 - R^2$	C_p
2	0.053	0.36	15.6
3	0.039	0.25	7.6
4	0.039	0.23	7.2

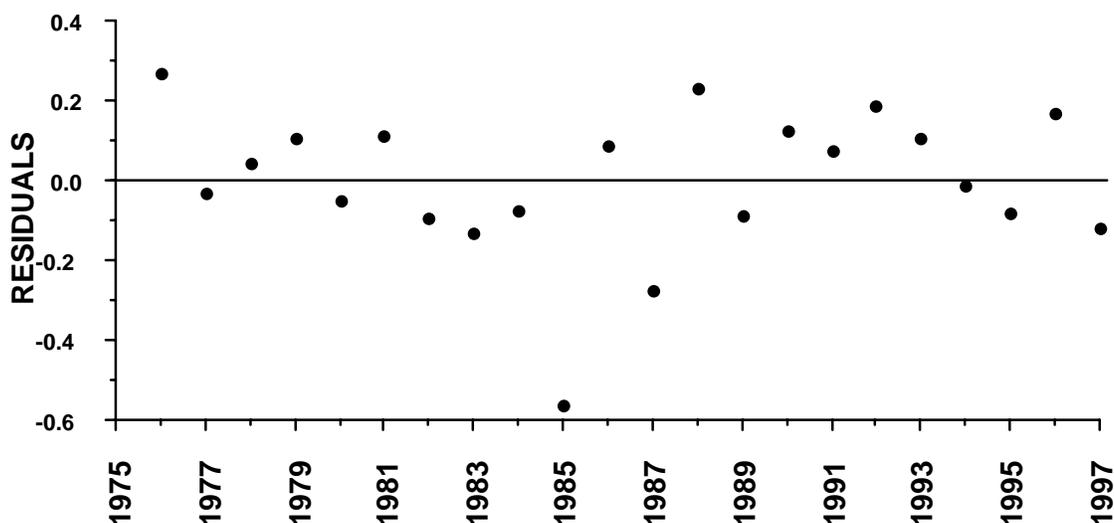


Fig. 9.17. Deviations from regression for the grizzly bear data of Fig. 9.16.

Bootstrapping was used to check the multiple regression calculations, with essentially the same results as for the manatee data. The year coefficient was 0.0287 with a 95% confidence interval of 0.017 to 0.040, while the mean of the bootstrap calculations (2,000 replications) was 0.0291, with 95% confidence interval of 0.018 to 0.040. The bias estimate of eq.(9.12) is thus $0.0291 - 0.0287 = 0.0004$, so there is again no evidence from bootstrapping of problems with the

regression approach. As with the manatee data, the bootstrap frequency distribution was symmetric about the estimate.

Example 9.7 An alternative approach to index models.

A different prospect for assessing trend data can be illustrated by using data on wolves and moose from Isle Royale (Peterson 1995), and a difference equation model used by Eberhardt (1998). The model is:

$$V_t = [1 + r_1] V_{t-1} - \left[\frac{r_1}{K^z} \right] V_{t-1}^{z+1} - c H_{t-1} \quad (9.19)$$

Where V_t denotes ungulate prey abundance at time t , and H_{t-1} denotes predator abundance the previous year, K is the asymptotic population level of prey, z is a constant for the generalized logistic equation (Eberhardt 1987), r_1 is the maximum rate of increase of prey, and c is the predation rate (prey taken per wolf per year). The above model can readily be fitted by multiple regression, giving the results of Fig. 9.18. DelGiudice et al. (1997) indicated that the moose population on Isle Royale was importantly affected by an epizootic of the winter tick (*Dermacentor albipictus*) in 1989, so only the data series through 1988 is used here. R^2 for the regression fit is 0.91, suggesting a good fit.

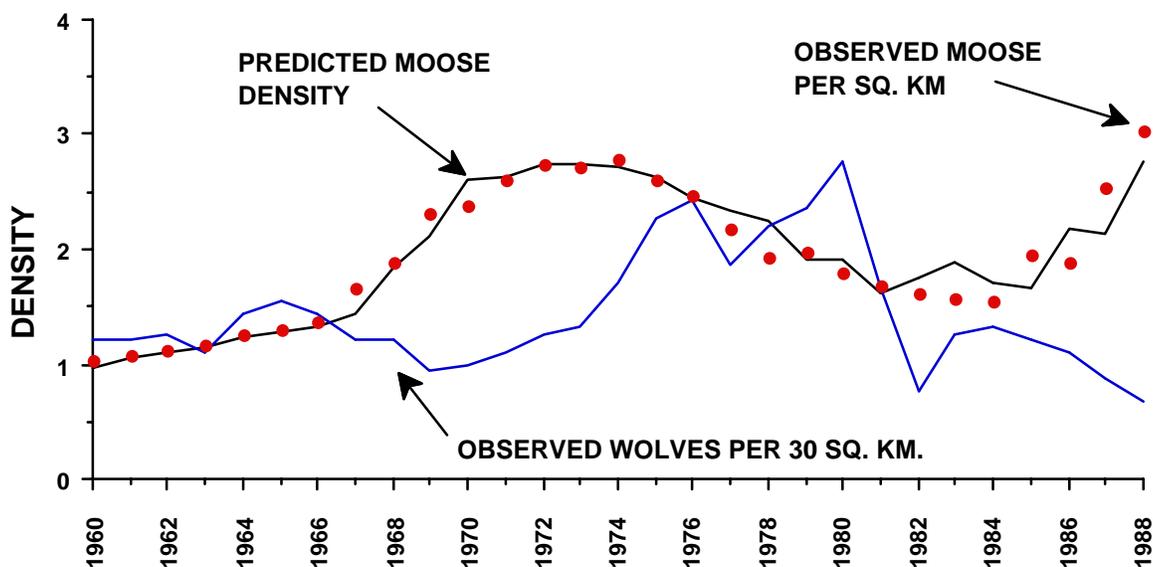


Fig. 9.18. Fit (solid line) of eq. (9.19) to observed data on moose (solid points) and wolves (broken line) on Isle Royale.

Bootstrapping was again applied to try to check validity of the approach. However, the structure of the model forces use of bootstrapping the deviations in which deviations from the fitted model are randomly sampled with replacement, attached to the model fitted to the original data, and refitted. This was done 2,000 times. Using bootstrapping pairs here poses problems, because the model is fit to observations taken sequentially. Results of the bootstrapping study appear to support the model, giving mean values for the 3 coefficients close to those from the original fitting. The original fit estimated $\lambda = 1 + r_1$ as 1.309, and bootstrapping gave 1.314. The second coefficient in the model was -0.00126 from the original data while bootstrapping yielded an average of -0.0013 . The third coefficient (c) was -7.626 in

the original fit, while bootstrap data averaged -7.651 . Frequency distributions of the bootstrapped data gave wider 95% confidence intervals than might be desired. Those for λ were about 1.19 to 1.44, and the interval for the moose kill per wolf (c) was -3.78 to -11.80 . The interval on R^2 was somewhat more satisfactory, being 0.88 - 0.96. The Durbin-Watson test on residuals was 2.01, indicating virtually no deviation from randomness. A plot of the deviations (Fig. 9.19) does, however, emphasize the variability evident in the fit to the later years evident in Fig. 9.18.

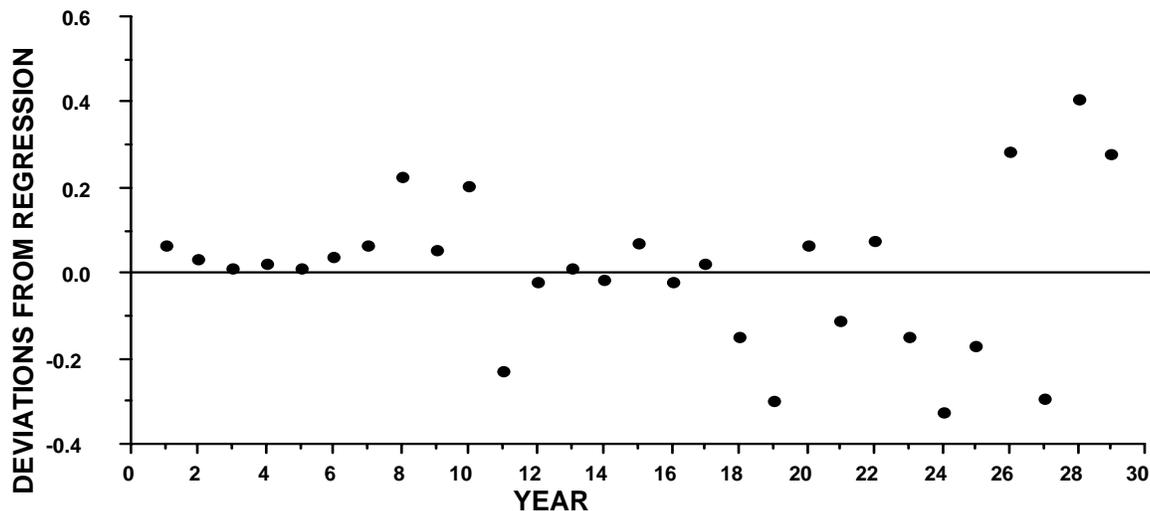


Fig. 9.19 Residuals from the multiple regression fit of eq.(9.19) to Isle Royale moose and wolf data.

One might thus be inclined to suppose that the model of eq. (9.19) may give a useful representation of the data. Unfortunately, the bootstrapping exercise may be of uncertain utility here, inasmuch as Efron and Tibshirani (1993) point out that results of bootstrapping of deviations depend on the assumption that the underlying model is correct, and this may not be true here. One problem is that the model may induce correlations, by virtue of the fact that all but one of the observations appears both in the dependent variable (V_t), and in the first independent variable (V_{t-1}).

The same kind of problem exists in other analyses of the data, but arises in a different way. Mech et al. (1987) used linear regression to relate snow accumulation to moose abundance on Isle Royale. Their results were critiqued by Messier (1991), followed by a response by McRoberts et al. (1995). The data were again studied by Post and Stenseth (1998). The difficulty is that these analyses depend on the ratio of successive population sizes (see, for example, eq.(1) and (2) of Post and Stenseth, 1998). Using such a ratio can induce correlations, as was pointed out by Watt (1964,1968) and further illustrated by Eberhardt (1970). Consequently, analyses should somehow use only the current observations as the dependent variable, and should not include the population index as an independent variable in the regressions. Further study is thus needed to determine just how to proceed. Mech et al. (1987) show various correlations between reproductive parameters and snow depth, so the question is not one of whether winter conditions have an effect, but is rather one of the impact on population trend.

9.11 Catch-effort methods

The catch-effort methods have been developed as a way to use information gained in the course of exploiting a population. By far the main use has been in connection with commercial fisheries. The methods have had very little use in appraising either sports fishing or hunting, although presumably suitable data has been increasingly available from surveys of hunting and fishing. Very much the same techniques are also encountered under the title of "removal methods", mostly in situations where animals are killed by traps, or removed from the study area for some reason.

Use of catch-effort methods in studies of exploited populations depends on the prospect for obtaining large volumes of data in return for a relatively small investment of research or management funds. Complications are, however, introduced by the nature of commercial operations. Times and places of sampling are largely not controlled by the investigators (which has led to increasing use of research vessels that can be operated in specified patterns). The exploitation is usually continuous within the seasons set by regulations, but the amount of effort may vary substantially in time, and may also be correlated with past success and local population densities. These and other difficulties have led research workers to incorporate tagging programs with catch-effort studies. Such tagging programs are almost always of the "single-recovery" type inasmuch as there is no prospect of releasing tagged individuals caught in large scale commercial operations.

In many fisheries situations, there may be more emphasis on estimating rates of exploitation than on population size. For the most part, exploitation rate is measured by estimating survival rates, and this almost always brings in the complications involved in separating fishing and non-fishing mortality rates.

A serious problem in using catch-effort methods is that the "catchability" may change as exploitation continues. One obvious prospect is that the vulnerability of several age-classes to exploitation may differ. As the more vulnerable age-groups are removed, catchability will appear to decrease. Normally such an effect can be studied by examining data pertaining to different age groups separately.

In sports hunting there may be two facets of changes in catchability. One is the higher vulnerability of younger animals, which leads to an apparent decrease in catchability as the season progresses. The other is that heavy hunting effort usually occurs early in the season. In some situations this may mean that hunters tend to interfere with each other and thus reduce the effective catchability early in the season. An additional factor may be that inexperienced hunters tend to give up after a few days; in combination these two factors increase catchability as the season goes on. This plus decreases in vulnerability could make for marked changes in time. However, there are also circumstances, such as deer-hunting in heavy cover, where the higher levels of effort may actually be more efficient in finding and harvesting the available animals.

Apparent changes in catchability may also be due to other circumstances. With relatively short seasons and high effort, sports hunting studies may be conducted as though the population were "closed" to other losses with the exception of "crippling" loss whereby animals are killed but not recovered. For the most part, such losses tend to be proportional to the recovered kill so that the recorded catch per unit of effort is less than actual, and the population size is underestimated. In any case, the effects of changes in catchability and other uncertainties have largely limited application of catch-effort methods to commercial fisheries. The approach should, however, be considered as a potential index method in other situations, so that the main features are discussed here.

9.12 Models for catch-effort data

The basic model for catch-effort studies is one developed by Leslie and Davis (1939) for animal trapping studies and by DeLury (1947, 1951) for fisheries work, and is thus sometimes called the Leslie-DeLury model. Work by Moran (1951), Zippin (1956), Ricker (1958), Chapman (1954), and Hayne (1949) has led to their names also being attached to various versions of the equations.

The population is assumed to be closed to all losses other than the source under study, and to any form of recruitment. One simple and useful way of approaching the method is to visualize random sweeps

of a net through a fixed unit of volume in some large region containing N animals. If the sweeps are of fixed size and are randomly executed, then the prospect that an individual animal is caught in one sweep can be regarded as a binomial-type situation, with the probability of capture (P) depending on the fraction of the total volume swept by the net. It is assumed, of course, that the animals are unable to escape the net and that the sweeps all constitute the same fraction of the total volume. If the sweeps are randomly located, there is no need to assume that the animals are distributed in any particular fashion -- the probability that any given individual is caught does not depend on position of the others if the sweep is done "at random". Such an argument does, of course, have to encompass the very unlikely prospect that all of the animals present could be accommodated in one sweep.

Assuming that the sweeping is done at random, we can determine the probability that a given individual is caught on the i^{th} trial very simply as $P(1-P)^{i-1}$, that is, the animal escapes $i-1$ sweeps, each with probability $(1-P)$ and is caught (with probability P) on the i^{th} trial. Again this pertains only to the fate of a particular individual. If, however, it is postulated that the chance that any one individual is caught does not depend on the fate of the other individuals, then the expected value of the total catch, C_i , in the i^{th} sweep can be written as:

$$E(C_i) = NP(1-P)^{i-1} \quad (9.20)$$

and the equation can be converted to a model relating number of captures in each sampling unit (sweep of the net) to the accumulated units, i.e.:

$$\log C_i = \log(NP) + (i-1)\log(1-P) \quad (9.21)$$

and this equation can be fitted to data by ordinary regression methods. This is the form in which the equation has been used to describe "removal" trapping.

DeLury used a "catchability coefficient", k , rather than P , and considered results in terms of the catch per unit of effort, rather than as catch per sweep of a single net, as above. This is a change necessitated by the continuous nature of a commercial fishery, in which there may be a large number of nets or other fishing "gear" in use simultaneously, and in which the records may be in terms of summaries for fixed lengths of time. It is thus necessary to assume that the various units of effort (which may be, for example, several vessels fishing for a week) are independent (i.e., vessels do not interfere with each other's success) and write the relationship as:

$$C_t = kN(1-k)^{E_t}$$

where E_t represents the cumulated units of effort up to the time when measurement of C_t began; that is, E_t is made as nearly analogous to $(i-1)$ in eq.(9.6) as possible. Also, when k is small and E_t is large, the equation can accurately be approximated by:

$$C_t = kNe^{-kE_t} \quad (9.22)$$

and written in logarithmic form as:

$$\log C_t = \log(kN) - kE_t \quad (9.23)$$

which can be fitted by simple linear regression of logarithms of catch per unit of effort on cumulative effort. It should be noted that the same regression fit can be used for eq.(9.21), with the main difference being in how one interprets the regression slope. In the present equation, the slope estimates k , the catchability coefficient directly, while in eq.(9.21) it estimates $\log(1-P)$. Of course, if P is small (as it will be when large populations are involved) there may be no practical difference, since $\log(1-P)$ is susceptible to series expansion in which the main term is $-P$. Seber(1982:302-303) calls eq.(9.21) Ricker's method, and eq.(9.23) DeLury's regression model.

A serious theoretical limitation on eq.(9.20) is that the successive C_i are by no means independent, a fact emphasized by Moran(1951). If the capture of any individual animal is assumed to be independent of that of any other individual, then a binomial model for the capture of X out of N individuals can be used:

$$P_T[\text{capturing } X \text{ individuals in one trial}] = \binom{N}{X} P^X (1-P)^{N-X}$$

and the model can be extended to cover n trials in which each of C_i individuals are caught. Estimating equations were obtained by Moran(1951) and various approximate solutions and methods for obtaining sample size and so on were obtained for these "removal" methods by Zippin (1956,1958).

Another way to approach the problem of non-independence of successive catches is to consider each such catch in terms of a "conditional" model. If the probability of catching a given individual is regarded as k (as in DeLury's development) then on the i^{th} trial the expected catch is that of the binomial expectation:

$$E(C_i) = k[N - \sum_{j=1}^{i-1} C_j]$$

where the term in brackets represents the number of individuals surviving in the population up to the time of the i^{th} sampling. DeLury generalized this model to represent an average catch per unit of effort, C_t and used K_t to represent cumulative removals up to the time period being considered so that:

$$C_t = kN - kK_t \quad (9.24)$$

This equation can also be fitted to data by simple linear regression methods. We thus have two elementary models, eq.(9.20) and eq.(9.22), that can be applied to data from populations that are "open" only to the removal method being used to estimate population size. Applications mostly have to be limited to rather short time periods and conditions where some other form of loss quite surely does not apply. In the great majority of cases, one has to assume losses from other causes, and often to account for various forms of recruitment to the population under exploitation. These problems, plus uncertainty as to the constancy of k , the "catchability coefficient", may require special auxiliary studies, often accomplished by tagging a number of individuals in the population being studied. Those facing such problems should consult the book by Ricker(1975) and the recent fisheries literature.

Example 9.8 Calculations for catch-effort models

There are a number of ways to estimate variances and confidence limits for the catch-effort models. These are summarized by Seber (1982: Ch.7). Most current usage is based on regression calculations, where the models are represented as:

$$y = a + bx$$

where y is either catch per unit effort (eq.9.21) or $\log C(t)$ as in eq.9.23, while x is cumulative kill or cumulative effort. We then recall that simple linear regression can be viewed as a straight line passing through the means of the observations:

$$y - \bar{y} = b(x - \bar{x})$$

or,

$$y = (\bar{y} - b\bar{x}) + bx = a + bx$$

and that b is estimated by:

$$b = \frac{\sum(y_i - \bar{y})(x_i - \bar{x})}{\sum(x_i - \bar{x})^2}$$

In both of the equations (9.21 and 9.23) k is estimated by b , while a is either kN or $\log_e(kN)$. Confidence limits for k can be calculated directly, as in the usual linear regression calculation for b . However, the regression intercept (a) estimates the product kN , or its logarithm, so that an approximation is required for a variance estimate of N , which is estimated from:

$$\hat{N} = \frac{a}{b} = \frac{kN}{k}$$

or by solving for N in:

$$\log_e(kN) = \bar{y} - \bar{x}$$

i.e.,

$$\hat{N} = \frac{\exp(\bar{y} - \bar{x})}{b}$$

(students should remember that b will be negative in the present situation).

Seber (1982:Ch.7) gives an approximate variance estimate for N appropriate for the logarithmic form (eq. 9.21) as:

$$V[\hat{N}] = s^2 \hat{N}^2 \left[\frac{1}{n} + \left\{ \frac{b\bar{x} - 1}{b} \right\}^2 \left\{ \frac{1}{\sum(x_i - \bar{x})^2} \right\} \right]$$

and for eq. 9.23 he gives:

$$V(\hat{N}) = \frac{s^2}{b^2} \left[\frac{1}{n} + \frac{(\hat{N} - \bar{x})^2}{\sum(x_i - \bar{x})^2} \right]$$

To use these equations in practice, we substitute the "variance about regression" for s^2 . It will be recalled that this is estimated as the "mean square" of the deviations of the observation from the regression line, i.e.,

$$s^2 = \frac{\sum(y_i - a - bx_i)^2}{n - 2}$$

The numerator can be rearranged by using $a = \bar{y} - b\bar{x}$, grouping terms, and then writing out the squared terms:

$$\sum[(y_i - \bar{y}) + b(x_i - \bar{x})]^2 = \sum(y_i - \bar{y})^2 - 2b\sum(y_i - \bar{y})(x_i - \bar{x}) + b^2\sum(x_i - \bar{x})^2$$

Substituting the estimated form of b reduces this to:

$$\sum(y_i - \bar{y})^2 - b^2\sum(x_i - \bar{x})^2$$

and the two "sums of squares" on the right can be calculated from, for example,

$$\sum(y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

Variances of a and b are calculated as:

$$v(b) = \frac{s^2}{\sum(x_i - \bar{x})^2} \quad \text{and} \quad v(a) = \frac{s^2 \sum x_i^2}{n \sum(x_i - \bar{x})^2}$$

That for $v(b)$ can be used directly to obtain approximate confidence limits on k , as:

$$k \pm t_{n-2}[v(b)]^{1/2}$$

Example 9.9 Variable-effort models

In the variable-effort models described above, the fishing mortality rate, F , was assumed constant. When this is an unrealistic assumption, an alternative is to consider mortality due to exploitation to be proportional to the effort, so that $F_i = kE_i$ where E_i is the effort in the i^{th} time period and k is a constant "catchability coefficient" (as used in eq.9.23 and 9.24).

In generalizing the model, one may as well also permit time intervals of varying length, so instead of a constant time period, t , we now use intervals $t_i - t_{i-1}$, where $i = 1, 2, 3, \dots$, and t_0 denotes the beginning

of exploitation. Thus where eq.9.25 contains a term s^{i-1} as the product of constant survival rates through the $i-1$ previous intervals, it now needs to be a product like $s_1s_2\dots s_{i-1}$, which would then be replaced by:

$$s_i = e^{-(kE_i + X)(t_i - t_{i-1})}$$

and the equivalent of eq. 9.27 becomes:

$$kE_i = \frac{kE_i N}{[kE_i + X]} [\exp(-S(E_i + X)(t_j - t_{j-1})) [1 - \exp\{(kE_i + X)(t_i - t_{i-1})\}]]$$

This new equation is not so readily treated by regression methods. One approach is to consider the ratio C_{i+1}/C_i , which, after taking logarithms (to base e) and rearranging gives a complicated equation that can be replaced by an approximate solution by dropping the second logarithmic term and rearranging the result so that a single linear regression equation ($y=a+bx$) results with:

$$y = \frac{1}{t_j - t_{j-1}} \log_e \left[\frac{C_i E_{i+1}}{C_{i+1} E_i} \right], \text{ and } a=X, b=k, \text{ and } x=E_i$$

Estimates of the slope (b) and intercept (a) provide approximate values of the unknown quantities X and k. An improved estimate is then obtained by substituting these trial values of X and k in the second logarithmic term, and then using the entire right-hand side of the equation as y in a new regression calculation. The resulting estimates of k and X can be again substituted and the regression calculated again to check whether the estimates change enough to justify another cycle. Readers familiar with the series expansion of e^{-x} will note that one might start with y values as:

$$y = \frac{1}{t_j - t_{j-1}} \log_e \left[\frac{C_i E_{i+1} [t_{j+1} - t_j]}{C_{i+1} E_i [t_{j+1} - t_{j-1}]} \right]$$

to give a somewhat improved first approximation.

A fault in the above procedure was noted by Chapman(1961), in that the successive y values will be correlated, again violating the essential assumptions for ordinary linear regression. To see this, one can examine the first 3 values of y:

$$\begin{aligned} y_1 &= \frac{1}{t_1 - t_0} \log_e \frac{C_1 E_2}{C_2 E_1} \\ y_2 &= \frac{1}{t_2 - t_1} \log_e \frac{C_2 E_3}{C_3 E_2} \\ y_3 &= \frac{1}{t_3 - t_2} \log_e \frac{C_3 E_4}{C_4 E_3} \end{aligned}$$

from which we see that successive terms contain common elements, e.g., E_2/C_2 appears in y_1 and again in y_2 (inverted). One way to avoid this is to restrict the regression calculations to every other data point; one could thus do two separate regression calculations and average the resulting estimates.

Evidently the above calculations will be somewhat involved and tedious. When effort varies markedly from period to period, an approach like this one seems to be the only realistic answer. Readers should note that when effort is constant from period to period, the basis for this procedure collapses. In fact, other things being equal, the wider the range of

effort encompassed by the study, the more information one gets about k and X .

Example 9.10 Catch-effort data on an "open" population

Some data on a population of tagged juvenile cottontail rabbits may be used to study the effects of natural mortality (Eberhardt et al.1963). The data apply to the 1955-56 hunting season, and were selected from 7 years of similar data as giving the best fit to a regression line. Hence the variance about regression (or correlation) should not be regarded as typical of such data. In the first 3 weeks of hunting, 32 tagged individuals were harvested, but the effort data are not usable, since this period encompassed heavy hunting for pheasants.

Weeks of season	Tagged animals shot C_i	Effort in gun-hours E_t	y $\log(C/E)$	x $E(t)$
4-6	23	984	9.151	0
7-9	22	1167	2.936	.984
10-11	13	1042	2.524	2.15
12-15	9	1059	2.140	9.193

A total of 230 tagged juvenile rabbits constitute the pre-season population. Regression calculations gave the line: $y = 9.196 - .3213x$

One would thus estimate the initial population size as:

$$\frac{kN}{k} = a/b = 9.196/.321 = 99$$

If we add in the 32 rabbits killed in the first 3 weeks ("pheasant season") this gives 131 rabbits as compared to the 230 tagged before hunting began.

Turning to eq. 9.28, we see that the intercept and slope now represent a more complicated expression:

$$a = \log\left[\frac{FN}{F+X}(1-\exp[-(F+X)])\right]$$

$$b = (F+X)^t$$

From this we can estimate:

$$\frac{FN}{F+X} = 88.92.$$

Some other data yield an estimate of the instantaneous rate of natural mortality of 0.089 (calculated on a monthly basis). Using this rate and assuming the above periods to be uniformly two weeks long, and 4 weeks in a month, students should estimate N . To compare the outcome with the initial population (230), assume 4 weeks of natural mortality between tagging and the beginning of "pheasant season" (which was 3 weeks in length).

9.13 Catch-effort models with non-harvest losses

Sources of mortality have commonly been divided into two categories, one due to exploitation and the other due to other causes, operating concurrently with the harvest, and usually described as "natural" mortality. When tagged individuals are included in the analysis it seems best to refer to rates due to

exploitation and to "other" causes in order to be able to provide for the additional mortality due to tagging, and the possible effects of loss of tags and non-reporting of tags.

The elementary model can be introduced by considering survival over some fixed period as a constant rate, s , and supposing that harvest takes a fixed fraction, f , of the population available during the interval. Letting the initial population size be N , we then have the first catch (C_1) as:

$$C_1 = f(1-s)N$$

which can be interpreted as saying that, of the proportion dying in the period, a fraction (f) are taken by the harvest method. The number surviving up to the next period is just sN (since $(1-s)N$ died in the first period), so that:

$$C_2 = f(1-s)sN$$

and, since s^2N survive to the third period:

$$C_3 = f(1-s)s^2N$$

and, in general:

$$C_i = f(1-s)s^{i-1}N \quad (9.25)$$

The notion of "competing risks" (see Ricker 1975), or instantaneous rates, leads to defining:

$$s = e^{-(F+x)t} \quad \text{and} \quad f = \frac{F}{F+X} \quad (9.26)$$

where F denotes the rate of exploitation (F for fishing) and X the "other loss" rate. When the other losses are assumed due only to natural mortality, the symbol M is commonly used rather than X . Since the rates are "instantaneous" and thus are independent of length of the time period, it is necessary to define an arbitrary interval length, t , and write equation (9.25) as:

$$C_i = \left[\frac{FN}{F+X} \right] e^{-(F+X)t(i-1)} [1 - e^{-(F+X)t}] \quad (9.27)$$

If t is now defined in terms of a "unit of effort" then equation (9.27) serves to replace eq.(9.22) in circumstances where the population is "open" to other sources of loss. In effect, units of effort are regarded as operating sequentially on the population, so that we have the effect of i time intervals, each of length t during which the losses take place. In practice, many units of effort are applied simultaneously to the population. The model may nonetheless be satisfactory, but the effect will be one of making estimates of F and X differ from year to year or place to place, depending on how the effort is applied (how it is distributed in time and space).

A number of schemes have been devised to estimate the parameters of eq.(9.26) from actual data. One of the simplest is to take logarithms (using natural logs) giving:

$$\log C_i = \log \left[\frac{FN}{F+X} \{1 - e^{-(F+X)t}\} \right] - (F+X)t(i-1) \quad (9.28)$$

and letting $y = \log C_i$, $a = \log \left[\frac{FN}{F+X} \{1 - e^{-(F+X)t}\} \right]$, $b = (F+X)t$, and

$x = i - 1$, which gives a simple linear regression model:

$$y = a + bx.$$

This can readily be fitted to data on catch per unit effort and time (or cumulative effort). Since t is presumably known, the regression slope (b) gives an estimate of $F+X$. and interest then centers on estimating F , and possibly N , from the regression intercept (a).

When a tagging or marking study is used, N is a known quantity and the regression intercept can be written as:

$$a = \log F + \log N + \log \left[\frac{1 - e^{-(F+X)t}}{F+X} \right] \quad (9.29)$$

so that an estimate of $\log F$ can be obtained by subtracting $\log N$ and an estimate of the quantity in brackets calculated from the slope estimate (b) of $(F+X)$ and t . An immediate problem with estimation from eq.(9.28) is that the conditions of such a study do not conform very well to the assumptions required by a simple linear regression estimate.

In many situations, it is unrealistic to regard F as being constant from time unit to time unit, because of substantial variations in the effort expended in each time unit. Although grouping of units of effort has been used to produce new units of roughly equal size, such a practice is not very satisfactory if X , the "other loss" rate, is constant over time.

9.14 Sampling for indices

The various uncertainties about interpretation of indices, combined with conditions of their use, do not make it any less important to conform to good sampling practice in the use of indices. By and large, most index methods are applied on rather large and heterogenous areas so there is generally a prospect for considerable gains in efficiency (and the concomitant reduction of effort required) through the use of standard methods, like stratification. Since indices are usually rather tightly tied to seasonal conditions, a sampling design needs to be arranged to permit study of seasonal effects, too.

Specific sampling methods need to be tailored to each index, and this is best done by writing a model of the kind already described but containing parameters relevant to the major factors that must be considered in the particular instance at hand. Where possible, it seems desirable to arrange the design in two or more stages. The first stage should be constructed in terms of those factors that are fairly well understood, so that the effects of uncertainties about a particular index are most likely to crop up in the second (or lower) stage. In nearly all cases, one might expect to do a reasonably efficient job of stratification by area and time, and to restrict the effect of other factors on an index to subsampling within the strata.

Stratification very likely will need to be more nearly of the kind generally described as "analytical" in view of the hazards of getting area effects entangled with strata. That is, generally one would make an individual stratum out of as many contiguous units as possible, and try to avoid having scattered units belonging to the same stratum. However, if calibration is attempted, one cannot usually have more than a few strata, or costs of the requisite independent density estimates become too high.

9.15 Transformations

Most of the few statistical analyses of index data done so far have employed transformations. The stated purpose of the transformation usually is to attempt to achieve normality. There are, however, other requirements for the analysis of data of this kind. One arrangement is:

- (1) additivity
- (2) constancy of error variance
- (3) normality of distributions
- (4) independence of observations.

The usual goal in making a transformation is to carry out an analysis of variance. Thus one might have results of, say, roadside counts carried out over several years and in a number of different areas, and wish to know if there are statistically significant differences in the underlying population levels between areas and among years. Nearly all of the published analyses of index data have been so handled. An example of the analysis of variance on index data appears in Example 9.4.

Securing independence of observations is largely dependent on how the observations are taken. Constancy of variance and normality of error distributions are often taken as one feature, while additivity implies that the underlying model is linear, that is, of the form:

$$x_{ij} = A_i + B_j + e_{ij} \quad (9.30)$$

so that an index value in year i and area j depends on an effect due to years (A_i), one due to areas (B_j) and a "random error component" (e_{ij}) which should be normally distributed with $E(e_{ij}) = 0$ for analysis of variance purposes. The usual analysis of variance table gives components associated with the elements of eq. (9.30) and uses the estimate of $V(e_{ij})$ as a basis for tests of significance of the "contributions" of year and area terms to the total variability.

Most of the available analyses of ecological data seem to proceed along the lines of seeking a transformation that tends to normalize the data, and then assuming that eq.(9.30) is satisfied. This is a dangerous course, if the investigator is really interested in sorting out the various effects influencing his data. In that case, the essential need is to achieve linearity (additivity) in the model. From the form of equations (9.22), and (9.27) it is evident that a logarithmic transformation will often be required. If the true relationship of index and density is as postulated in these equations, taking logarithms effectively converts them to additive (linear) models like eq.(9.30). If the results are to conform fully with the analysis of variance requirements, then one must also assume that an error component is multiplicative, i.e., for a roadside count conducted in the same area for a number of years and areas one would write the model as:

$$x_{ij} = B_j D_{ij} e_{ij}$$

so that:

$$\log x_{ij} = \log B_j + \log D_{ij} + \log e_{ij} \quad (9.31)$$

and investigate the prospect that there are area effects on the proportionality coefficient (B_j) as well as assessing population differences. If the error component is assumed to be multiplicative, and it is further assumed that taking a logarithm converts it to a normal distribution, then those two assumptions imply that the original distribution was lognormal in form. However, such assumptions may be somewhat fictitious as evidenced, for example, by the previous remark that pellet-group data follow the negative binomial distribution reasonably well. However, the difference between negative binomial and lognormal distributions may not have much of an effect after log transformations.

Use of the analysis of variance technique on index data calls for rather more detailed study than is possible here. The major point to be made is that blind use of a transformation seems extremely unwise. In most practical cases, the experienced investigator really has little interest in some of the tests of significance. He will have long since concluded that there are real differences in population density between areas, and will largely be concerned with year to year changes on individual areas, and with studying the magnitude of differences between areas. He will also want to look for effects due to observer, as might be identified in Example 9.3 in logarithmic form (where i pertains to observer and j to area):

$$\log X_{ij} = \log B_i + 2 \log r_j + \log D_j + \log e_{ij}$$

Analysis of some of the other equations might be considered in similar terms, but, as already noted, the usual approach is through regression methods. Regression equations were previously mentioned here as tools for estimating some parameter (e.g., population size) in the model, but they can also be studied in an analysis of variance format.

One further aspect of transformations is that investigators tend to be somewhat uncomfortable about presenting results of an analysis in terms of the transformed variable. For the roadside count example, one may originally have a variable expressed as so many animals seen per mile of driving (or per hundred miles, hours of observation, etc.). The logarithmic transformation yields an "unnatural" kind of datum. The usual advice is to transform back to the original scale, neglecting the prospect that such a transformation will introduce a bias. Sometimes this may be the best procedure, particularly if the results are to be used extensively for administrative guidance. However, it should be remembered that we are really considering the use of an index and there should not be any special disadvantage to using an index in logarithmic units. In fact, the behavior of the index (additivity) may be much more suitable on the transformed scale. Consequently, It is often not desirable to transform back to the original scale. Staying in the transformed scale is particularly desirable if several indices are combined, as described in the next section.

9.16 Combining indices

The problem of combining population indices has not been investigated in any detail in ecological studies as yet. Related problems exist in econometrics, but the results obtained there have not been applied in ecology as yet. We will assume that several independent indices are available, and that the need is to combine them to construct a single, overall index. Independence here means derived from different and unrelated sources, and thus refers largely to the sampling methods.

Part of the problem is that the individual indices will generally have different scales of measurement, in consequence of the kinds of information on which they are based. One possible structure for an index value is:

$$X_{1i} = \beta_1 D_i + e_{1i} \quad (9.32)$$

where D_i represents the true density and β_1 represents a "scale factor" or a "proportionality constant", while e_{1i} represents a random error component. We thus assume that a given index value is proportional to the true density, but has associated with it a randomly selected fluctuation, due presumably to a variety of influences on the observational process. The usual approach in problems of this type is to assume that $E(e_{ij}) = 0$.

Equation 9.32 can be used to indicate what results when correlation or regression techniques are applied to indices. Supposing we have a second index, with structure:

$$X_{2i} = D_i + e_{2i}$$

Then an indication of the behavior of the correlation and regression coefficients, assuming large samples, can be given as follows:

$$E(s_{x_1}^2) = \beta_1^2 \sigma_D^2 + \sigma_{e_1}^2$$

$$E(s_{x_2}^2) = \sigma_D^2 + \sigma_{e_2}^2$$

$$E(s_{x_1 x_2}) = \beta_1 \beta_2 \sigma_D^2$$

where σ_D^2 refers to the true variation in population density over the set of areas being investigated. Then the sample correlation coefficient (r) is approximately:

$$r = \frac{s_{x_1 x_2}}{[s_{x_1}^2 s_{x_2}^2]^{1/2}} \quad (9.33)$$

$$= \frac{\beta_1 \beta_2 \sigma_D^2}{[\{\beta_1^2 \sigma_D^2 + \sigma_{e_1}^2\} \{\sigma_D^2 + \sigma_{e_2}^2\}]^{1/2}}$$

so that r approaches unity only if $\sigma_{e_1}^2$ and $\sigma_{e_2}^2$ are nearly zero. Consequently, a very high correlation between two indices might be taken to mean that both are nearly directly proportional to true density. However, this depends on the indices being obtained from independent sources (otherwise the two indices may simply have highly correlated errors), and on the correctness of equation 9.32.

The regression coefficient (b) is approximately:

$$b = \frac{s_{x_1 x_2}}{s_{x_2}^2} = \frac{\beta_1 \beta_2 \sigma_D^2}{\beta_2^2 \sigma_D^2 + \sigma_{e_2}^2} \quad (9.34)$$

so that it is necessary for $\sigma_{e2}^2 = 0$ before the regression coefficient will reflect the true ratio of the two coefficients. This is the situation where one index is exactly proportional to true density, and one that is not likely to be encountered in practice. It is also the usual condition for regression analysis (i.e., that the independent variable be measured without error). If the ratio of σ_D^2 and σ_{e2}^2 is somehow known, then the regression coefficient can be estimated without bias. In the usual circumstances, one does not know the ratio, and the problem becomes difficult to handle.

In many cases, the investigator will have at least a rough idea of the sampling effort that went into each index. Such information can serve as a source of weights for combining indices -- usually just by converting the "sample sizes" to proportions and multiplying the transformed index values by these weights. Before doing so, it is necessary to convert the indices to the same scale. It also seems essential to have roughly the same spread of values for each index. One simple way to achieve this result is to transform the several indices to have about the same mean and variance. If the mean and variance are selected as some convenient values, say Z and S^2 , then the coefficients A and B for the transformation of the i^{th} index:

$$Z_i = BX_i + A$$

are obtained by noting that the variance of the transformed index should be equal to the original index adjusted by a constant:

$$\frac{\sum(z - \bar{z})^2}{(n-1)} = S^2 = \frac{B^2 \sum(x - \bar{x})^2}{(n-1)}$$

so that:

$$B = S/s, \quad (9.35)$$

and since $Z = B\bar{x} + A$ we have:

$$A = Z - (S/s)\bar{x} \quad (9.36)$$

A new set of coefficients (A, B) has to be calculated for each set of index data. The transformed index values are then combined by weights obtained as suggested above:

$$Y = W_1 z_1 + W_2 z_2 + \dots + W_n z_n \quad (9.37)$$

where there are n indices and the weights (W_i) sum to unity. A transformed value will, of course, be obtained for each area and time period under study, so that the y of eq.9.37 might be written as y_{ij} to pertain to the i^{th} area and j^{th} time period. Then the index values would have to be written as x_{ijk} (k for k^{th} index), and so on.

All of the above pertains to a model (eq.9.32) which assumes that the errors are additive. In Sec.9.12 it was pointed out that the logarithmic transformation implies that the model really is:

$$X_i = \beta D_i e_i$$

and the logarithmic transformation presumably makes it possible to investigate index behavior over a set of areas. Combining different indices will then partly be conditioned by the results of the statistical analysis of individual indices, and clearly can become quite complicated.

9.17 Converting indices

It has already been noted here that there are circumstances where an index value is not suitable for management purposes. This suggests a need for ways to convert an index to an estimate of actual density -- or for "calibration" of an index. If a set of areas exists on which true population density can be estimated, then an apparent solution is just to compute the regression of true densities on index values. However, it seldom is possible to measure true density without error -- ordinarily some sampling process is involved. This then puts us in exactly the same position as led to eq. 9.34, except that it may now be assumed that $\beta_1 = 1$, so we have:

$$b = \frac{1}{\beta_2 + \frac{\sigma_{e_2}^2}{\beta_2 \sigma_D^2}} \quad (9.38)$$

and a knowledge of the ratio of the two variances is required to obtain an estimate of $1/\beta_2$ for conversion or calibration purposes. Some elements of strategy for planning calibration work are evident from the above relationship, though, i.e., keep $\sigma_{e_2}^2$ as small as possible and choose the set of population densities used to have as wide a range as possible. Very likely these may be conflicting aims, since the choice of low densities may tend to increase $\sigma_{e_2}^2$.

When the regression relationship can be assumed to go through the origin, and if the variance of y increases proportionately with increasing x , then the ratio estimate (cf. Cochran 1977) is known to be optimum and unbiased. Using the same model as above, we have:

$$E(\Sigma y_i)/E(\Sigma x_i) = 1/\beta_2$$

so this estimate is approximately unbiased. In most real-world situations, these are the more likely assumptions than those of linear regression, so ratio methods are to be preferred to regressions, here.

As mentioned before, there are reasons to doubt the accuracy of the usual approximations to variance estimates (and hence confidence intervals) arising from the presence of errors in the independent variable (X). Very possibly there should be another component of variance in the equations. A related issue has to do with the distinction between ratio estimators (total of the X_i assumed known) and double sampling (only a sample of the X_i assumed known). With measurement errors in the X_i , one can have an observation on X_i on every study unit and still not "know" the total. That is, were a new survey possible, it would not give the same total.

A somewhat pessimistic view of what is known about double sampling with errors in the independent variable can be tempered, however, by another look at eq.9.33. As remarked there, the sample correlation coefficient (r) approaches unity only if measurement errors in X and Y are nearly zero. Hence observing sample correlations on the order of 0.9 or better with index data, gives one some considerable encouragement to think that the measurement errors are small, at least in relation to σ_D^2 . Hence, there are some grounds to suppose that the bias in eq.9.38 may not be unreasonably great, supposing β_2 is not very small.

9.18 Comments on the use of index data

Although Fig. 9.1 suggests a variety of ways to approach the use of indices, a great deal of research may be required before really suitable methodology is available for dealing with indices. Some of the problems may be worth mention here. One very important issue is selecting a model for the analysis. Using a simple multiplicative model has several advantages. One is the log-transformation which tends to "normalize" the data. The second is that it produces a linear relationship with time, if the population is changing at a constant rate (e.g., growing exponentially). One very important question is whether prospective auxiliary variables do in fact have a multiplicative effect on the index, so that the log transform yields a simple linear model. Perhaps an even more important question is whether the actual population is changing at a constant rate.

At present, the main approach to dealing with a situation where a number of variables may be involved seems to be to identify any variables that may possibly be relevant (and, of course, that can be measured!). Because the underlying relationship may not be linear, a squared term is sometimes introduced in the set of variables to be studied. An effort is then made to find out which of the several variables considered may be worth keeping in a final model. One might, for example, want to consider 5 candidate auxiliary variables. Including squared terms for each, then yields 10 variables. The immediate problem is

one of reducing the set to some smaller number of "significant" variables. One approach is to compute all possible regressions with each of the candidate variables left out in turn, and with that variable included, and to use one or more criteria to see whether inclusion of a given variable is worthwhile. Three such criteria are frequently considered, one being the regression mean square, the second R^2 [eq.(9.11)] and the third C_p [eq.(9.17)].

With as many as 10 variables, this approach gets out of hand, insasmuch as there are 2^f equations to study ($2^{10} = 1024$). An alternative is to start with an equation containing all 10 variables and use a pre-determined procedure to work back through the list and to eliminate those variables that do not meet certain criteria. A third procedure is stepwise regression, starting with the "best" variable (most highly correlated with the index) and work through the set, adding a new variable if it meets a significance criterion, and stopping when there isn't a significant improvement.

Because a lot of computing is involved, these procedures generally depend on a computer program. Different results may be obtained from different procedures, and depending on the order in which variables are introduced, and opinions on how to proceed may differ from reference to reference. One should always examine (plot) the residuals from regression to see whether they provide any hints as to possible improvements.

The C_p criterion seems to me unlikely to be very useful for index studies. As described by Draper and Smith (1998), it may be most useful in situations where there is some reason to believe that most of the variables relevant to the process being studied are included in the set to be analyzed. C_p may then serve in picking out sets that are somehow "adequate" to describe the process. It is unlikely that one can hope to measure most of the variables that influence a population index.

The likely situation can be suggested for the bear data of Example 9.6. Fig. 9.20 shows the data of Fig. 9.4 with the regression mean square of the adjusted bear index [eq. (9.18)] added. The inclusion of auxiliary variables has reduced the regression mean square to about 40% of that of the unadjusted index, and thus presumably yields a better index of population trend. However, there are a lot of smaller regression mean squares in the data set of Fig. 9.4, presumably because these populations were more accurately enumerated. Using the delta method to approximate the variability of the bear population (assuming survival rates and population sizes suggested by Eberhardt et al. (1994) and Eberhardt and Knight (1996) gives roughly the "true" variability shown at the far right side of the figure. Clearly a good deal of the variability in the index is unaccounted for. Sampling or measurement error in determining two of the auxiliary variables (frequency of sighting and snow depths) may account for some of the difference, but certainly not all.

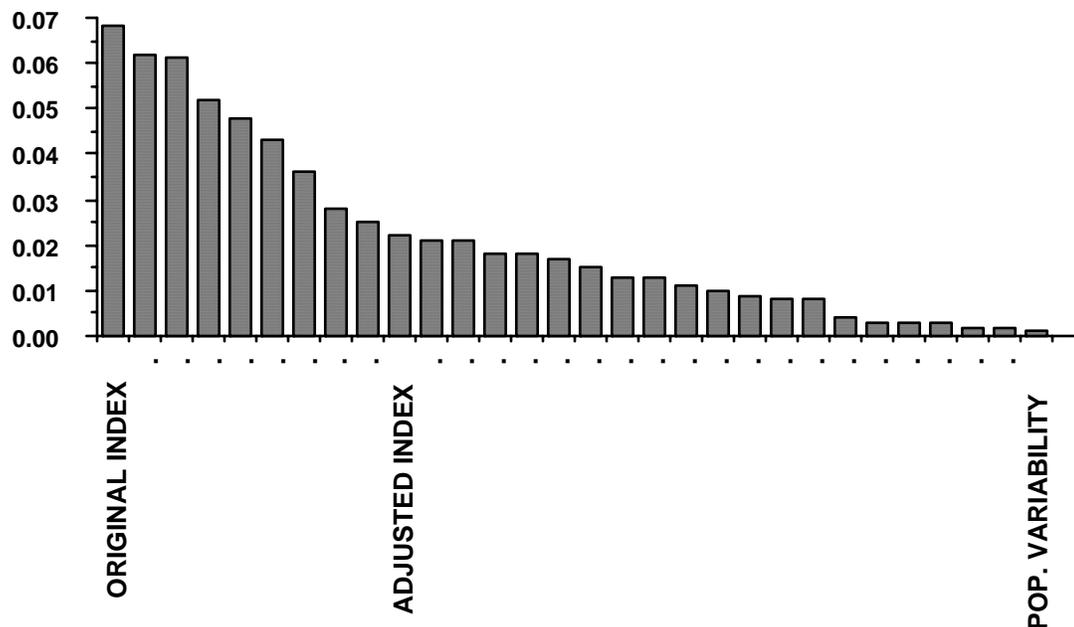


Fig. 9.20. Data of Fig. 9.4 with adjusted bear index and estimates of actual population variability added.

One may thus be inclined to doubt the utility of C_p in ecological index studies. Because it utilizes RSS_p , the regression sum of squares, it will likely follow the trend indicated by the regression mean square. Also, RSS_p for the model with all significant parameters included is the same as $RSS(\text{full})$ of R^2 . Hence, it may be best to use regression mean square and R^2 as criteria for studying potential indices by multiple regression. One can interpret the resulting regression mean square by comparison to other values plotted in Fig. 9.4. It might also be noted that an earlier comment (Sec. 9.9) that " R^2 cannot become unity" is strictly true, but the very low value of estimated population variability for bears (Fig. 9.20) suggests that R^2 will become very close to unity when variability in regression mean squares is mainly from stochasticity. A number of such examples are present in the data used for Fig. 9.4, as shown in fig. 9.13.

9.19 Exercises.

9.19.1 Plot the data of Example 9.2 and suggest how one might go about trying to determine whether the two estimates of b given in the example are different.

9.19.2 Plot the data on Gambel quail from Example 9.3 and the ratio and regression lines. Which of the two lines appears to fit the data best? Why?

9.19.3 Conduct the analysis of variance described in Example 9.4

9.19.4 Data for counts of bison and for muskox (Fig. 9.3) are given below. Using natural logarithms of the data fit a straight line and the "quadratic" of eq. (9.4) and use the F-test of eq. (9.6) to test for curvilinearity. Discuss the difference in results between bison and muskox. Plot the residuals to see if they help in explaining the difference.

YEAR	NO. OF MUSKOX	YEAR	NO. OF BISON
1950	61	1909	37
1951	76	1910	48
1952	77	1911	70
1953	90	1912	85
1954	100	1913	104
1955	116	1914	130

1956	126	1915	164
1957	143	1916	194
1958	181	1917	240
1959	206	1918	295
1960	256	1919	367
1961	293	1920	420
1962	353	1921	479
1963	406	1922	554
1964	467		

9.19.5 Using the data on brown bears of Table 9.3 without year 6 (only one point) conduct a one-way analysis of variance for differences among years using untransformed data and then using a log-transform. Which seems to be the best approach? Why?

9.19.6 Estimate lambda for the gray whale data of Table 9.4 using eqs. (9.7), (9.8), and (9.9). Compare the resulting estimates with the rate obtained from log-linear regression on the data.

9.19.7 Calculate the Durbin-Watson test [eq.(9.13)] on the residuals from log-linear regression from the gray whale data of Table 9.4. Report the value of d and comment on its meaning. What is the mathematical relationship between the slope and lambda for Table 9.7?

9.19.8 Repeat the calculations for Example 9.5 (sea otter census). Estimate the ratios for each day from eqns. (9.7), (9.8), and (9.9) and compare with the ratio estimates obtained by pooling the data from all 3 days. Do they suggest some differences? How might you test for significant differences between days?

9.19.9 Data for the counts of Yellowstone grizzly family groups (Example 9.6) appear below (note that the snowpack data have been “centered”—deviations from the mean are shown).. (1) Compute R^2 and s^2 for loglinear regression on the original counts. (2) Then compute R^2 and s^2 for the full data set, i.e., using the two auxiliary variables along with year. (3) Extend the computations to include t^2 (year-squared) as an auxiliary variable. Discuss your results. What is the effect of t^2 ?

LNCOUN	YR	FREQ	SNOPAK
T			
2.833	1	1.64	7.568
2.565	2	1.5	10.368
2.197	3	1.28	-18.432
2.565	4	1.08	3.868
2.485	5	1.4	1.568
2.639	6	1.58	-4.432
2.398	7	1.62	-9.832
2.565	8	1.2	9.568
2.833	9	2.29	2.868
2.197	10	2	-4.332
3.219	11	3.12	-0.632
2.565	12	1.64	3.568
2.944	13	2.12	-15.732
2.773	14	1.86	-3.132
3.219	15	1.95	8.968
3.178	16	2.65	-4.732
3.135	17	1.65	0.468
2.996	18	1.67	-6.332
2.996	19	1.47	3.268

2.833	20	1.47	-5.332
3.497	21	1.96	12.668
3.434	22	2.95	8.168

9.19.10 Do part (2) of Exercise 9.19.9 but use the actual years (1967, 1968, 1969,...) instead of 1,2,3 and compare the results. Some references recommend “centering” the data. That is, instead of using 1967, 1968, 1969,... find the mean of this column and use the deviations from the mean as the x-variable. Try this and compare your results.

9.19.11 Note that (Fig. 9.16) the observed count for 1985 differs considerably from the predicted value. Do you think that value should be dropped from the index? Can you justify your answer statistically? How?

9.19.12 Compute C_p and AIC for the grizzly bear data using all of the auxiliary variables in the table. To compute AIC you need to know that the log-likelihood value for a linear model with normal errors is just the usual s^2 , but calculated with n as divisor rather than $n-p$. That is, compute the sum-of-squares and divide by n . The assumption of normal errors is not strictly defensible statistically, but the log-transform seems to result in quite symmetrical distributions from regression given reasonably large samples.

9.19.13 Lobster catch data

DeLury (1947) gave the following catch data for lobsters:

<u>Date</u>	<u>C(t)</u>	<u>K(t)</u>	<u>E(t)</u>
May 23	82	0	0
24	75	7	8
25	94	13	16
26	80	16	19
27	83	22	27
29	89	25	32
30	70	32	40
31	58	37	48
June 1	64	40	53
2	55	45	61
5	52	50	69
6	45	53	76
7	45	54	77
8	49	55	79
9	45	57	85
10	48	60	90
12	43	62	96

Using eq.9.23 he estimated $k = .008348$ and $N = 112.34$, using simple linear regression as outlined above. From eq.9.24, he obtained $k = .007984$ and $N = 116.33$, also by linear regression. Students should repeat the calculations, and calculate variances and confidence limits.

9.19.14 Apply the expression for confidence limits given at the end of Example 9.5 to the 2 values of k obtained in the lobster catch data of Exercise 9.19.3. Report your results.

Exercise 9.19.15 Removal trapping

Zippin (1956) illustrated removal methods by assuming catches of 165, 101, and 54 animals were caught and removed in 3 nights of trapping. Calculations for eq. 9.21 are then based on the following data:

i	C_i	$y_i = \log_e(C_i)$	$x_i = i-1$
1	165	5.1060	0
2	101	4.6151	1
3	54	9.9890	2

Regression calculations will then proceed just as in the examples above, but we now have $b = \log(1-p)$. When p is small, we can represent $\log(1-p)$ by $-p$ and the calculations are essentially those already described. However, in this case, p is clearly not small, so that eq. 9.21 is appropriate. Students should carry out the regression calculations and compare the estimate of N obtained by assuming $b = p$ and $b = \log(1-p)$. Improved methods of estimation are available (Zippin 1958; Seber 1973,1982:Ch.7) but require an iterative solution or the use of graphs (given in both references). Variance calculations are similarly complicated, and should be approached through the references cited.

Attention to variability is important in designing a study based on removal trapping, as a substantial fraction of the population must be caught in order to obtain a reasonably precise estimate. If we let $q = 1-p$, then the following equation (Zippin 1956:171) approximates the variance:

$$V[N] = \frac{N(1-q^n)q^n}{(1-q^n)^2 - q^{n-1}}$$

Students should try this equation for a few values of p , n , and N . Increasing n (beyond 3) doesn't do much to reduce the variance, which means that p has to be increased (by using more traps). This has a considerable practical significance, in that it usually isn't sensible to run a removal trapping program very long, since immigrants will soon show up, violating the assumption of a closed population. Tagging some animals before the study starts is always a wise precaution, if feasible.

Exercise 9.19.16 Combining and comparing indices

In section 9.16 it was suggested that several different indices might be combined by: (1) transforming the individual indices to a common scale, and (2) weighting the transformed values by some independent measure of their variability. An example was given by Eberhardt (1960), portions of which are reproduced here to exhibit calculations. Four individual indices were used: (1) Accidental highway kills (recorded by Conservation Department staff. (2) "Camp kill" (estimates of deer taken on a special "camp" license, obtained through a mail survey of hunters). (3) July deer counts (a roadside tally by Conservation Department personnel. (4) Archery kill (deer taken on a special "bow and arrow" hunting license, estimated from an independent mail survey of those hunters).

An arbitrary transformation to a variance of 9.0 and mean (Z_i) of 4.0 was used as given by eqs.9.35 and 9.36. An example of the calculations appears in the second table below. Students should check their understanding of the equations by repeating the calculations. The transformed data were then combined into a single index as in eq.4.15, by using a set of weights, W_i that summed to unity. One possible choice of weights was the sample sizes for the various indices:

<u>Index</u>	<u>Units used</u>	<u>Average number per year</u>	<u>Square root</u>	<u>Weight</u>
July count	Number of deer seen	5000	70	.409
Archery kill	Number of hunters in samples	2000	44.7	.258
Camp kill	Number of parties in samples	200	14.1	.081
Highway kill	Number of deer tallied	1900	49.6	.252
				1.0000

Transformation of roadside deer counts to standardized values

Deer seen per 100 hours

District	<u>1952</u>	<u>1953</u>	<u>1954</u>	<u>1955</u>	<u>1956</u>	<u>1957</u>	<u>1958</u>
1	18.4	16.4	29.5	19.3	20.7	17.6	17.1
2	29.9	29.4	34.8	29.5	37.0	30.9	40.3
3	30.9	35.1	37.1	35.2	39.9	32.8	35.5
4	27.8	27.4	42.5	26.5	37.0	30.1	30.5
5	49.6	32.8	28.1	30.4	29.8	32.9	35.0
6	39.7	11.3	16.6	8.5	19.7	11.5	19.2
7	71.4	78.6	77.8	58.2	38.1	49.2	59.7
8	20.6	19.1	14.2	14.0	18.9	16.9	24.8
9	15.0	12.2	19.2	9.9	16.8	22.1	16.7

Transformed values

1	7.93	7.54	4.09	8.10	8.38	7.77	7.68
2	9.00	4.07	11.12	4.09	11.55	4.36	12.19
3	4.36	11.18	11.57	11.20	4.95	4.73	11.26
4	9.76	9.68	12.62	9.50	11.55	4.21	4.28
5	12.83	4.73	9.82	4.26	4.15	4.75	11.16
6	12.08	6.54	7.58	6.00	7.01	6.58	6.92
7	18.25	19.65	19.49	15.68	11.76	12.76	15.97
8	8.36	6.90	7.11	7.07	8.02	7.64	9.17
9	7.27	6.72	6.92	6.27	7.62	8.65	7.60

Data for transformation

$$s^2 = 237.315$$

$$s = 15.405$$

$$x = 29.041$$

$$n = 63$$

$$B = S/s = 9.000/15.405 = .1947$$

$$A = 10 - .1947(29.041) = 4.3457$$

Check on transformation

$$s^2 = 8.9991$$

$$z = 10.0002$$

Several checks on behavior of these weights were used, including: (1) error mean square in an analysis of variance of the index (compared as coefficients of variation), (2) mean square deviation from regression (against pellet count data), and, (3) correlations among the 4 indices. Since none of these analyses provided a measure that would be independent of the index observations, they were only used to check on behavior of the weights, i.e., to show that the weights were roughly correct.

Execution of a transformation to the same scale and appropriate weighting does not provide much evidence about validity of the combined index. In the present example, there were two independent measures that could be used for this purpose. One was the pellet group counts (already described here), and the other a population estimate based on sex composition, age structure, and kill (harvest) data, labelled the S-A-K method. Correlations between the 3 sources were:

	<u>Combined index</u>	<u>S-A-K</u>
S-A-K	.934	--
Pellet counts	.954	.951

There is thus independent evidence that the index did indeed provide a good measure of population levels.

In a Wisconsin study, McCaffrey (1976) introduced another index, a count of deer trails intersecting 0.4 km transects. Correlations were:

	<u>Trail index</u>	<u>S-A-K</u>
S-A-K	.94	--
Pellet counts	.89	.781

It thus seems that this index is also well-correlated with independent measures of deer population size.

The question of how one uses a "good" index once it has been analyzed and shown to be well-correlated with direct measures of abundance is one needing further research. In broad outline, it has been suggested here that, when the direct estimate is available for only portions of the study area, while the index is available for the entire area, then one might use double sampling (ratio method) to use the index to arrive at an estimate for the entire study area. Presumably predictions of population levels might be made for sub-areas from the same relationship. These will, however, be quite variable, and just how confidence limits on these predictions might be obtained seems to me to be an unresolved question as yet. In areas where both index and direct measure are available, it does not seem feasible to use the index, unless it is converted to an estimate on the basis of prior (not current) experience. The current estimates should, however, be used in checking to see that the index is still "in calibration".

10.0 ESTIMATING SURVIVAL

10.1 Introduction

There is a considerable variety of ways to estimate survival. Many of the available methods come as by-products of methods designed primarily for measuring population size (Chapter 8), and are thus best discussed in reference to that problem. However, knowledge of survival is sufficiently important to justify specific studies without special reference to determining population size. Such efforts are usually dependent on some form of marking or tagging. Since marking is almost always a very expensive and time-consuming operation, it should, wherever possible, also result in estimates of population numbers. However, we will here be concerned mainly with measuring survival.

A survival rate is necessarily defined for some unit of time, thus:

$$\begin{aligned} \text{Survival rate} &= \frac{\text{Number alive at end of period}}{\text{Number alive at start of period}} \\ &= \text{proportion surviving the period.} \end{aligned}$$

When dealing with populations exploited by hunting or fishing it is frequently necessary to consider two components of mortality, one due to the exploitation and the second to all other causes of death ("natural" mortality). When the two different forces of mortality exist at the same time, only one can actually result in a given death, so it is customary to speak of "competing" sources of mortality or "competing risks."

If a population is "closed", that is, has no gains or losses from immigration, births, or emigration, the simplest model for mortality or survival is just the binomial distribution. If N represents initial population size, and p the proportion dying during some interval, then the expected number of deaths is simply Np , and the probability of various numbers of deaths is given by terms of the binomial distribution. Similarly if fractions can be assigned to exploitation and to natural mortality, then the model becomes a multinomial distribution.

In practice it is necessary to be cautious in using a simple model. Mortality seldom operates as a simple random variable as required by the model. In exploitation in particular, there is usually a tendency for events resulting in deaths to come in "bunches", violating the postulate of independence. Furthermore, mortality almost always varies seasonally, and exploitation usually is restricted to particular seasons. Also, when animals are tagged in groups, they are likely to stay clustered to some extent.

Another point is that we rarely know the total population size or the total numbers of deaths, and must estimate survival or mortality through some sampling process. In spite of all of this, the binomial model frequently provides a useful starting place (but usually not with total population as a parameter). Certain special populations do fulfill many of the requirements for the binomial and multinomial models. These are normally experimental populations being subjected to, e.g., chemicals to determine lethal effects of various dosages ("bioassay").

The principal methods for estimating survival or mortality may be categorized as follows:

I. From tagging or marking

A. Multiple recoveries of marked individuals

1. Capture-recapture methods. As mentioned above, estimates of survival or mortality are usually auxiliary to population estimation. These methods usually involve a series of captures and releases of individuals.

2. Multiple recoveries of individuals all tagged at the same time. Not a commonly used technique.

B. Single recovery of a marked individual. Normally the tag recovery comes coincident with death of the individual.

1. Recovery effort variable. For the most part, this pertains to exploitation, and the effort expended in hunting or fishing is that considered as devoted to recovery.

2. Constant recovery effort. Very often this is a necessary, but untested assumption. It may be reasonably acceptable where tag recovery depends on accidental death of individuals.

a. Exact time of death known. This is an area not much explored in biological studies, but well known industrially under the heading of "life-testing", where the survival of interest is that of some manufactured product (e.g., light-bulbs).

b. Time of death not known exactly. This is the category into which the bulk of methods in current use fall.

II. Other methods

A. Catch-effort methods. In commercial fisheries studies, much attention has been paid to assessing population size through measuring the rate of change in catch with accumulated effort. It may also be possible to estimate mortality at the same time.

B. Change-in-ratio methods. Also known as "dichotomy" or "survey-removal" methods, these depend on knowing the magnitude of a removal and observing a ratio before and after the removal. Survival may be concurrently estimated.

C. Catch-curves or kill curves. The age structure of animals taken in harvests may provide some notion of the year-to-year survival rate. It is usually necessary to assume constant population size and constant annual survival.

It is perhaps worthwhile to mention life tables here. In the original application, a life table served to record survival information down through the history of a particular group of individuals or to reflect the current state of survival rates in a population. In ecological studies there has been a

common tendency to arrange the returns from tagging in a life table, and this has led to some rather dubious calculations.

10.2 A single-recovery model

If a group of tagged individuals are released together and have a constant survival rate thereafter, then a fairly simple model suffices and permits various useful estimates. What happens when the assumption of constant survival does not hold is not known in any detail, but some rough guidance is available. It should be noted that the frequent practice of estimating annual survival from tag recoveries may not be importantly influenced by seasonal changes in survival, so long as the annual rate remains relatively constant. The other major use of single-recovery data is for relatively short-term periods (usually in seasons of exploitation) wherein it may reasonably be assumed that survival is approximately constant.

The most thorough analysis to date of the relevant model is one by Chapman and Robson (1960) which actually pertained to catch-curves, but also applies directly to tag recoveries, as pointed out by Paulik (1962). We will follow that analysis here. Let:

N_0 = initial population (number tagged or number alive at the time analysis starts -- this last assumption may be used to avoid effects of a higher early mortality).

s = constant survival rate (per unit of time)

λ = fraction of number currently alive that are caught and reported (in a given time unit). This may be the product of two constants, one being "catchability" and the second, reporting rate. Note that λ is used elsewhere here as a rate of population change, but is used in this case to correspond to the paper of Chapman and Robson (1960).

n_x = number of tagged individuals recovered in x th time unit. Confusion in notation can arise here, depending on how time units are numbered. One naturally speaks of the "first" time unit, but it is useful to let $x = 0, 1, 2, \dots$, and to suppose that n_0 represents those individuals who fail to survive through the first time unit. Under the above assumptions and definitions, it turns out that the probability of recovering a given individual in the x^{th} time unit follows a geometric probability distribution:

$$p_x = (1-s)s^x \quad x = 0, 1, 2, \dots \quad (10.1)$$

and the expected numbers of recoveries are:

$$E(n_x) = N_0 \lambda s^x \quad x = 0, 1, 2, \dots \quad (10.2)$$

This seems intuitively obvious, since if λ represents "catchability" the entire population is available in the first year, but only survivors are "at risk" in subsequent years.

The actual recoveries in any particular time unit constitute a binomial random variable, with probability given by (10.1) above, so that:

$$\Pr\{n_x \text{ individuals recovered}\} = \binom{N_0}{n_x} (\lambda s^x)^{n_x} (1-\lambda s^x)^{N_0-n_x} \quad (10.3)$$

If one wants to consider the distribution of recoveries over a set of time periods, then it is possible to consider the conditional distribution as being approximately a multinomial distribution. That is, for a fixed total sample $n = n_0 + n_1 + \dots + n_k$, the joint probability distribution

$$\Pr\{n_0, n_1, n_2, \dots, n_k \mid n\}$$

is a multinomial with the P_i corresponding to equation (10.1) above.

Chapman and Robson (1960) have shown that the best estimate for s is:

$$\hat{s} = \frac{T}{n+T-1} \quad (10.4)$$

where:

$$\begin{aligned} n &= n_0 + n_1 + n_2 + \dots + n_k \\ T &= n_1 + 2n_2 + 3n_3 + \dots + kn_k \end{aligned}$$

That is, n is just the total recoveries, and $n + T$ is effectively the "total years survived" after tagging. An estimate of the variance is:

$$v(\hat{s}) = \hat{s} \left(\hat{s} - \frac{T-1}{n+T-2} \right) \quad (10.5)$$

It is important to note that the method assumes that essentially all tagged individuals are dead when the analysis starts--a modified formula for partially complete recoveries will be described later.

The Chapman-Robson estimator is nearly identical to one generally attributed to D. Lack, but first studied mathematically by Haldane (1955). It is:

$$s = \frac{T}{n+T} \quad (10.6)$$

so the two estimators will differ only for small samples. The variance obtained by Haldane is (method of maximum likelihood):

$$v(s) = \frac{s(1-s)^2}{n} \quad (10.7)$$

a little algebra permits comparison of the two variance estimates:

$$\frac{\text{Chapman-Robson variance}}{\text{Lack-Haldane variance}} = \frac{n(n+T-1)}{(n-1)(n+T-2)} \quad (10.8)$$

which shows that the Lack-Haldane estimate will be a little smaller than the "best" estimate, but not importantly so if the number of recoveries is at all sizable.

A simple way to obtain the Lack-Haldane estimator is to use equation (10.2) and obtain expected values for n and T as follows:

$$E(n) = E(n_0) + E(n_1) + \dots = N_0\lambda(1 + s + s^2 + \dots)$$

and summing the infinite geometric series, we have:

$$E(n) = \frac{N_0\lambda}{1-s} \quad E(T) = E(\sum x n_x) = N_0\lambda(0 + s + 2s^2 + \dots) = \frac{N_0\lambda s}{(1-s)^2}$$

If the "mean age at death" is defined as:

$$\bar{x} = \frac{T}{n}$$

then substituting expected values for T and n :

$$\frac{E(T)}{E(n)} = \frac{s}{1-s} \quad \text{and} \quad \hat{s} = \frac{T}{n+T} \quad (10.9)$$

Example 10.1 Chapman-Robson survival estimate

As an example of the use of equations (10.4) and (10.5) we consider some data on band recoveries from adult Canada geese (Rutherford, 1970:27) banded as adults in the Arkansas Valley of Colorado in the 1950-51 season. A total of 344 geese were banded, with recoveries in subsequent years as follows:

<u>Year of recovery following banding</u>	<u>Number of recoveries</u>
x	n_x
0	45
1	23
2	19
3	10
4	7
5	12
6	5
7	3
8	3
9	2
10	2
11	0
12	1
13	1
14	0
15	1
Total recoveries (n)	134

$$T = 23 + 2(19) + 3(10) + 4(7) + 5(12) + \dots + 15(1) = 332$$

$$\hat{s} = \frac{T}{n+T-1} = \frac{332}{134+332-1} = 0.714$$

$$v(\hat{s}) = \hat{s} \left(\hat{s} - \frac{T-1}{n+T-2} \right) = \frac{332}{465} \left(\frac{332}{465} - \frac{331}{464} \right) = 0.0004401$$

so that the standard error of the estimated variance is $(.0004401)^{1/2} = 0.021$ and approximate 95 percent confidence limits on the survival estimate are given by $.714 \pm 2(0.021)$ or 0.672 to 0.756.

10.3 Further estimators

One very simple procedure is to consider the fraction n_0/n , which represents the fraction dying in the first time interval and thus estimates a mortality rate. Conversely, survival can be estimated as:

$$s_0' = 1 - \frac{n_0}{n} = \frac{n-n_0}{n} = \frac{n_1 + n_2 + \dots + n_k}{n_0 + n_1 + n_2 + n_k} \quad (10.10)$$

where n_k represents the oldest group from which any recoveries are obtained (we again assume all tagged individuals are dead before calculations begin, so no new recoveries can be expected). The above estimate is attributed to Heincke (1913) and a minor modification (dropping n_k from the denominator of the right-hand expression) yields an estimator proposed by C. H. N. Jackson (1939).

The Heincke estimate can readily be treated statistically, inasmuch as, for a given total number of recoveries (n), the probability that a particular individual is recovered in the first year is just $p = 1-s$ so the distribution of the number of first year recoveries (n_0) is binomial:

$$\Pr\{n_0 \text{ recoveries in 1st year}\} = \binom{n}{n_0} (1-s)^{n_0} s^{n-n_0} \quad (10.11)$$

so that the expected number is:

$$E(n_0) = n(1-s)$$

which gives the estimate of equation (10.10). The usual binomial estimate of a variance of a proportion then gives a variance as:

$$v(s_0') = \frac{s(1-s)}{n} \quad (10.12)$$

The importance of this result arises by comparison with equation (10.7). The ratio of the two variances is:

$$\frac{v(s_0')}{v(\hat{s})} = \frac{1}{1-s} \quad (10.13)$$

which amounts to a rather remarkable state of affairs, if one stops to consider the substantial costs involved in tag and recovery studies. Supposing annual survival to be on the order of 50 per cent, equation (10.13) shows that the

variance obtained from the Heincke method will be about double that obtained from the preferred methods (Chapman-Robson or Lack-Haldane). This then says that choice of the wrong equation amounts effectively to throwing away half of one's hard earned information, and the older literature abounds with examples where the Heincke estimate is used when the better estimate is available.

10.4 Analysis of incomplete data

In many applications it is necessary to estimate survival from only part of the data. Perhaps the most frequent case is when the investigator cannot afford to wait until nearly all recoveries are back, but must proceed with results from the early recovery periods. Calculations depend on an equation analogous to (10.9) which is:

$$\frac{T}{n} = \frac{s}{1-s} - \frac{(k+1)s^{k+1}}{1-s^{k+1}} \quad (10.14)$$

in which k represents the last recoveries available, and

$$n = n_0 + n_1 + n_2 + \dots + n_k$$

$$T = n_1 + 2n_2 + \dots + kn_k$$

If only two years of recoveries are available $k=1$, and

$$\frac{T}{n} = \frac{n_1}{n_0 + n_1} = \frac{s}{1-s} - \frac{2s^2}{1-s^2} = \frac{s}{1+s}$$

and this gives an estimate of s as

$$\hat{s} = \frac{n_1}{n_0} \quad (10.15)$$

which is what one might intuitively expect--that recoveries in two successive years differ only by s (or this can be obtained by use of equation (10.2)).

If $k=2$ a somewhat more complicated equation can be obtained for estimating s directly (Robson and Chapman, 1961:188), but for larger values of k one must resort to trial and error (iterative) methods or to tables (Robson and Chapman, 1961). Hand calculations are not too awkward, and a starting value can quickly be obtained from equation (10.15).

An estimate of the variance of the estimate is obtained from:

$$v(s^*) = \frac{\frac{1}{n}}{\frac{1}{s(1-s)^2} - \frac{(k+1)^2 s^{k-1}}{(1-s^{k+1})^2}} \quad (10.16)$$

An examination of equation (10.16) shows that as k becomes large, the term involving k tends to become very small, so that the equation reduces to equation (10.7), as it should. With these results, the investigator can determine approximately what the effect will be of estimating survival from only part of the recoveries, as opposed to waiting until all tagged individuals are quite

surely dead. If s is not quite small, there may be a substantial difference in the variability between the two situations. An estimate of the reduction in variability resulting from using all of the recoveries can be obtained by dividing equation (10.7) by (10.16):

$$\frac{\hat{v}(s)}{v(s^*)} \doteq 1 - \frac{(k+1)^2 s^k (1-s)^2}{(1-s^{k+1})^2} \quad (10.17)$$

Some selected values are as follows:

s	k						
	1	2	3	4	6	8	10
0.10	0.669	0.927	0.987	0.998	1.0	1.0	1.0
0.20	0.444	0.766	0.918	0.974	0.998	1.0	1.0
0.30	0.290	0.581	0.785	0.900	0.982	0.997	1.0
0.50	0.111	0.265	0.431	0.584	0.806	0.921	0.970
0.70	0.031	0.081	0.145	0.219	0.384	0.544	0.680
0.90	0.003	0.007	0.014	0.022	0.043	0.071	0.104

Inspection of the above table shows the advantage of basing survival estimates on as many periods of recovery as possible.

Example 10.2 Analysis of incomplete data

For an example of "segment" calculations (Eq. 10.14), we again use Rutherford's data, and suppose that only the first 4 classes are available for estimation. Then:

$$n = 45+23+19+10=97$$

$$T = 23+2(19)+3(10)=91$$

$$\bar{x} = \frac{T}{n} = \frac{91}{97} = 0.9381$$

and we enter the tables given by Robson and Chapman (1961) and Seber (1982:584). In the column corresponding to $k = 3$ (4 classes) and find 0.938 to correspond to $s = 0.62$. A convenient alternative is to use SOLVER in EXCEL. A variance estimate (Eq. 10.16) is:

$$\hat{v}(s) = \frac{\frac{1}{97}}{\frac{1}{0.62(0.38)^2} - \frac{16(0.62)^2}{(1-(0.62)^4)^2}} = 0.003816$$

which is much larger than the variance for the complete set of data (0.00044).

10.5 Testing assumptions

The assumption of constant survival and recovery rates is one likely to be violated in many actual situations. We will look briefly at some of the effects of changing rates in a later section. Here we consider ways of looking

for bias. Probably the most useful technique is to have a thorough knowledge of the circumstances in which the data were obtained, and to examine the prospects for any major change that might influence either survival or recovery rates. Another considerable help is to have a rather long series of comparable data to examine for any persistent trends. Most bird-banding experience, for example, shows that juveniles have a markedly lower survival than do adults. Hence one either tries to band mostly adults or deals with juveniles separately.

Some statistical tests may also be considered. One is just the ordinary chi-square goodness-of-fit test, wherein the expected values are computed from $n_x = n\hat{s}^x$ and compared with the observations (n_0, n_1, n_2, \dots) . Substantial departures from constant survival may show up in this test, but it isn't very sensitive, and failure to obtain a significant chi-square needs to be regarded as a necessary condition for use of the estimate but should not be regarded as any substantial evidence against bias.

Fortunately there is a test for one of the major sources of error, i.e., a difference in survival rate between the first recovery - class and subsequent classes. It thus serves very well to examine the effects of, for example, tagging young animals. The test is due to Chapman and Robson (1960) and is a chi-square with one degree of freedom:

$$\chi^2 = \frac{(\hat{s} - s'_0)^2}{\frac{T(T-1)(n-1)}{n(n+T-1)^2(n+T-2)}} \quad (10.18)$$

where \hat{s} is obtained from the Chapman-Robson estimate of equation (10.4) and s'_0 is the Heinke estimate of equation (10.10).

Example 10.3 Testing compatibility of first recoveries

The chi-square test for compatibility of the first recovery-class with the remaining classes (Eq. 10.18) is readily illustrated on the data of Example 10.1. It requires Heinke's estimate of survival (Eq. 10.10):

$$s'_0 = \frac{n - n_0}{n} = \frac{134 - 45}{134} = 0.6642$$

The chi-square test then is:

$$\chi^2 = \frac{(0.714 - 0.6642)^2}{\frac{332(331)133}{134(465)^2 464}} = \frac{0.00248}{0.00109} = 2.28$$

which is well short of the 95 percent significance level of chi-square.

The above test is "two-tailed", that is, operates so as to protect one against either over- or under-representation of the first age-class. If the investigator has good grounds for expecting a departure to be in only one direction (e.g., if mesh size of nets is known to be the critical factor, or greater vulnerability of young game animals), then a one-tailed test may be in order. This can be obtained just by

considering the ratio of $\hat{s} - s'_0$ to the square root of the denominator of Eq. (10.18), and referring to tables of the normal distribution for significance levels (the critical level for a one-tailed test at the 95 percent level of significance is 1.65 rather than the 1.96 used for two-tailed tests). If there were grounds to support doing a one-tailed test here, one could calculate:

$$z = \frac{0.714 - 0.6642}{(0.00109)^{1/2}} = \frac{0.0498}{0.033} = 1.5$$

which is again short of the 95 percent level (1.65).

The above are "large-sample" tests. Chapman and Robson (1960:364) point out that the appropriate "exact" distribution is the hypergeometric, but also that for all practical purposes one may substitute a table of binomial confidence limits, using the statistic $T=(n_1+n_2+n_3+...)$ as "sample size" and $T-n+n_0$ as "number of successes". If the selected confidence interval (corresponding to chosen level of significance) excludes s then one rejects the hypothesis of compatibility of survival of the first age group with that of the remaining recoveries. If the above tests suggest incompatibility of the first recovery class, one simply drops that class and recodes his data, so that n_0 is now the first recovery-class retained, n_1 the second, and so on. Calculations of survival and variance, etc., proceed just as before, but its a wise precaution to repeat the above testing procedure to make sure that the new initial recovery group is compatible with the remainder.

Example 10.4 Combining years

In Example 10.1, a large number of geese were banded in one year, resulting in a substantial number of returns and quite a precise estimate of the annual survival rate. Often one must deal with much smaller samples, and the question of combining a series of estimates may then come up. When all, or, virtually all, of the recoveries are in, the procedure is very simple -- just add up the recoveries, and proceed as before (Chapman and Robson 1960:357). Some data on adult Canada geese banded in Washington (Hanson and Eberhardt, 1971) provide an example:

Year banded	Number banded	<u>Recoveries during subsequent years</u>										Recov. Tot. Rate		
		n_0	n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	n_9			
1950	43	1	2	1			1				1	6	.14	
1951	40	2	2	2			1					7	.18	
1952	133	14	8	7	6	1		2	1	1		40	.30	
1953	13				1							1	.08	
1954	25	3	2	2	1							8	.32	
1955	30	3	2				1					6	.20	
1956	49	1	2	3	1			1		1		9	.18	
1957	38		2	1	2		2	1				8	.21	
1958	45	5	2			2	2		2			13	.29	
1959	34		1	1	1	1				1	1	6	.18	
1960	18	2						1				3	.17	
-----		468	31	23	17	12	4	7	5	3	3	1	107	.23

$$T = 23 + 2(17) + 3(12) + 4(4) + 5(7) + 6(5) + 7(3) + 8(3) + 9(1) = 228$$

$$\hat{s} = \frac{T}{n + T - 1} = \frac{228}{334} = 0.683$$

$$v(\hat{s}) = \frac{228}{334} \left[\frac{228}{334} - \frac{227}{333} \right] = 0.0006506$$

We thus obtain an estimate of annual survival of $.684 \pm 2(0.0006506)^{1/2}$ or confidence limits of 0.634 - 0.735 on the combined estimate.

When several years are combined, as above, it seems prudent to examine the band recovery rates to see whether there is any evidence to suggest a change that might invalidate the underlying assumptions (e.g., a change in reporting rates). One simple procedure is just to do a chi-square test against the overall recovery rate. The calculations here are:

<u>Year</u>	<u>Number banded</u>	<u>Number recovered</u> (q _i)	<u>Recovery rate</u> (p _i)
1950	43	6	0.1395
1951	40	7	0.1750
1952	133	40	0.3008
1953	13	1	0.0769
1954	25	8	0.3200
1955	30	6	0.2000
1956	49	9	0.1837
1957	38	8	0.2105
1958	45	13	0.2889
1959	34	6	0.1765
1960	18	3	0.1667
	468	107 (A)	0.2035 = \bar{p}

and a handy formula for chi-square is (Cochran and Snedecor, 1967:240):

$$\chi^2 = \frac{\sum p_i a_i - \bar{p}A}{\bar{p}\bar{q}} = \frac{26.5809 - 21.7738}{0.1621} = 29.66$$

with 10 degrees of freedom, indicating strong evidence of a departure from chance fluctuations in proportion recovered.

10.6 Life tables

The original and major use of a life table is to provide a convenient summary of survivorship data over the history of a given group of animals, all born at nearly the same time (a "cohort" and hence a cohort table) or to provide a cross-section of current mortality experience in a population (a "current" life table). As such, there should be no objection to life tables. Unhappily many such tables have been constructed from band-recovery and

age structure data in ways that may greatly reduce the value of the data. Most of this section will be devoted to life tables as constructed from band recovery data.

The basic structure of the table is quite simple, although a variety of modifications and extensions has to be derived for various special purposes. The structure of a "cohort" table is obtained as follows. Most life tables start out with an arbitrary number of newly born animals (usually 1000, or 10,000), and have 4 main columns. The first (x) column lists age at the beginning of the time interval (normally a year), the second (l_x) gives the number alive at the beginning of the year, while the third (d_x) gives the number of deaths during the year, and the fourth column (q_x) represents the mortality rate. Entries in a particular table might thus be:

x	l_x	d_x	q_x
0	1000	620	0.620
1	380	167	0.439
2	213	85	0.399
3	128	51	0.398
4	77	32	0.416
5	45	19	0.422
6	26	14	0.538
7	12	6	0.500
8	6	4	0.667
9	2	2	1.000

Since the table begins with an arbitrary number of animals, the only really important data are those contained in the q_x column, which summarizes the available information on mortality by age. Some cohort tables do start out with the actual number of individuals in a real group, and thus provide more information than does the "standard" table. "Current" life tables are usually constructed in exactly the same way, but depend on different sources of data (i.e., on current mortality experience of a population).

Construction of a cohort life table (also known as a "time specific" life table) from the recovery of marked animals at their death rests on the intuitively plausible concept that the recoveries constitute a representative sample of the fate of the entire group initially marked. The method in common use is to sum up the total recoveries (n) over the years, and to regard this total as a cohort, or "number initially at risk". Thus n may be the first entry in the l_x column -- unfortunately many workers convert it to 1,000 or 10,000, making it impossible (without additional information) to reconstruct the actual data from their table. Those individuals recovered in the first year are then regarded as mortalities (appearing in the d_x column) and the first year mortality rate is estimated as n_0/n . In the second year it is assumed that there are $n-n_0$ survivors (l_1 entry) and the recorded deaths in that year (n_1) appear in the d_1 position and become the numerator in the mortality (q_1) estimate:

$$q_1 = \frac{n_1}{n - n_0}$$

continuing the process yields mortality rates for each year up to the last recoveries (n_k) whereupon the last year's mortality rate becomes $n_k/n_k = 1$.

It thus turns out that the first entry in the table corresponds to the Heincke estimate of equation (10.10), and each successive entry corresponds to the same estimate, with previous recovery-class disregarded. Thus the estimates are not what they have been claimed to be, that is, age-specific mortality rates. They are instead estimates of a constant rate for the whole set of recovery data, and the supposed life table is not at all what it is normally expected to be. Furthermore, as we have seen from equation (10.13) using the Heincke estimate amounts to discarding a sizable fraction of the available data. Some workers did "combine" the several estimates by weighting each q_x estimate by the apparent sample size (n_x), giving:

$$\frac{\frac{n_0}{n} n + \frac{n_1}{n-n_0} (n-n_0) + \dots}{n_0 + 2n_1 + 3n_2 + \dots} = \frac{n}{n + T} \quad (10.19)$$

which is the Lack-Haldane estimate, and thus may well be the only really useful estimate in the whole table. If the actual numbers recovered are entered in the table (rather than converting to 1000 or 10,000 for the l_0 entry) then it is feasible to go back and use the chi-square test of equation (10.18) on the data. Sometimes the total number of recoveries is given separately, so that it is possible to reconstruct the original data from the table, even though l_0 is 1000 or 10,000. At worst, one can look down the q_x column to see whether the apparent mortality rates do change -- but this is seldom a very helpful practice in view of the substantial variability of the Heincke estimates.

Perhaps the best summary statement about cohort tables based on recovery data is that they should not be constructed in the first place. If the basic assumption is fulfilled (constant survival) then each entry estimates exactly the same quantity! If survival rates are not constant over the years, then the structure of the various estimates can be illustrated by writing out the value for the first one:

$$q_0 = \frac{n_0}{n - n_0} = \frac{pN(1 - s_1)}{pN(1-s_2)s_1 + pN(1-s_3)s_1s_2 + \dots + pN(1-s_{k+1})s_1s_2\dots s_k} \quad (10.20)$$

where p = recovery rate, N = population tagged, and the s_i are annual survival rates. Thus only if the s_i are all equal to s , do we get conditions for a valid estimate.

When one uses data based on a group of animals all marked at the same time, the applicable life table concept is indeed that of a cohort. However, the data are nonetheless at times used to construct a current life-table (also called a "dynamic" table, and perhaps more accurately so in the present context).

The usual practice seems to be to regard the ratios of successive recovery- classes as estimators of current survival rates, so that:

$$q_x = \frac{n_x}{n_{x-1}}$$

and using the notion of equation (10.20):

$$q_x = 1 - \frac{pN(1-s_x)s_1s_2\dots s_{x-1}}{pN(1-s_{x-1})s_1s_2\dots s_{x-2}} = 1 - \frac{(1-s_x)s_{x-1}}{1-s_{x-1}} \quad (10.21)$$

so that, if survival is constant in two successive years, the method does provide a valid estimate. The existence of changing survival rates evidently will result in a bias. Perhaps a more serious drawback is that the samples usually obtained are so small as to result in substantial variability in the year to year estimates. In the majority of cases, it is likely that only the first two or three recovery groups are large enough to justify much confidence in the mortality estimates.

One technique used to reduce the effect of small samples is to combine the recoveries from a number of years of marking, producing what is commonly called a "composite" life table. Sometimes this is done with partially incomplete data. That is, there may be one (or more) sets where all of the recoveries are available, plus several sets where banding was done recently enough so that there is likely to be a number of marked individuals yet alive. In such circumstances an adjustment may be introduced to compensate for the incomplete nature of the data. The adjustment takes the form of computing for each year a "number of marked individuals available." The trick is simply to add up the total number actually banded during the period of interest. This total is regarded as the number "available", and the first year recoveries are divided by that number to generate a " d_0 " class. If the most recent year of banding did not contribute to the second year recoveries then the number banded in that year is subtracted from the total banded, and the remainder is used as a divisor to generate " d_1 " from the second year recovery. The process continues until the number banded in the first year is used as divisor for the last recovery class. Some better methods for calculating survival rates from incomplete data are available, and will be described below.

In summary, many of the methods for generating life tables from banding data are of very dubious utility. The "cohort" method does not do what it is supposed to do at all, and the "current" method suffers seriously from small samples. There is, of course, no reason to object to the construction of life tables from data obtained by recovery of banded animals. The point is that survival estimates should be obtained by the best methods available and then used to construct a life table. The methods already described are quite simple and should normally be less trouble to use than the supposed life table method.

10.7 Catch-curves and kill-curves

It was mentioned above that the Chapman-Robson analysis was originally developed for catch curves but applies equally well for banding data. It thus follows that there is little need to recapitulate the methods of estimation already described. It is necessary, though, to note some pitfalls for the unwary and changes in assumptions. In band-recovery studies, one usually has a known population (the total number banded) under study, and obtains recoveries over a number of time intervals. The essential assumptions are that survival remains constant (or nearly so) and that recovery rates of dead animals also remain constant. On the other hand catch-curves or kill-curves usually depend on an age-classification of a group of animals taken nearly simultaneously (by fishing or hunting) and thus born over a number of years. Constant survival is again a necessary assumption, while the

assumption of constant recovery rates is now replaced by one of equal "catchability" (vulnerability to the harvest method) for each age group. Unfortunately a third assumption has also to be added in that it is also required that each age-group initially exist in equal numbers.

In fisheries studies there is often a problem with the second assumption - - younger fish may be less readily taken by the fishing methods (e.g., mesh size of nets). This problem is usually dealt with by restricting the analysis to start with the first age group considered to be "fully vulnerable to the gear" (which can be checked out with the chi-square test of equation (10.18)). Most fish have a tremendously large reproductive potential so that it seems that the link between size of mature stock and numbers of new recruits may at least be tenable for several years in a row. This is not to deny, of course, the importance of spawner-recruit relationships. The major point is that catch-curves seem to be rather more useful in fisheries work than do kill-curves in studies of hunted populations, and it seems likely that the chief difficulty may lie in the third assumption. Most game populations do not appear to have a sufficiently constant recruitment to permit use of survival estimates from kill-curves for more than very rough estimates.

10.8 Combining estimates from several years

In many cases the available recovery data will come from a number of years of banding. There then may arise the need to combine several sets of data to produce a single survival estimate. As always, one should not blindly combine data from different years but should first examine the individual estimates (along with the standard errors), and review the prospects of any major change between the years in question. All too often recoveries from a single year's banding constitute a very small sample, and they may have to be combined with other years.

The simplest situation arises when there are complete sets of recoveries for a number of years. One then simply adds together the corresponding year-classes (i.e., numbers of recoveries 1,2,3,... years after banding) and uses equation (10.6) to estimate survival, and (10.7) to estimate variance of the estimate.

When the data are incomplete, as discussed in Section 10.4, the calculations become somewhat more tedious. The usual set of data may be represented by the following table:

Maximum age at recovery	<u>"Age" at recovery</u>				Totals
	1	2	3	x	
k	n_k
3	d_{13}	d_{23}	d_{33}		n_3
2	d_{12}	d_{22}			n_2
1	d_{11}				n_1
<hr/>					
Totals	d_1	d_2	d_3	d_x	$\Sigma n_k = \Sigma d_x$

Here the rows are sets of recoveries from a given year of banding, with the most recent year at the bottom -- hence the maximum "age" at recovery is one year. The row totals (n_k) are thus the total number recovered from a given year of banding, while the column totals reflect the total numbers that had been "at large" for a given number (x) of years. "Age" is necessarily reckoned from banding, so that animals banded as adults may in fact be appreciable older than implied by the terminology here.

Actual survival estimates from the above table are obtained from the following equations (Gilmartin et al. 1993):

$$\sum_{j=0}^k \frac{T_j}{n_j} = \sum_{j=0}^k \frac{1}{1-s} - \frac{(k_j+1)s^{k_j+1}}{(1-s^{k_j+1})} \quad (10.22)$$

$$\text{Var}(s) = \left[\sum_{j=0}^k \left\{ \frac{n_j}{s(1-s)^2} - \frac{(k_j+1)^2 s^{k_j-1} n_j}{(1-s^{k_j+1})^2} \right\} \right]^{-1} \quad (10.23)$$

where T_k is computed for each row in just the same way that T was obtained for equation (10.6), i.e.;

$$T_3 = d_{13} + 2d_{23} + 3d_{33}$$

The solution of equation (10.22) depends on iterative methods just as it did for equation (10.14), to which it is closely related -- in fact (10.22) may be obtained by summing equation (10.14) over the total number of years of recovery (i.e., sum on k). Similarly, equation (10.23), stems from summing the equation (10.16).

The method described above was obtained by Haldane (1955) and is fairly well-known as "Haldane method." It may be repeated that a prudent investigator will use equations (10.14) and (10.16) to investigate at least part of the individual years before settling on a combined estimate. If no banding was done in one or more years it is still feasible to use equations (10.22) and (10.23); one just enters zeroes for the years of no banding. Gilmartin et al.(1993) used monte carlo methods to study the variance equation and found that it gave confidence limits that were too wide, so resorted to bootstrapping to estimate confidence limits.

Example 10.5 An example of Haldane's method

For an example of combining incomplete data, we use some band returns data on Canada geese banded in Washington (Hanson and Eberhardt, 1971) as adults:

Year <u>banded</u>	No. <u>banded</u>	Maximum age <u>at recovery</u>	"Age at recovery"							Tot.
			0	1	2	3	4	5	6	
1961	14	6	3	0	2	1	0	0	0	6
1962	15	5	0	2	0	0	1	0		3
1963	28	4	4	4	1	1	0			10
1964	45	3	3	2	2	6				13
1965	10	2	0	1	0					1
1966	0	1	0	0						0
1967	15	0	2							2
			12	9	5	8	1	0	0	35

We first compute T_k and n_k for each row in just the same manner as T was computed before:

$$\begin{aligned}
 T_6 &= 2(2) + 3(1) & = 7 & n_6 &= 6 \\
 T_5 &= 2 + 4(1) & = 6 & n_5 &= 3 \\
 T_4 &= 4 + 2(1) + 3(1) & = 9 & n_4 &= 10 \\
 T_3 &= 2 + 2(2) + 3(6) & = 24 & n_3 &= 13 \\
 T_2 &= 1 & & n_2 &= 1 \\
 T_1 &= 0 & & n_1 &= 0 \\
 T_0 &= 0 & & n_0 &= 0
 \end{aligned}$$

and s is estimated by solving equation (10.22) by trial and error. One sets up the table above in EXCEL and computes the components of equation (10.22) in two columns, sums these and varies s until the two column sums are equal. If there are several examples to compute, the solution might be obtained with SOLVER.

10.9 Regression methods

Recalling equation (10.2):

$$E(n_x) = N_0 \lambda s^x$$

If we take logarithms:

$$\log E(n_x) = \log(N_0 \lambda) + x \log s \quad (10.24)$$

and ordinary linear regression methods might be used to fit observed data and to estimate s as the slope of the regression of $\log n_x$ on years since banding. Such a procedure has the additional merit that one can examine the plotted data for any evidence of non-linearity, and hence for prospects of a change in survival with time.

A disadvantage of the regression method is that band recovery data do not conform to the requirements for regression estimation. Chapman and Robson (1960) suggest an improvement in some cases may be obtained by using

$$\log n_x - \frac{1}{n_x + 1}$$

as the independent variable (rather than just $\log n_x$). They further recommend dropping all observations beyond the point where n_x is less than 5.

It is important to realize that the regression equation estimates $\log s$, and not s . If natural logarithms are used, the quantity

$$i = \log_e s$$

is the "instantaneous" rate of survival, and is rather widely useful in fisheries work as such. A natural temptation is to antilog the estimate of $\log s$ for a direct estimate of s ; unfortunately this yields a biased estimate and there doesn't seem to be much to do about it unless the samples are large. But with quite large samples the bias may not be important.

Paulik (1963) developed another regression-like method based on the fact that recoveries in a number of years approximately follow the multinomial distribution (this was mentioned in Section 10.2). His results seem to offer an improvement over the simple regression estimate (10.24). His estimate is:

$$\log \hat{s} = \frac{\sum j n_j y_j - [(\sum n_j y_j)(\sum j n_j) / \sum n_j]}{\sum j^2 n_j - [(\sum j n_j)^2 / \sum n_j]} \quad (10.25)$$

where $j = 1, 2, 3, \dots$ denoted the year of recovery (with the first year's recoveries now numbered as n_1 rather than n_0 as before). Also;

$$y_j = \log_e \left(\frac{n_j}{N} \right)$$

which is the natural logarithm of the proportion (of N banded) in the j^{th} year. Paulik recommended deleting observations beyond the point where $n_j = 10$.

Since this method also estimates the logarithm of s and not s directly, it may be best applied in circumstances where interest is not so much in estimates of annual survival, but is rather directed towards components of that survival rate, expressed in terms of the instantaneous rates. That is, fisheries workers customarily write

$$s = e^{-(F+M)}$$

where F = mortality rate (instantaneous) from fishing and M = mortality rate (instantaneous) from other causes (the "natural" mortality rate), and the regression methods thus estimate the quantity $F+M$ directly. Any auxiliary information on the intensity of exploitation during the time periods when the data were collected may potentially permit an estimate of the relative contribution of F to total mortality.

One other reason for considering estimates of $\log(s)$ rather than s is that $\log(s)$ may be roughly normally distributed. This may be a considerable advantage in making tests of comparisons or in setting confidence limits. If

$\log(s)$ is normally distributed, then the antilog $e^{\log s}$ will be lognormally distributed (and thus have a skewed distribution -- awkward for many purposes.)

10.10 Effect of changes in survival

We recall that one of the key assumptions in the analysis of both band recovery and kill-curve data is that the survival rate remain constant over the period of study. In practice, survival undoubtedly does vary from year to year. If the annual fluctuation are not large and are more or less "chance" events, presumably the estimates may provide some sort of average value. Very often, though changes in survival rate may persist for several years. This is especially true for exploited populations, where shifts in management regulations may induce changes in survival.

Not a great deal is known about the effect of persistent changes on estimation of survival, but some indication of trend can be supplied. We first consider band recoveries, and assume the recoveries come from an exploited population (tags are recovered from the harvested animals). Suppose fishing or hunting (F_x) and natural (M_x) mortality operate together over the year (and remain constant during the year) but change from year to year. If we let

$$f_x = \frac{F_x}{F_x + M_x}$$

then the estimate of survival from recoveries is approximately:

$$\begin{aligned} \frac{E(T)}{E(n+T)} &= \\ \frac{E(n_1) + 2E(n_2) + 3E(n_3) + \dots}{E(n_0) + 2E(n_1) + 3E(n_2) + \dots} &= \qquad (10.27) \\ &= \frac{f_1(1-s_1)s_0 + 2f_2(1-s_2)s_0s_1 + 3f_3(1-s_3)s_0s_1s_2 + \dots}{f_0(1-s_0) + 2f_1(1-s_1)s_0 + 3f_2(1-s_2)s_0s_1 + 4f_3(1-s_3)s_0s_1s_2 + \dots} \\ &= s_0 \frac{f_1(1-s_1) + 2f_2(1-s_2)s_1 + \dots}{f_0(1-s_0) + 2f_1(1-s_1)s_0 + \dots} \end{aligned}$$

Thus if the samples are large enough to permit the operation of equation (10.27) -- one would really prefer to find $E(T/n+T)$ which is not the same as what we have above -- it appears that the "dominant" quantity in the estimate is s_0 , or survival in the first year. One might expect that result by inspection of the recovery data, since the first two classes (n_0 and n_1) are necessarily expected to be the largest. The effect of some postulated change in survival rates can be determined by considering the sequence of survival rates in equation (10.27). For example, suppose survival has been increasing so that $s_0 > s_1 > s_2$. Then $(1-s_1) > (1-s_0)$ and $(1-s_2) > (1-s_1)$, and so on. Consequently the overall estimate will be less than s_0 . One might thus write:

$$\hat{s} = s_0(\text{bias})$$

Any quantitative appraisal of the bias term will evidently depend on a numerical evaluation of equation (10.27) for which one needs to know survival rates. Furthermore, changes in survival are likely also to affect the f_x terms, inasmuch as the fraction taken by exploitation will no doubt also change.

While the situation above is rather complex from an analytical standpoint, the circumstances are not difficult to mimic via computer simulation. A limited effort along those lines (Eberhardt, 1972) led to the following set of recommendations:

(1) The Chapman-Robson (Lack) equation for estimation of survival from banding recoveries is clearly the best of those studied. The other methods have variances which are usually at least twice as large.

(2) Both Chapman-Robson and Heincke estimates should always be computed and used to test the compatibility of the first year of returns actually used with those of later years. If the test is significant, one may simply drop the first year, and proceed with the remaining data.

(3) In the case of incomplete data (banded individuals still alive at analysis) or a prior evidence of a marked change in survival rates, the segment method is available, but variability of the results is discouragingly large unless the data are nearly complete.

(4) Combining results of a series of years of banding may take two routes.

(a) If it appears that there is a marked change in survival, one probably has to depend on a year-to-year analysis--this is an area needing further investigation.

(b) If it appears that survival rates have not changed appreciably, then those years for which complete data are available can simply be added together and analyzed with the Lack formula (justification for this statement appears in the appendix to Eberhardt (1972)). If part of the data is incomplete, then the Haldane formula is called for.

(5) It was provisionally suggested that, even in the face of changing rates, the Chapman-Robson and/or Heincke methods give estimates of first-year survival which likely are not very much biased. An exception is the case where rates change in the second year, as may be expected if first-year recoveries of juvenile banding are used. However, the Chapman-Robson test provides good protection against that eventuality.

(6) If all else fails, and samples are indeed large, one might look at ratios of successive groups of returns (i.e., use the "current" life table) and obtain an average survival rate. Even so, it is probably advisable to simply drop the smaller frequency groups (say below 10 returns).

(7) There seems to be little reason to continue the practice of life-table analysis of banding data. Under the requisite assumptions, all of the information in a "cohort" table appears in its first line, and in the form of a

Heincke estimate, which has a variance appreciably larger than that of the preferred method. Use of the method thus amounts to throwing away data that are often rather expensive to obtain. As suggested above (6), it may be necessary to examine data on a year-to-year basis--essentially in a "current" life table, but the variances of such estimates are discouragingly large. The evidence from this study suggests that the Chapman-Robson (Lack) method will often give a useful estimate of first-year survival, and that very likely will be all that can effectively be salvaged.

Much the same sort of analysis can be carried out for kill-curves. In analogy to equation (10.2):

$$E(n_x) = \lambda N s_1 s_2 \dots s_x \dots$$

$$\frac{E(T)}{E(n+T)} = \frac{\lambda N (s_1 + 2s_1 s_2 + 3s_1 s_2 s_3 + \dots)}{\lambda N (1 + 2s_1 + 3s_1 s_2 + 4s_1 s_2 s_3 + \dots)} \quad (10.28)$$

$$= s_1 \left[\frac{1 + 2s_2 + 3s_2 s_3 + \dots}{1 + 2s_1 + 3s_1 s_2 + \dots} \right]$$

and we can again consider changes like $s_1 > s_2 > s_3$ and their qualitative effect on a bias term:

$$\hat{s} = s_1(\text{bias})$$

It is perhaps worth noting what happens if survival remains constant for several years, e.g.:

$$s_1 = s_2 = s_3 > s_4 = s_5 = s_6$$

if we let s_1 represent the recent series and s_2 the older series of rates, we have

$$\frac{E(T)}{E(n+T)} = s_1 \frac{1 + 2s_1 + 3s_1^2 + 4s_1^3 s_2 + \dots}{1 + 2s_1 + 3s_1^2 + 4s_1^3 + \dots}$$

so that, unless survival is unusually high so that the later terms have an influence, it seems evident that the estimate will be very close to s_1 , apart from the effects of sampling variation.

A similar analysis can be carried out for the Heincke estimate (equation (10.10)):

$$\frac{E(n-n_0)}{E(n)} = s_1 \frac{1 + s_2 + s_2 s_3 + \dots}{1 + s_1 + s_1 s_2 + \dots} \quad (10.29)$$

A crucial assumption in the above is that recruitment remains constant, i.e., that we can consider N to be a constant in developing equations like (10.28). As we have already remarked, this is frequently an unwarranted assumption, and this largely negates direct use of the methods thus far described for kill-curves of many populations.

10.11 Use of telemetry to estimate survival

Radiotelemetry offers substantial advantages over depending on tag returns for estimating survival. Conceptually, the actual time of death will be

known and it may be possible to determine the cause of death for non-harvest mortalities. Under such circumstances a simple binomial model may suffice for analysis. Very likely the most widely known analysis is that of Heisy and Fuller (1985). They proposed estimating a daily survival rate as:

$$s_i = \frac{x_i - y_i}{x_i} = 1 - \frac{y_i}{x_i} \quad (10.30)$$

Where x_i is the total number of transmitter-days, and y_i is the total number of deaths in some interval, i . An important precaution is to select an interval in which conditions are likely to be relatively constant. It may be necessary to consider several such intervals that span the total period of interest. For a single such interval the survival rate may be estimated as:

$$S_i = s_i^{L_i} = \left(1 - \frac{y_i}{x_i}\right)^{L_i} \quad (10.31)$$

Combining rates for several intervals is then accomplished by multiplying the rates for the several (I) intervals

$$S^* = \prod_{i=1}^I S_i \quad (10.32)$$

10.12 Kaplan-Meier estimates of survival

An important problem in survival estimation is lack of knowledge of the fate of individuals. The difficulty is severe when tags or bands are recovered by sampling as in hunting and fishing, but is also important in the capture-recapture methods. Conceptually, it should be a minor concern in radiotelemetry studies where one might suppose there is virtually absolute knowledge of the fate of individuals. However, this is not always the case. Radios do fail on occasion, as do the attachments, and the lost radiotag may not always be promptly recovered. Also, radiotagged individuals may emigrate out of the study area, which may pose a special problem that needs to be considered in the study design.

The problem of the loss of identity of individuals turns up in various other circumstances. It is particularly important in medical studies, where the long-term fate of patients given some treatment needs to be followed for many years. Various forms of "life-testing" also bring up the issue, inasmuch as some items may be removed from the study for reasons that have no bearing on the duration of life of the item. These difficulties led to a study of ways to deal with the incomplete observations by Kaplan and Meier (1958). It has been proposed that their methods for compensating for inadvertent losses be used in radiotelemetry studies. (Pollock et al. 1989).

The basis for the Kaplan-Meier approach is to handle the individuals lost from the study ("censored" observations) by breaking the records down in intervals and estimating survival separately for each such interval. Overall survival for a longer period is estimated from the product of the individual intervals just as in equation (10.32). It turns out that the approach of Heisy and Fuller (1985) described above can be used to give essentially the same results,

purpose. Pollock et al. (1989:Table 1) illustrate the method using weekly data. Their Table 1 with Heisy-Fuller calculations added shows the agreement between the two methods:

Week	Number at risk	Deaths	Censored	Kaplan-Meier Survival	Cumulative weeks	Cumulative deaths	Heisey- Fuller Survival
1	18	0	0	1	18	0	1
2	18	0	0	1	36	0	1
3	18	2	0	0.8889	54	2	0.8930
4	16	0	0	0.8889	70	2	0.8905
5	16	0	0	0.8889	86	2	0.8890
6	16	1	0	0.8333	102	3	0.8360
7	15	0	0	0.8333	117	3	0.8337
8	15	1	1	0.7778	132	4	0.7818
9	13	1	2	0.7179	145	5	0.7292
10	10	1	1	0.6462	155	6	0.6738
11	8	0	0	0.6462	163	6	0.6620
12	8	0	1	0.6462	171	6	0.6514
13	7	0	0	0.6462	178	6	0.6403

The data above are cumulated over the span of the study, which, as noted previously, assumes constant survivorship. If there is reason to suppose that survival is not constant, then the data should be analyzed in blocks that are more likely to have constant survival within a block. In many studies, captures may go on throughout the course of the study. Kaplan and Meier (1958) noted that these individuals can be added to the number at risk as they enter the study, and Pollock et al. (1989) dubbed such an approach as a “staggered-entry” model, and illustrated it with weekly data on bobwhite quail (*Colinus virginianus*) survival. They provided a convenient variance estimate [Heisey and Fuller (1985) resorted to a computer program for variance estimation]:

$$v[S(t)] = \frac{S(t)^2[1 - S(t)]}{r(t)} \quad (10.35)$$

where $r(t)$ is the number at risk at time t . Data to illustrate the approach from their Table 2 appear below with variance estimates. In this case sequential survival estimates [equation(10.33)] are multiplied [equation (10.32) to obtain an estimate of survival at the end of the current period. To use the Heisey-Fuller approach on the data, one would have to determine how long the individuals dying during the period (week) were alive during the period, so that it is desirable to record radio-days as best possible. In the table below it is assumed that individuals censored were lost at the end of the period and any added were introduced at the end of the period.

Kaplan-Meier estimates of survival for bobwhite quail [from Pollock et al. (1989: Table 2)] with variance estimates.

Week	Number				Kaplan-Meier	
	at risk	Deaths	Censored	Added	Survival	Variance
1	20	0	0	1	1	0
2	21	0	0	1	1	0
3	22	2	1	0	0.909	0.0034
4	19	5	0	0	0.670	0.0078
5	14	3	0	0	0.526	0.0094
6	11	0	0	0	0.526	0.0119
7	11	0	0	0	0.526	0.0119
8	11	2	0	0	0.431	0.0096
9	9	1	0	0	0.383	0.0100
10	8	0	1	0	0.383	0.0113
11	7	0	0	3	0.383	0.0129
12	10	0	0	6	0.383	0.0090
13	16	4	0	10	0.287	0.0037
14	22	4	0	5	0.235	0.0019
15	23	4	1	6	0.194	0.0013
16	24	4	0	0	0.162	0.0009
17	20	2	0	0	0.146	0.0009

10.13 More complex models for analysis of survival data.

Much of this Chapter has been devoted to what might now be termed “classical” methods for estimating survival. The older methods remain valuable as a background for understanding the newer approaches, and they provide tools for initial analyses and for planning studies. The major drawback in using these earlier methods is the restrictive assumptions that are required to justify the estimates. Section 10.10 provides some rough notions about the outcome of using these estimates when the underlying assumptions are violated.

The newer, “modern”, methods make it possible to use less restrictive and more realistic assumptions, and to take advantage of the computing power now available on most desktops. A disadvantage of the newer methods is the nearly universal dependence on computer programs. A major goal of the present effort has been one of presenting the basis for many quantitative methods without requiring the reader to use any more complex programs than are available on most spreadsheets, with specific reference to EXCEL. For the most part, only basic statistical methods have been used. To use the newer methods, one needs to rely on computer programs, along with more sophisticated statistical methods.

A good starting-place for approaching “modeling survival and testing biological hypotheses using marked animals” is the paper by Lebreton et al. (1992). They recommend (1) starting from a “global” model (essentially one that includes all the parameters that may be relevant) and assess its fit, then (2) select a more parsimonious model (fewer parameters) using Akaike’s Information Criterion (AIC), and (3) testing for the more important biological questions by comparing this model with “neighboring ones” using likelihood

ratio tests, and then (4) find maximum likelihood estimates of model parameters with estimates of precision. Papers utilizing this general approach are now steadily appearing in the ecological literature, and it appears likely that the approach has become the “standard” method for studying survival.

Using AIC in biological studies is the subject of a recent book (Burnham and Anderson 1998) that contains a good deal of practical advice on model selection and inference. It appears, however, that these authors have developed doubts about the importance of a “global” model, at least in contrast to an earlier essay (Burnham and Anderson 1992), and they particularly warn of the dangers of “data-dredging”, i.e., starting with a model with many parameters and discarding parameters by using AIC. The problem basically is that chance fluctuations in the data may seriously influence the outcome, essentially in the manner that plagues stepwise regression (Section 1.12). Burnham and Anderson (1998) thus recommend choosing a “parsimonious” model on largely a priori grounds. How one does that is still an open question in many situations. However, the rapidly developing literature on modeling survival offers opportunities to select a few models that seem to be among the top candidates and thus useful starting-places.

Lebreton et al. (1992) list a sizable number of computer programs that have been used to estimate survival and used RELEASE which is described in detail by Burnham et al. (1987) and SURGE 4.0, which was (and still may be) a proprietary program. Fortunately, Dr. G. C. White has invested a great deal of effort in producing and updating program MARK which contains most of the features of these earlier versions and is freely available through the Web Site listed in Section 8.4.

10.14 Exercises

10.14.1 Calculate the Heinke estimator of survival and its variance for the data of Example 10.1. Calculate the ratio of its variance to that of the Haldane variance estimate [eq.(10.7)]. Compare this ratio to the value you get from eq.(10.13). Is there a difference? If so, explain.

10.14.2 Repeat the analysis of Example 10.2, then extend it to use the first 6 classes and calculate a variance estimate from eq.(10.16). Compare this variance estimate with that of example 10.1, and with the value of the ratio tabulated below eq.(10.17). Discuss your results.

10.14.3 Check on the values tabulated for eq.(10.17) for $k = 6$ and explain the trend from $s = 0.1$ to $s = 0.9$.

10.14.4 Suppose that there were 55 recoveries in the first year after banding in Example 10.1 (instead of 45 recoveries). Compute the chi-square test for compatibility of first recoveries and state the p-value.

10.14.5 Calculate the chi-square value for Example 10.4.

10.14.6 The following resighting data are for female monk seals on Laysan Island. Calculate a survival rate as in Example 10.5. Compute a variance from eq.(10.23). As noted in Section 10.8, it likely is best to bootstrap for confidence intervals, but the estimate should serve for comparison to the results of Exercise 10.14.7 below.

	Tagged	0	1	2	3	4	5	6	7
1983	10	10	10	9	8	9	9	6	6
1984	13	12	10	10	8	7	6	6	
1985	14	10	9	9	7	5	5		
1986	17	12	12	9	7	6			
1987	15	13	11	9	7				
1988	17	11	5	4					
1989	13	6	6						

10.14.7 The following data are resightings for female monk seals banded on French Frigate Shoals in 1984. Calculated survival from eq.(10.24) along with a variance about regression (you can do this with EXCEL's regression tool). Compare this variance with that found in Exercise 10.14.6. Also calculate s from eq.(10.25) and compare with the results of eq.(10.24) and Exercise 10.14.6.

Tagged	1	2	3	4	5	6	7
39	35	35	33	30	28	26	26

11.0 THE LOTKA-LESLIE MODEL

11.1 Introduction

The life table is a useful technique for studies of a stationary population or of a single cohort. Changing populations require the additional features provided by the Lotka-Leslie models which add age-specific reproductive rates (m_x) to the survivorship (l_x) entries from which the classical life table is developed. A further feature of the Lotka-Leslie approach is the calculation of a stable age structure (c_x).

The use of life tables in ecology became popular after the classic paper of Deevey (1947). Most of the illustrations for natural populations were unfortunately based on observed age structures or band returns with little assurance that the age data came from stationary populations or that survival rates were constant in the banding data. More recently it has become evident that these are risky assumptions. Burnham and Anderson (1979) and Anderson et al. (1981) discussed the problems and gave tests of the underlying assumptions for banding data. Tait and Bunnell (1980) noted that ages of animals found dead could be used with m_x data to estimate λ , and Van Sickle et al. (1987) have conducted further exploration of that approach. Although dependable survival rates cannot be extracted from age structure samples (if used alone) of non-stationary living populations, some useful information can be obtained from the "apparent" survival rate calculated from such populations (Eberhardt 1988).

Since fluctuations in population size caused by human influences on natural populations have become nearly ubiquitous, the classical life table has become largely an abstraction and laboratory tool. Practical purposes are now better served by the Lotka-Leslie model. However, "life tables" now appear in the literature that contain reproductive data as well as survival data, and are often the basis for calculation of a rate of change of a dynamic population. It may be worthwhile to denote these tables as "augmented" life tables to distinguish them from the classical life table, which contains only data on survivorship.

11.2 Discrete and continuous population growth

Many large mammal populations exhibit what Caughley (1977) termed a "birth-pulse" growth pattern. That is, reproduction takes place in a relatively short period each year. In between these pulses of growth, mortality takes its toll, and the population decreases. However, most textbooks treat growth as a continuous process and use an exponential curve to represent the growth pattern. I believe that a more realistic model is that of compound interest, with the compounding done annually through the birth-pulse. Fig. 11.1 shows the

two models as compared to the likely actual trend of a population.

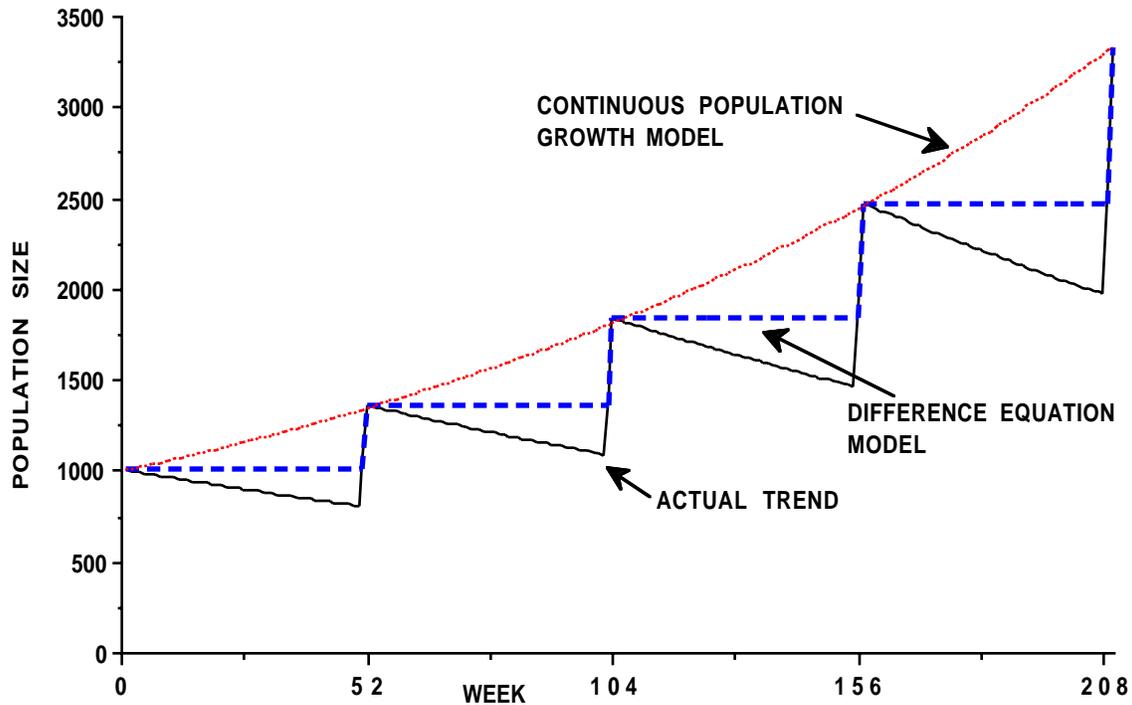


Fig. 11.1 Population growth modeled by continuous (exponential) and discrete (geometric) growth curves and the likely actual course of a population's size.

Very likely the best approach for most purposes is to use a "finite growth" model, using λ to denote the rate of change and writing:

$$N_t = N_0 \lambda^t \quad (11.1)$$

where N_0 denotes initial population size and N_t is population size t years later. We can then interpret λ as $\lambda = e^r$ for continuous growth and $\lambda = 1 + r$ for the compound-interest or geometric growth model. Fig. 1 suggests that the geometric growth model is closer to reality for birth-pulse populations, but the frequent use of the exponential model in texts makes it desirable to have the two interpretations in mind.

Another consideration is that we need methods to estimate rates of change, and using different methods can lead to some confusion. When there is a sequence of annual measurements of population size, the usual approach is that of log-linear regression. That is, we take the natural logarithm of population size and convert eq.(11.1) to:

$$\log_e N_t = \log_e N_0 + t \log_e \lambda \quad (11.2)$$

and fit a linear regression of the form $y = a + bx$. Then the slope (b) estimates $\log_e \lambda$, and if we use the interpretation in which $\lambda = e^r$, then $\hat{r} = b$ inasmuch as $\log_e e^r = r$. Using the interpretation in which $\lambda = 1 + r$, then b estimates $\log(1+r)$, and we have to take the antilog and subtract unity to get r .

In some circumstances, however, we may want to estimate rate of change from two successive measurements of population size, getting:

$$\hat{\lambda} = \frac{N_{t+1}}{N_t} \quad (11.3)$$

We thus need to keep in mind the quantity being estimated. Using log-linear regression the logical estimate is r , while the ratio estimator gives λ directly. How important is the difference? If the rate of change is small, the two interpretations are nearly indistinguishable. But for larger rates of increase, the two interpretations are sufficiently different to make an appreciable difference in, say, projections of population size into the future. Fig. 11.2 exhibits the values of the two interpretations of λ .

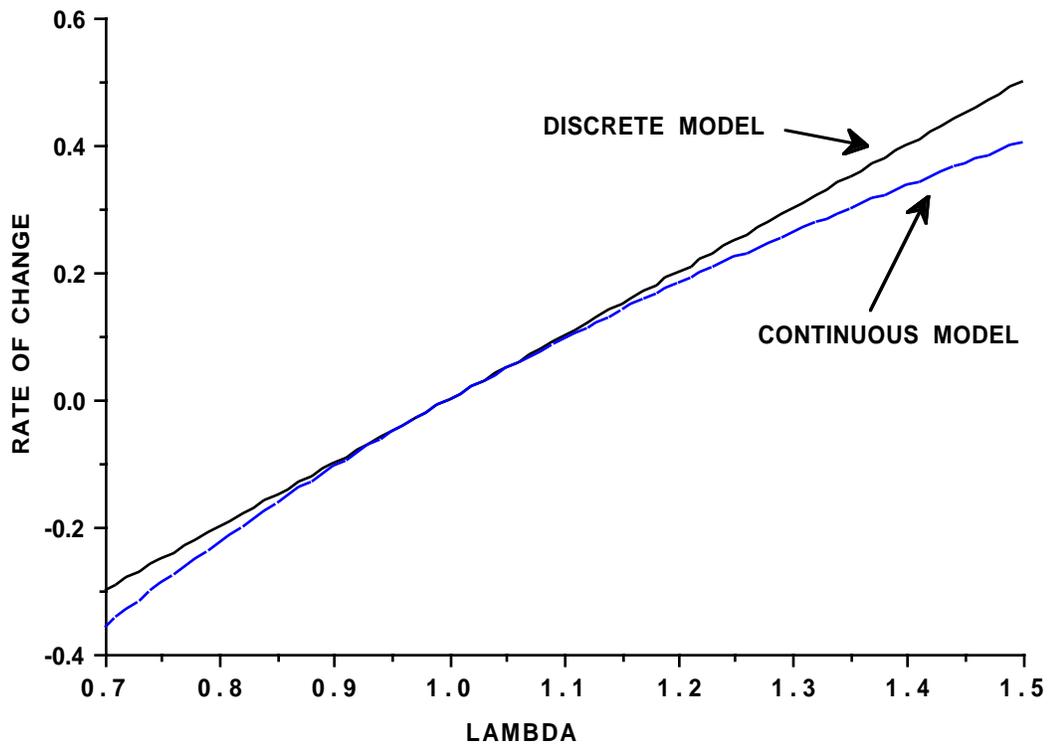


Fig. 11.2 values of r for the two interpretations of λ . The continuous model ($\lambda = e^r$) yields values appreciably lower than those for the discrete or geometric model ($\lambda = 1 + r$) when λ is significantly below or above 1.0.

Some discussions of the continuous model present the rate of increase as having two components, $r = b - d$, where b represents births and d , deaths. These are "instantaneous" rates, i.e., appropriate to indefinitely small increments of time. However, the model of Fig. 11.1 doesn't work that way. The model for the actual trend for one year could be written as:

$$N_{t+1} = N_t s^52 R = N_t \lambda \quad (11.4)$$

where s is a weekly survival rate. In actuality, the model would be more complex, with s being a blend of rates over the year and R , the birth-pulse,

usually doesn't occur instantaneously. Nonetheless, eq. (11.4) describes the general pattern -- mortality occurs throughout the year, while reproduction occurs in a short annual period. The geometric approximation operates as a "step-function", assuming the population remains constant throughout the year, and then jumping to the new level at year-end.

One very important consequence of the birth-pulse model is that the basic unit of time is one year. This means that models for population growth need to be difference equation models, rather than the continuous models implied by the differential equations commonly presented in ecology textbooks. Differential equations are often derived as limiting expressions for difference equations. That is, a relationship is written as a difference equation with time unit Δt and then Δt is assumed to become very small, resulting in an "instantaneous" rate or differential equation. Thus, if the units of time become very small (days, perhaps, instead of years) the discrete model becomes nearly indistinguishable from the continuous model. This does not work with birth-pulse models, due to the fact that births occur only once a year. Consequently, the more realistic difference equation models are not necessarily suitably approximated by continuous models. However, it is often possible to obtain a perfectly reasonable fit of a continuous model to the data. Thus, in Fig. 11.1, if we observe population size annually just after births take place (or just before), the continuous exponential model (dotted line) fits the data very nicely. However, the rates of increase (r) implied by the two models may be quite different unless λ is in the neighborhood of unity (Fig. 11.2). In many situations, one can ignore these details by using eq. (11.1). Problems may come in at two points: (1) when estimating λ by eqs. (11.2) or (11.3), and when a more complex model for population growth is fitted.

11.3 Lotka's equation and the Leslie matrix model

The underlying model was derived by Lotka (1907) for continuously reproducing populations. Leslie (1945) derived his matrix formulation as an approximation to a continuously reproducing population, and the demographic textbook approaches (e.g., Keyfitz 1968) also focus on the continuous situation, since they are largely concerned with human populations, where reproduction occurs the year around. Keyfitz (1968:Ch.8) discussed convergence of the matrix equation to Lotka's integral equation. For practical purposes, it is useful to be able to inter-relate the discrete version of Lotka's equation (eq.(11.1)) with the Leslie matrix. This is readily done through the characteristic polynomial (characteristic equation) of the matrix, if the "maternity values" (F_x) of the Leslie matrix are expressed as $F_x = s_0 m_x$, where s_0 is survival from birth to age 1. This points up the main difference between the two formulations -- Lotka's equation pertains to the population just after reproduction occurs, while the Leslie matrix pertains to the population just before reproduction. That is, the first term of the stable age distribution (eq.(11.6)) is $c_0 = B$, the proportion of newborns in the population, while the first element in Leslie's population vector (n_1) represents the youngest age class, just before parturition takes place, and one year after births occur.

Cole (1954) showed that Lotka's integral equation can be replaced by a summation for populations in which births are concentrated in a relatively short period of the year. In practice, solutions of Lotka's equation have

necessarily been approximated by a summation, in any case. The basic equation is:

$$1 = \sum_{x=a}^w \lambda^{-x} l_x m_x \quad (11.5)$$

where l_x denotes survivorship from birth (or from recruitment to some older portion of the population) to age x , m_x represents age-specific reproductive rates, and the summation ranges from the age of first reproduction (a) to the maximum age (w) assumed represented in the population. Lotka's intrinsic rate of increase (r) is widely used, but is replaced here by using λ since a finite rate of increase is appropriate for the many species populations that reproduce in a short annual period ["birth pulse" populations in the useful terminology of Caughley (1977)]. One can thus interpret λ as e^r or $1+r$.

The results given here were derived for "birth-pulse" populations. Continuous breeding may take several forms. If the reproductive rate is constant throughout the year, then equation (11.5) can be used as an approximation, as was done by Birch (1948). If constant, continuous breeding occurs only over part of the year, with no reproduction at other times, then it may be necessary to resort to the approach of Leslie and Ranson (1940). In a third pattern, reproduction may take place at any time of the year, but there is a definite seasonal peak. Murray and Garding (1984) presented a general approach for populations with "discontinuous breeding seasons" that depends on an "average m_x " but this rate is defined only by the statement (Murray and Garding 1984:325) " m_x is the average fecundity of individuals in age class x , in the sense that the $C_x(t)$ individuals produce $C_x(t)m_x$ offspring between t and $t+1$ that are observed alive at $t+1$...", and no method for calculating this "average m_x " was given in the paper. Consult Caswell (1989, 2001) for more explicit approximations.

Lotka's model, in the form given by Cole (1954), represents the population just after births take place. If we consider the population structure just before the annual period of parturition, the characteristic polynomial of the Leslie (1945) matrix gives the same result. The Lotka equation (eq. (11.5)) is somewhat easier to write out and preserves the central (l_x) entry of the traditional life table. To construct a Leslie matrix, one only needs to note that Leslie's reproductive rates can be written as $F_x = s_0 m_x$, where s_0 = survival from birth to age 1 (Eberhardt 1985). Annual survival rates appear in the Leslie matrix as $s_x = l_{x+1}/l_x$ (hence $s_0 = l_1$). It is important to note that the calculations here assume first-year survival (s_0) to be independent of fate of the adult female. Caswell (1989, 2001) gave definitions of the F_i that permit structuring the Leslie matrix for censuses of "birth-pulse" populations taken at any time of the year, but this requires some approximations. He also considered birth-flow populations, and devoted most of the books to "stage-structured" models, which are applicable when ages cannot readily be assigned to individuals.

11.4 The stable age structure

The stable age distribution is calculated as:

$$c_x = B \lambda^{-x} l_x \quad (11.6)$$

because c_x represents proportions summing to unity, we have

$$B = \frac{1}{\sum_{x=0}^w \lambda^{-x} l_x}$$

1.5 Augmented life tables

An "augmented" life table is simply a convenient way to present the data for Lotka-Leslie calculations in table form. Basic data for an augmented life table are a column of ages, and the l_x , m_x , and c_x columns. Iterative solutions of eq.(11.5) are needed to determine the rate of increase (λ or r) and to calculate c_x . Such solutions are readily obtained with a simple computer program (or on a programmable calculator). A convenient alternative is to incorporate the data in a spreadsheet program that includes a column of values of $\lambda^{-x} l_x m_x$. Trial and error calculations can then be used to satisfy eq.(11.5). However, many of the currently available spreadsheet programs include iterative procedures, so the table can be set up to provide a direct solution of eq.(11.5). This can readily be done with the "SOLVER" routine in EXCEL (found in the TOOLS menu). The column of $\lambda^{-x} l_x m_x$ values should nonetheless be maintained, since it is both a convenient way to examine approximations and provides useful information, described below. One of the convenient features of the spreadsheet approach is that the data can often be transferred directly into graphics and word processing programs. At least one of the commercially available spreadsheet programs (EXCEL) also incorporates routines for using matrix equations, so that various further operations on the Leslie matrix are feasible directly from an augmented life table, if desired. For most practical purposes, the Leslie matrix is not needed. The term, "augmented", is likely superfluous, as most contemporary writers now simply use "life table" to include reproduction, although the original meaning concerned only survivorship.

11.6 An augmented life table

The example used here (Table 11.1) is based on data on elk (*Cervus elaphus*) reported by Houston (1982). The survivorship data given by Houston (1982:Table 5.8) were fitted to an equation given by Eberhardt (1985:eq.(6)) by nonlinear least-squares on logarithms of the survival data (Fig. 11.3). The equation used is:

$$l_x = \exp[-a_1 - a_2 x - a_3(\exp(b_3 - 1))] \quad (11.7)$$

Reproductive data (Fig. 11.2) from Houston (1982: Tables 5.1 and 5.2) were used with eq (11.8), developed by Eberhardt (1985):

$$m_x = a[1 - \exp(b(x - c))]\exp[-a_3 \exp(b_3 x - 1)] \quad (11.8)$$

Pregnancy rates for older animals were obtained as a weighted average of the 1962-1967 data given by Houston (1982: Tables 5.1 and 5.2). Data for the older elk were reported as an average for all animals 16 years of age and older, and are arbitrarily plotted here at age 18.

There is now a considerable degree of confusion in the literature about parameter definitions, which is too extensive and too pervasive to discuss in detail here. Jenkins (1988) provided a good review of the confusion in textbooks. David (1995) more recently reviewed different formulations of eq. (11.5). Much of the difficulty stems from the fact that both the Lotka and the Leslie approaches actually refer to continuously-breeding populations, but are computed with discrete approximations. Definitions of continuously breeding populations thus involve averaging over some time period, usually one year in length. The present approach depends on Cole's (1954) development of the "Lotka" equation as a discrete model where reproduction is assumed, in effect, to occur instantaneously at the same time every year.

Applications of the augmented life table to actual field data must, in nearly all cases, use that data which can be obtained on a particular population. "Birth rates" are then often actually observed as pregnancy rates. Much of the confusion as to definitions (e.g., "natality", "fecundity") and differences between the Lotka and Leslie models can be avoided in practice simply by computing both approaches. This is readily accomplished by spreadsheet methods.

All that is needed is to multiply the c_x column (stable age distribution) of Table 11.1 by an initial population size (N_0). These values can then be projected forwards one year by multiplying by entries in the s_x column to create a column (N_1) of projected numbers in age classes 1,2,3,... . The first entry in N_1 (age class zero) is obtained as the sum of products of the m_x entries (Table 11.1) with those in column N_1 (in most spreadsheets it is convenient to enter the individual products in an intermediate (B_1) column). The same procedure may be used to project N_1 to N_2 , N_2 to N_3 , and so on. The overall procedure is essentially that of the Leslie projection matrix, yielding the corresponding series of population vectors (N_1, N_2, N_3, \dots). The only difference, apart from a lack of matrix notation, is that newly-born individuals are represented by an entry, whereas they do not appear in the age vectors of a Leslie matrix projection.

The projection calculations should carry forward fractional "animals". If this is done, then the ratio of successive sums will yield the same value of λ as obtained by solution of eq (11.5). It is instructive to round the projection vector (N_j) entries (readily done by a standard spreadsheet command) and calculate a series of λ_j . For small and moderate population sizes such a calculation gives a worthwhile illustration of the desirability of including no more than 2 decimal places in reporting estimated values of λ . For example, the data of Table 11.1 were used to project an initial elk population of 1,000 individuals. With fractional entries, the successive ratios of column sums are all 1.20133, but if rounding to the nearest "individual" in each age-class is used, successive values are 1.2010, 1.1998, 1.1964, 1.1972, 1.2010, 1.2000, 1.2013, 1.2012, 1.1989, and 1.2005. Note that, in calculating λ by eq.(11.5), one will

usually carry more than 2 decimal places in order to be sure that eq.(11.5) "balances".

Table 11.1 Augmented life table for elk data from Houston (1982)

Age	Reproductive rate $m(x)$	Survivorship $l(x)$	Adjusted net maternity rate	Stable age distribution $c(x)$	Survival rate $s(x)$
0	0	1	0	0.2426	0.6745
1	0	0.6745	0	0.1362	0.995
2	0.0749	0.6711	0.0348	0.1128	0.995
3	0.4682	0.6677	0.18031	0.0934	0.9934
4	0.4732	0.6634	0.1507	0.0773	0.994
5	0.4728	0.6594	0.1246	0.0639	0.9936
6	0.4721	0.6552	0.1029	0.0529	0.9929
7	0.4711	0.6505	0.0849	0.0437	0.9918
8	0.4696	0.6452	0.0698	0.0361	0.9903
9	0.4674	0.6389	0.0573	0.0297	0.988
10	0.4641	0.6312	0.0468	0.0245	0.9845
11	0.4592	0.6215	0.0379	0.02	0.9795
12	0.452	0.6087	0.0305	0.0163	0.9719
13	0.4416	0.5916	0.0241	0.0132	0.9608
14	0.4264	0.5685	0.0186	0.0106	0.9445
15	0.4048	0.5369	0.0139	0.0083	0.9208
16	0.3746	0.4944	0.00983	0.0064	0.8865
17	0.3337	0.4383	0.0065	0.0047	0.8377
18	0.281	0.3671	0.0038	0.0033	0.77
19	0.2174	0.2827	0.0019	0.0021	0.6791
20	0.1484	0.192	0.0007	0.0012	0.5632
21	0.084	0.1081	0.0002	0.0006	0.4261
22	0.036	0.0461	0	0.0002	0.2813

The "adjusted net maternity rate" is:

$$\lambda^{-x} l_x m_x$$

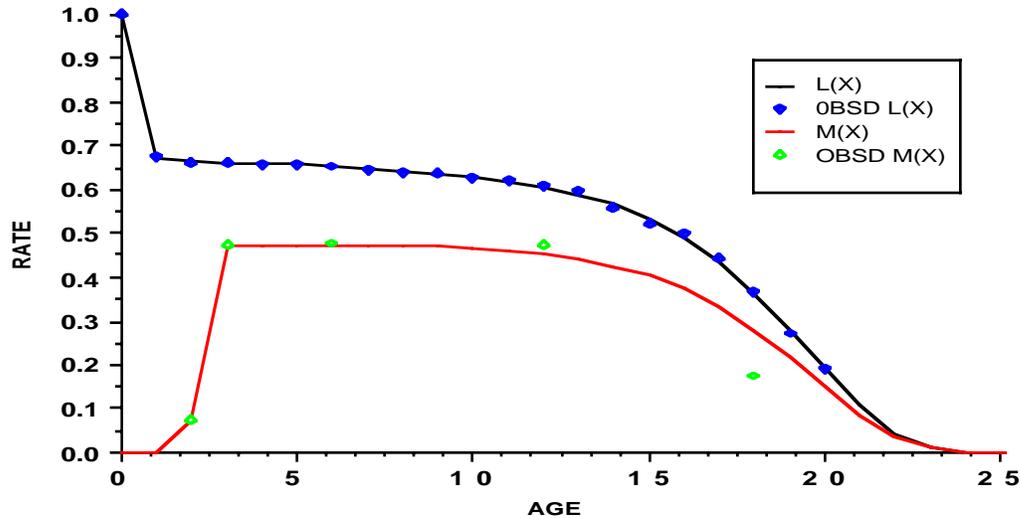


Fig. 11.3. Survivorship (upper curve) and reproductive data for elk from Houston (1982). Fitted curves (eqs. (11.7) and (11.8)) are shown, along with estimated age-specific rates (m_x rates were averaged for some age groups).

11.7 Reproductive and survival curves

The Lotka-Leslie model depends on age-specific rates. Ideally, λ would be calculated from eq. (11.5) using estimates of reproductive (m_x) and survivorship (l_x) rates for each age class, and the models are often presented as if such individual age-class values were available. In practice, however, we seldom have the large volume of data necessary to estimate rates for each age-class, and some kind of averaging has to be employed. Because the l_x values are calculated as the product of individual survival rates such curves always appear smooth, even though the actual survival estimates may be very erratic. It is thus desirable to also plot individual survival rates.

In practice, fitting curves like those of eqs. (11.7) and (11.8) may not be particularly useful or desirable. Very likely their main utility lies in comparing different data sets and in dealing with the issue of senescence. Because sample sizes are usually small, very few older individuals will be observed so that little will be known about senescence in a given species. In many studies, the possibility of lower reproductive and survival rates in the older age classes has simply been ignored. For large vertebrates with relatively high adult survival rates this practice can introduce a significant bias.

The practical way to limit such a bias is simply to truncate the age distribution, i.e., to choose a value of the maximum age (w) used in eq. (11.5) below the maximum age actually observed for a given species. Very likely the main value of fitting curves to reproductive and survivorship data is just to demonstrate the impact of various choices of w on the resulting estimate of λ . In many cases there may not be enough data to make a guess at w , so that it may be helpful to review data on other species, as shown in Fig. 11.4 below.

11.8 Some useful approximations

The approximation used here serves when individual age-specific rates cannot be obtained for all age-classes, as usually is the case in practice. It represents survivorship by $l_x = l_a s^{(x-a)}$ (when $x \geq a$) and reproductive rate by $m_x = m$, so that the net maternity curve is $l_x m_x = m l_a s^x$, where l_a denotes survival to the age of first parturition, and s denotes adult survival rate. Senescence is approximated by truncating the net maternity curve, i.e., w in eq.(11.5) is reduced to compensate for dropping the senescence terms of eqs. (11.7) and (11.8). Lotka's equation (eq.(11.5)) then has the solution used by Eberhardt (1985: eq.(9)), expressed here as a polynomial in λ :

$$\lambda^a - s\lambda^{a-1} - l_a m \left[1 - \left(\frac{s}{\lambda}\right)^{w-a+1} \right] = 0 \quad (11.9)$$

when w becomes very large we have:

$$\lambda^a - s\lambda^{a-1} - l_a m = 0 \quad (11.10)$$

If $\lambda = 1$, then $l_a m = 1 - s$, so that recruitment to reproductive age just balances annual mortality $(1-s)$, as required for stationarity. Because 1 or 2 of the younger age classes often show sharply lower reproductive rates than older animals, one may need to use 2 or 3 values of m , rather than the single value of eq.(11.9) for accuracy in the approximation. The polynomial of eq.(11.9) then becomes somewhat more complicated, but is not needed in practice, as the relevant entries can simply be introduced in a spreadsheet table like Table 11.1 for a solution.

When some of the subadult age classes can be assumed to have the same survival rate (s) as adults, eq. (11.9) can be written in terms of survival to age k (as l_k) after which survival is at the adult rate, so that $l_a = l_k s^{a-k}$, giving:

$$\lambda^a - s\lambda^{a-1} - l_k s^{a-k} m \left[1 - \left(\frac{s}{\lambda}\right)^{w-a+1} \right] = 0 \quad (11.11)$$

Utility of the approximations can be assessed by comparing the adjusted net maternity curve for the approximate values with values for the fitted equations (eqs.(11.7) and (11.8)). For the elk data discussed above, the approximation is very good (Fig. 11.4). Using the data on fur seals (*Callorhinus ursinus*) of Eberhardt (1985) gives a little poorer fit (Fig. 11.5), but nonetheless a good approximation.

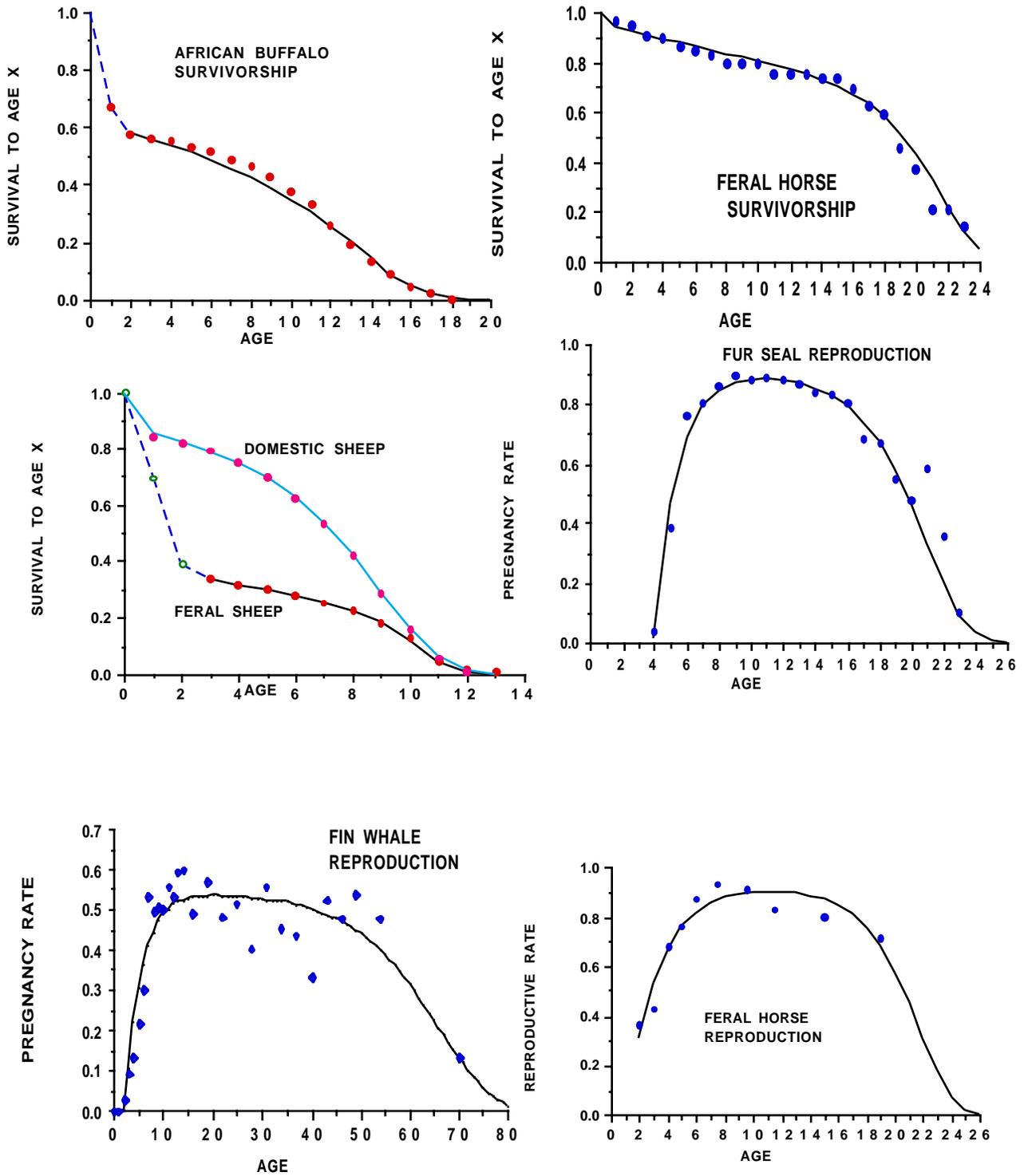


Fig. 11.4. Examples of survivorship and reproductive data. The domestic sheep data are from Hickey (1960), while the feral sheep data are from Grubb (1974). These curves show the impact of mortality early in life on a wild population

very nicely. The feral horse survivorship data are from Garrott and Taylor (1990), while the horse reproductive data are from Berger (1986). The African buffalo survivorship data are from Sinclair (1977), the fur seal data from Eberhardt (1981), and the fin whale data from Mizroch 1981.

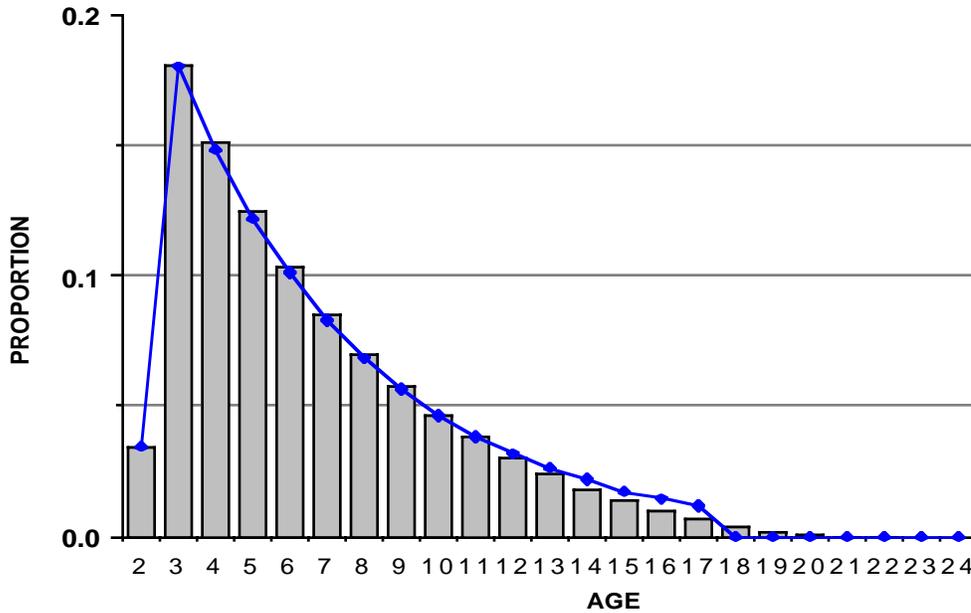


Fig. 11.5 Approximation (solid points) compared to values (bars) from fitted curves (eqs.(11.7) and (11.8)) for elk data from Houston (1982).

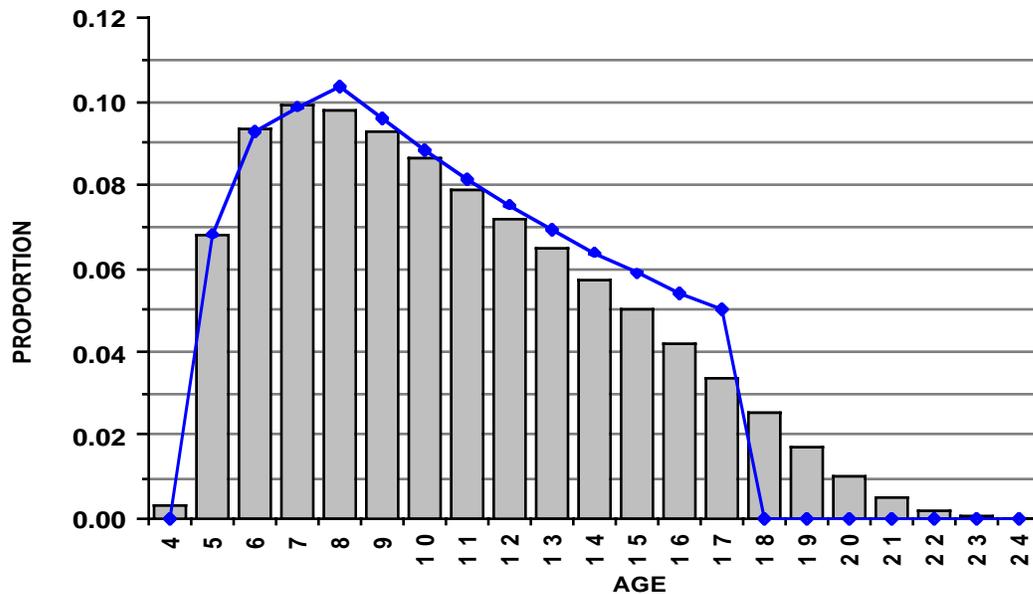


Fig. 11.6. Approximation (solid points) compared to data from fitted curves obtained by Eberhardt (1985) for data on northern fur seals.

11.9 Approximate variance of λ

Setting confidence limits on λ s likely best done by the statistical technique of bootstrapping. However, bootstrapping does not supply estimates

of the relative magnitude of the several components contributing to the overall variance. Such estimates are essential in planning studies, i. e., how much effort should be devoted to estimating each component, and what total effort is required to produce a given width to the confidence interval? Questions of this kind can be approached by approximating the variance of λ by the "delta method".

A general expression for obtaining a variance by the delta method is (Seber 1982):

$$V[g(x)] \approx \sum_{i=1}^n V(x_i) \left(\frac{\partial g}{\partial x_i} \right)^2 + 2 \sum_{i < j} \text{cov}(x_i, x_j) \left(\frac{\partial g}{\partial x_i} \right) \left(\frac{\partial g}{\partial x_j} \right) \quad (11.12)$$

where $g(\mathbf{x})$ is some function of several variables, x_i , $v(x_i)$ is the variance of a given x_i , $\frac{\partial g}{\partial x_j}$ represents the partial differential of $g(\mathbf{x})$ with respect to x_j , and $\text{cov}(x_i, x_j)$ denotes the covariance of two variables x_i and x_j . In the cases considered here, it is assumed that these covariances are zero or negligible.

The partial derivatives provide a useful byproduct, in that their numerical values give an indication of the relative importance of each component variable in determining λ . The partial derivatives have to be obtained by implicit differentiation, due to the fact that there is no "closed form" expression for λ , i. e., $\lambda = g(\mathbf{x})$ cannot be written as a simple mathematical expression (λ is estimated by iterative methods from eq.(11.5), (11.9), (11.10), or (11.11)). Inasmuch as the partial derivatives are complicated, they are most conveniently obtained by one of the computerized routines now available for equation solving. Results for eq (11.9) are:

$$\frac{\partial \lambda}{\partial s} = \lambda [(w-a+1)l_a m \lambda \left(\frac{s}{\lambda}\right)^{w-a+1} - s\lambda^a] / sA \quad (11.13)$$

$$\frac{\partial \lambda}{\partial l_a} = m \lambda^2 \left[\left(\frac{s}{\lambda}\right)^{w-a+1} - 1 \right] / A$$

$$\frac{\partial \lambda}{\partial m} = l_a \lambda^2 \left[\left(\frac{s}{\lambda}\right)^{w-a+1} - 1 \right] / A$$

where $A = (w-a+1)l_a m \lambda \left(\frac{s}{\lambda}\right)^{w-a+1} + \lambda^a (as - s - a\lambda)$.

Recently there has been a considerable amount of interest in "sensitivity" and "elasticity". Sensitivity has been defined as the partial derivative of λ with respect to a_{ij} and elasticity as:

$$e_{ij} = \frac{a_{ij}}{\lambda} \frac{\partial \lambda}{\partial a_{ij}}$$

These definitions refer to the individual elements, a_{ij} , of a Leslie matrix, whereas equations (11.13) are based on implicit differentiation of eq. (11.9). Nine papers in a Special Feature in vol. 18(3) of the journal, Ecology, describe

recent developments. These papers nearly all require matrix algebra and use "stage-structured" models extensively.

Using the delta method requires variance estimates for each component of the equation used. The main use at present appears likely to be in planning a study, or in deciding how to improve a given data set by collecting more observations. One then might assume a binomial distribution for the survival rates for planning purposes. Obtaining a simple variance formula (like the binomial) for reproductive rates is a subject that needs further investigation. More research is also needed to evaluate the accuracy of both the delta method and bootstrapping when used with the Lotka-Leslie model. Two approaches are used here. One is to compare the two methods on actual data sets for several species, and the second is to test the methods by monte carlo simulations.

Bootstrapping is a relatively new statistical technique that takes advantage of the "number-crunching" ability of modern computers. The ability to do many thousands of computations very quickly makes it feasible to "resample" a data set and extract variance estimates. The surprising feature of bootstrapping is that useful results appear to be available for quite small samples.

The approach used here to set confidence limits on λ is reasonably simple, and can be illustrated with reference to eq. (11.9). For simplicity, suppose n_1 observations are available for estimating s , n_2 for l_a , and n_3 for m , and stored in computer files. Exactly n_1 random draws with replacement are made from the file of data for s ("with replacement" means that the same data point can be drawn more than once). An estimate of s is then made from this sample by whatever technique was used for the original estimate. The same process is carried out for l_a and m , using n_2 and n_3 samples drawn with replacement, and an estimate of λ formed from the resulting data.

The technique for forming confidence intervals used here consists in repeating the above scheme a large number of times, say 1,000. The resulting 1,000 estimates are arranged in a frequency distribution, and confidence limits are determined as the points on the frequency distribution that cut off approximately $\alpha/2$ of the observations in each "tail" of the distribution, with α usually set at 0.05. Thus for 1,000 bootstraps, the confidence limits are the points cutting off the smallest 25 observations (lower confidence limit) and the largest 25 observations (upper confidence limit). This is the "percentile" method (Efron and Tibishirani 1993: Ch. 13).

An alternative approach is to use the set of bootstrap estimates to calculate a standard deviation (s), and set approximate 95% confidence limits as the original estimate of $l \pm 1.96 s$. This technique has also been applied to the examples described below, and in the monte carlo simulations, and gave results very close to the percentile method.

Most of the applications described below have been based on eqs. (11.9) through (11.11), and values of a and w were assumed known. Given enough appropriately collected data, one can bootstrap eq. (11.5) directly. Results for one example appear below, along with examples depending on the approximations given above.

11.10 Examples of bootstrapping

The following figures show examples of bootstrapping for a number of species.

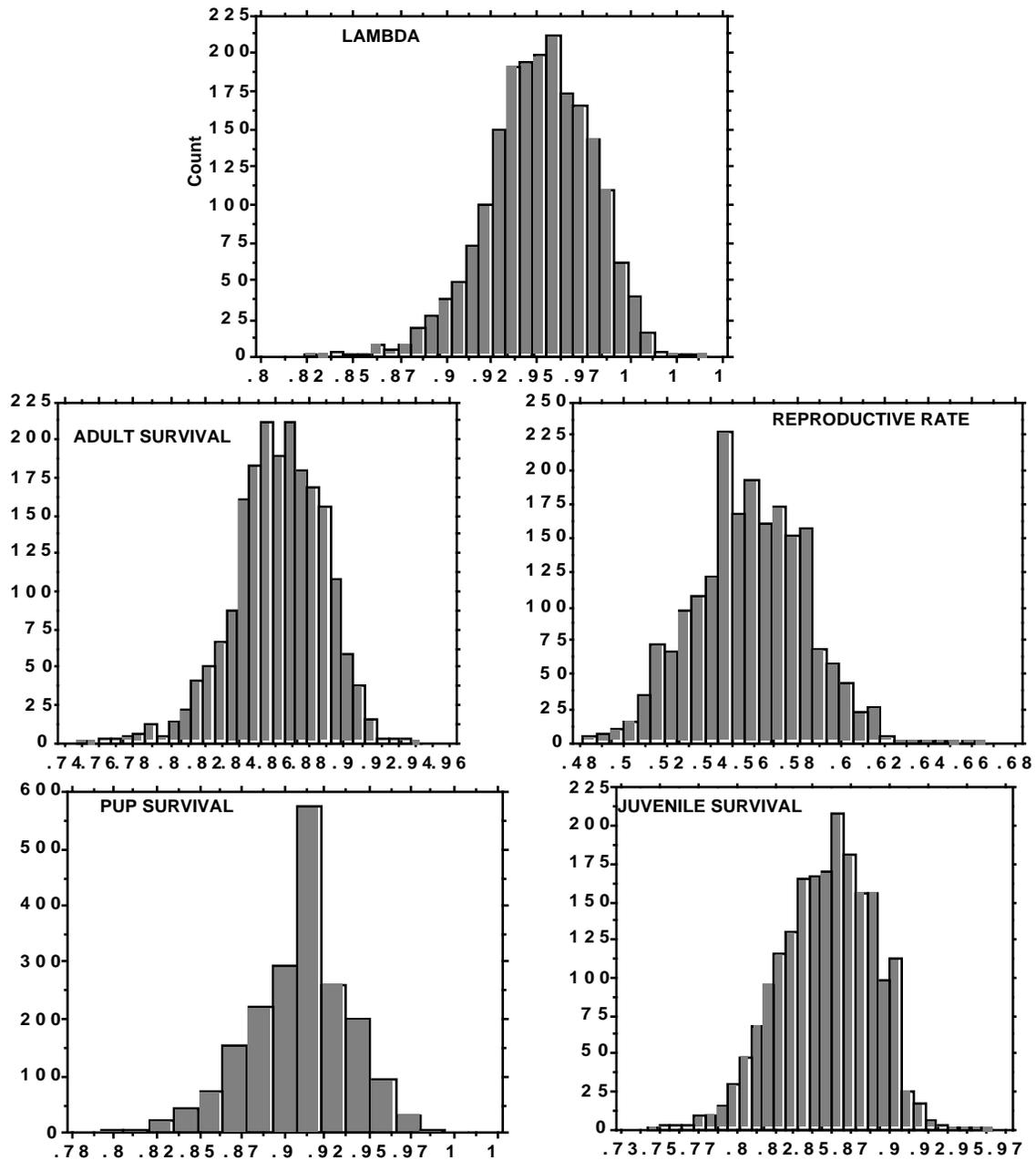


Fig. 11.7. Outcome of bootstrapping survival and reproductive data for the Hawaiian monk seal on Laysan Island. The reproductive and survival data were described by Johanos et al (1994) and Gilmartin et al. (1994). Frequency distributions for the components of eq. (11.5) are shown here. In this instance, the adult survival rate was assumed to apply from age 1 onwards, so that $k = 1$. Early survival has 2 components, survival from birth to weaning, and from weaning to age 1. The estimated value of λ was 0.97, with 95% confidence limits of 0.91 to 1.01. The indicated rate of decline in the population was supported by a log-linear regression on trend data, which gave essentially the same value of λ .

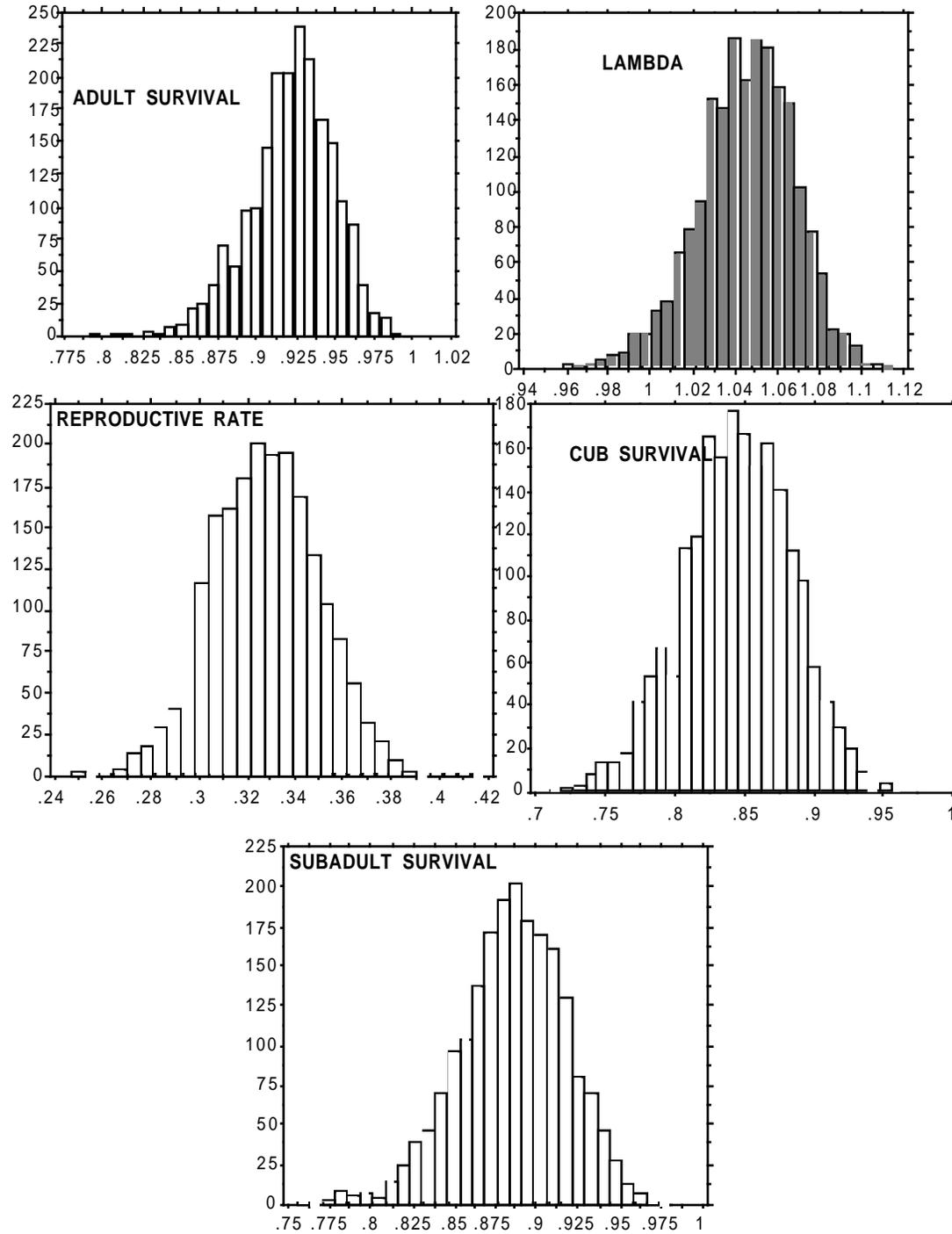


Fig. 11.8. Bootstrapping data from grizzly bears in the greater Yellowstone area. Details of the study were reported by Eberhardt et al. (1994). In this example, two subadult rates were used, so that $l_a = s_0 s_1^4$, where s_0 = cub survival, and s_1 = annual subadult survival rate. Cub survival is appreciably less than subadult survival, which in turn is less than adult survival.

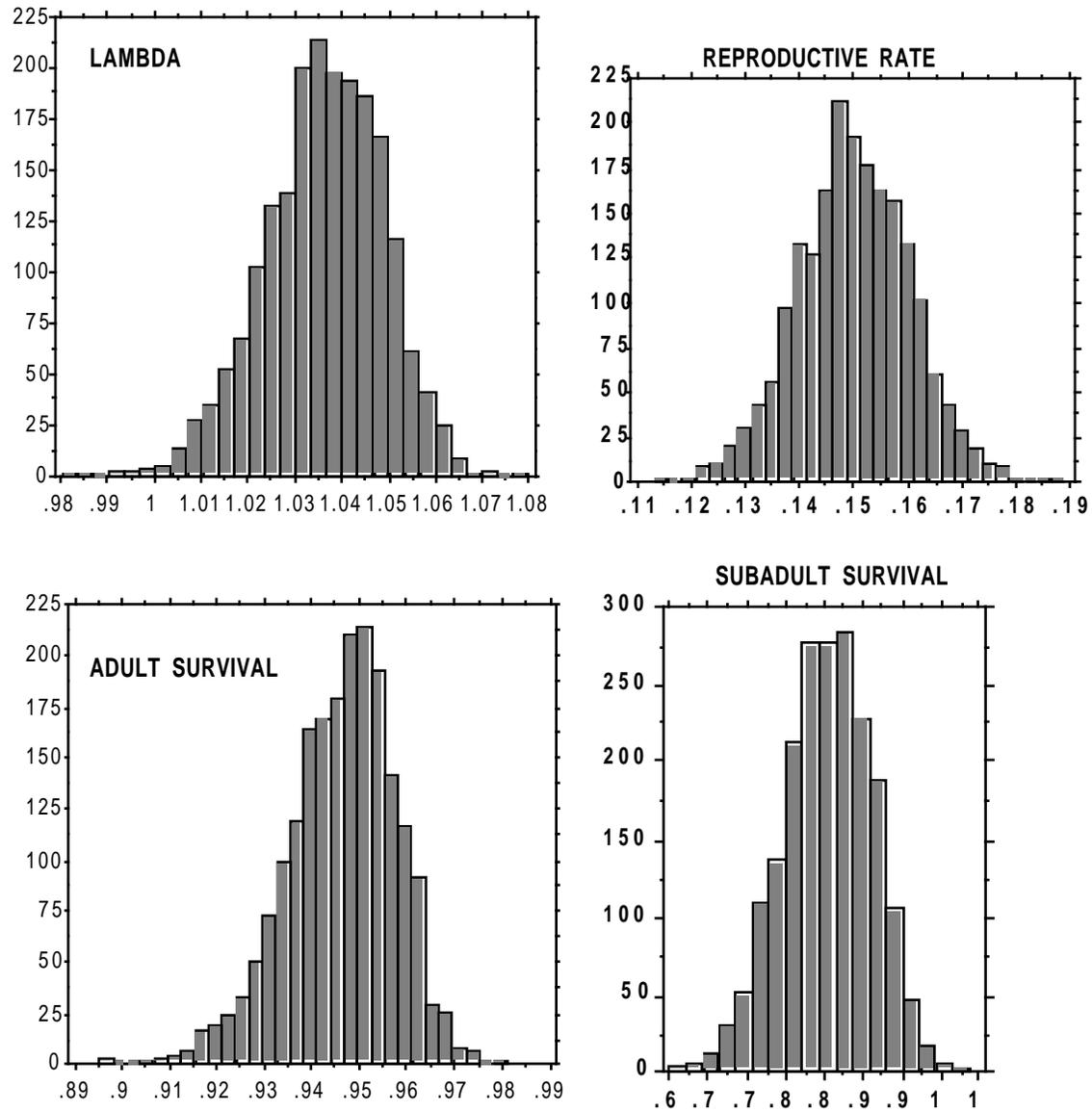


Fig. 11.9. Bootstrapping data for Florida manatees. Manatees reproduce at about 3 year intervals, and thus high adult survival is essential for population growth.

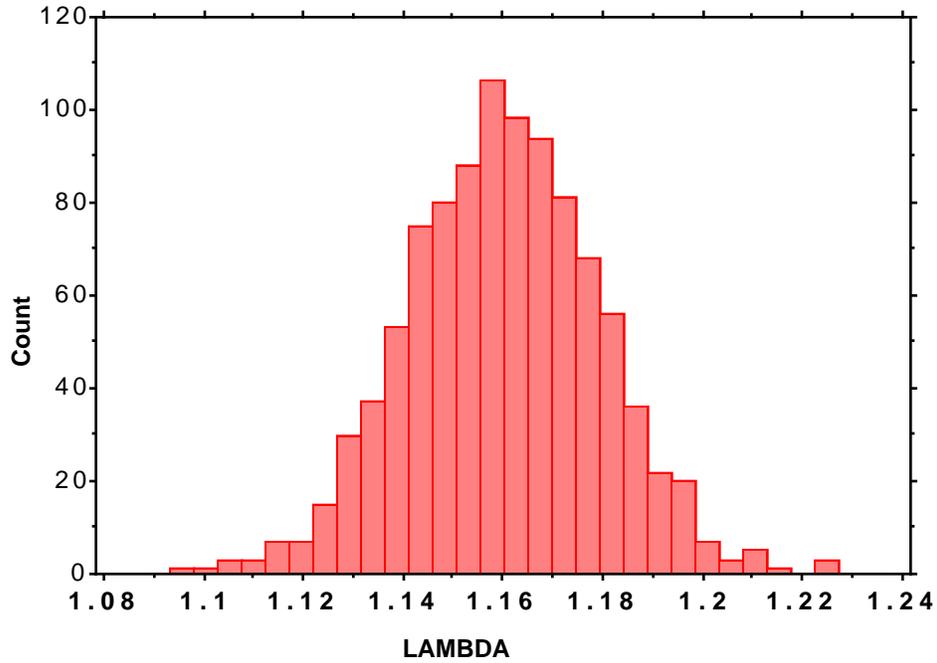


Fig. 11.10. A study of feral horses by Garrott and Taylor (1990) provides sufficiently extensive data for bootstrapping eq. (1) directly. The frequency distribution of λ above is from 1,000 bootstraps of Pryor Mountain horse data. The 95% confidence limits were 1.122 to 1.196.

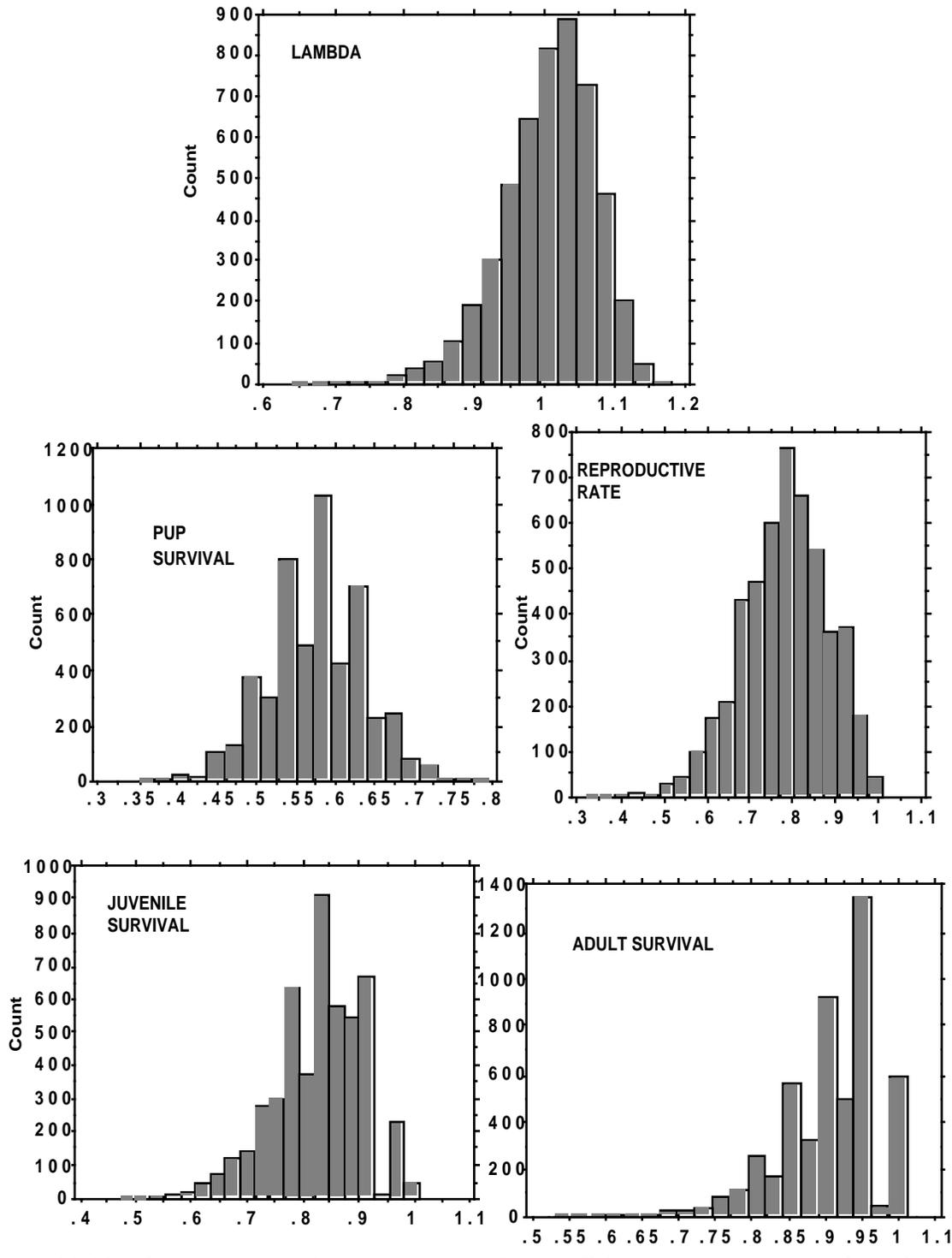


Fig. 11.11. Bootstrapping data from a study of California sea otters. Data largely from Siniff and Ralls (1991).

11.11 Delta method examples

Applying the delta method of eq.(11.12) to data on actual populations gives the results shown in Table 11.2.

Table 11.2 Data from the delta method for actual populations.

		Monk seals	Grizzly bears	Manatees	Feral horses	Sea otters
Survival to age k	l k	0.38	0.52	0.68	0.968	0.63
Reproductive rate	m	0.28	0.33	0.15	0.282	0.23
Adult survival	s	0.87	0.92	0.95	0.974	0.91
Age first reproduction	a	6	5	4	3	3.5
Maximum age	w	20	20	50	20	15
Lambda						
Bootstrapping		0.97	1.046		1.161	1.013
Delta method		0.97	1.046		1.158	1.006
Coefficients of variation						
Delta method		0.03	0.02	0.01	0.009	0.07
Bootstrapping		0.03	0.02	0.01	0.017	0.07
Components of variance						
Survival		0.96	0.44	0.58	0.21	0.81
Early survival		0.02	0.46	0.18	0.03	0.1
Reproduction		0.02	0.1	0.24	0.76	0.09
Total		1	1	1	1	1
Partial derivatives						
Survival		1.1	0.57	1.02	1.03	0.9
Early survival		0.11	0.2	0.09	0.16	0.17
Reproduction		0.3	0.32	0.5	0.55	0.56

11.12 Comparisons with direct estimates of λ

From Table 11.2 it appears that the delta method and bootstrapping yield much the same estimates of λ and variability. One might thus use bootstrapping to obtain confidence limits and utilize the delta method to examine components of variance as a guide to improving estimates of λ by obtaining larger samples of data on the parameters that contribute most to overall variability. Thus in Table 11.2 the main improvement for monk seals will come from obtaining more information on adult survival, which seems often to be the case. However, the grizzly bear data suggest improvements might be equally important for both early and adult survival. For the particular feral horse herd used in the example, it appears that more data on reproduction should be obtained, but this likely is an exception to the general rule due to circumstances.

An important further check on the estimates of λ obtained from the approximations obtained above is to compare estimates of λ from reproductive and survival data with those obtained from direct measures of population size, usually by log-linear regression following the examples of Chapter 9. Fig. 11.12 shows results of such a comparison. The line shown in Fig. 11.2 is a 1:1 line, i.e., shows exact agreement of the 2 estimates. The major discrepancy appears to be the data for caribou, and it seems likely that it is due to underestimation of adult female survival rates, inasmuch as they were reported to be significantly lower than those for subadults.

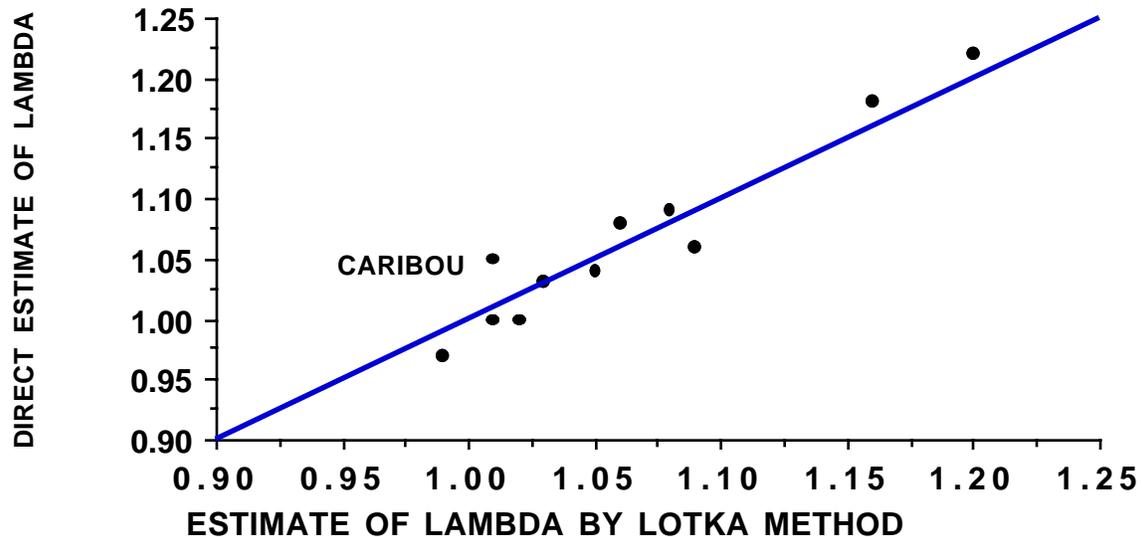


Fig. 11.12. Comparison of direct estimates of lambda with those obtained from reproductive and survival data by the Lotka method. Thirteen examples are plotted of which 2 coincide with other points.

11.13 Development of stable age-structure

The following material gives a simple approach to development of the Lotka equations. A formal proof requires methods of the calculus on a Complex domain. The derivation used here assumes the end result, so does not consist of a "proof" but is rather a simple demonstration of the mechanics of the process. A further simplification is to assume time units of one year and to regard reproduction as occurring instantaneously at the beginning of each year. We also assume that the population is growing geometrically at a rate, r . Thus we have:

$$N_0 = \text{initial population}$$

$$rN_0 = \text{number added in first year}$$

$$N_1 = N_0 + rN_0 = N_0(1 + r) = \text{population at end of first year}$$

$$N_2 = N_1 + rN_1 = N_0(1 + r)^2 = \text{population at end of second year}$$

and, generally:

$$N_t = N_0(1 + r)^t = \text{population at end of } t^{\text{th}} \text{ year} \quad (11.14)$$

The above equation admits the possibility of fractional animals, which is not a matter of great concern if the population size (N_i) is large enough to

avoid the influence of chance fluctuations. There are advantages in keeping fractional animals and thus regarding the values calculated as expectations.

A convenient tabulation of the process has time (in years) along the left margin and the series of year-classes along the top. A further simplification is achieved by supposing that a constant fraction (q) of the animals are lost each year so that $1-q = p$ survive to the next year. It should be noted that this is not a necessary assumption, but is used here to avoid complications. The first two years are shown as:

<u>Time</u>	<u>1</u>	<u>2</u>	<u>3</u>	...	<u>Total</u> <u>population</u>
0	1				1
1	p	$q+r$			$1+r$

Thus the first age-class, at time zero, might be regarded as a "pioneer" generation in a new habitat -- or, equivalently, as some particular segment of a larger population. A fraction, p , of this initial group survives to year 1, as shown at time 1, under year-class 1. Since it is postulated that the population increases to $1+r$ at the end of the first year, it is necessary to add $q+r$ new recruits -- q to replace mortality in the "pioneer" group, and r to achieve the requisite increase. Going into the second year, the table is:

<u>Time</u>	<u>1</u>	<u>2</u>	<u>3</u>	...	<u>Total</u> <u>population</u>
0	1				1
1	p	$q+r$			$1+r$
2	p^2	$p(q+r)$	$(q+r)(1+r)$		$(1+r)^2$

Fate of the pioneers is simply that another fraction (p) survives to year 2, as happens to year class 2. The entry in column 3 contains items to replace the losses, that is $p-p^2$ and $q+r - p(q+r)$. Since $q + p = 1$, this reduces to

$$p-p^2 + q+r - p(q+r) = q(1+r)$$

and we require that the previous year's total ($1+r$) be increased by a fraction r to sustain the geometric rate of increase. Hence year-class 3, which contains all of that year's recruits, has to be:

$$q(1+r) + r(1+r) = (q+r)(1+r)$$

which gives the second year total (adding along the second row) as $(1+r)^2$.

The third year population is constructed in the same way, giving:

<u>Time</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>Total</u> <u>population</u>
0	1				1
1	p	$q+r$			$1+r$
2	p^2	$p(q+r)$	$(q+r)(1+r)$		$(1+r)^2$
3	p^3	$p^2(q+r)$	$p(q+r)(1+r)$	$(q+r)(1+r)^2$	$(1+r)^3$

and a general formula for the i th year can be written down as:

$$\text{Total population} = p^i + p^{i-1}(q+r) + p^{i-2}(q+r)(1+r) + p^{i-3}(q+r)(1+r)^2 + \dots + p^{i-k}(q+r)(1+r)^{k-1} + \dots + p(q+r)(1+r)^{i-2} + (q+r)(1+r)^{i-1} = (1+r)^i$$

where i is the last year considered and k denotes a general term for some intermediate year class. Dividing through by the total population, $(1+r)^i$, gives the proportion in each year-class in year i :

$$\left[\left(\frac{p}{1+r}\right)^i + \frac{q+r}{1+r} \left\{ \left(\frac{p}{1+r}\right)^{i-1} + \left(\frac{p}{1+r}\right)^{i-2} + \dots + \left(\frac{p}{1+r}\right)^{i-k} + \dots + \left(\frac{p}{1+r}\right) + 1 \right\} \right] = 1 \quad (11.15)$$

Note that p is less than unity, so $p/1+r$ is also a fraction, hence large powers of this quantity tend to become vanishingly small. Hence, once i becomes large, the "pioneer generation" virtually vanishes, as do the other early year-classes, and the quantity in brackets approaches an infinite geometric series (but written backwards). Using the equation for the sum of a geometric series:

$$1 + x + x^2 + x^3 + \dots = (1-x)^{-1} \quad (x < 1)$$

we thus can write equation (11.15) as approximately:

$$\frac{q+r}{1+r} \left[1 - \frac{p}{1+r} \right]^{-1} = \frac{q+r}{1+r} \frac{1+r}{q+r} = 1$$

and the proportion corresponding to the x^{th} year class is:

$$c_x = \frac{q+r}{1+r} \left[\frac{p}{1+r} \right]^x \quad (11.16)$$

The quantity $\frac{q+r}{1+r}$ in equation (11.16) is also the ratio of any year's recruits to that year's population total (consider the last year-class entry in any row of the tables above), hence we denote it as

$$b = \frac{q+r}{1+r} \quad (11.17)$$

and regard this as the "birth-rate" per head (really a recruitment-rate).

Since survival from year-to-year is constant, the survival for x years is p^x , and equation (11.16) can be rearranged to

$$c_x = b(1+r)^{-x} p^x \quad (\Sigma c_x = 1) \quad (11.18)$$

11.14 Lotka's equations

In the development thus far, we have each year arbitrarily introduced the number of recruits required to sustain a population increasing geometrically, and this turned out to be a constant fraction (b) of the

population total. If we now assign age-specific reproductive rates, m_x , to each age-class ("year-class" above), an equation for b is:

$$\text{reproductive rate} = b = \sum_{x=0}^{\infty} c_x m_x \quad (11.19)$$

That is, b is the recruitment-rate per head and it is contributed to at a rate m_x , by each proportion of the population, c_x . Of course m_x may be zero, since the early and late age-classes may not reproduce. However, whether or not reproduction occurs in the first age-class depends on the species and the definition of time of recruitment, which might conceivably be set at sexual maturity. Using infinity for the upper limit of summation in equation (11.6) is a convenient convention and avoids the necessity for defining an upper limit to survival. When actual examples are considered, we replace it with w , the maximum age considered in an analysis.

If the definition of c_x given by equation (11.14) is inserted in equation (11.15) we have:

$$b = \sum_{x=0}^{\infty} b(1+r)^{-x} p^x m_x$$

or:

$$1 = \sum_{x=0}^{\infty} (1+r)^{-x} p^x m_x \quad (11.20)$$

and this now gives a general equation containing the several quantities involved in population analysis -- survival, reproduction, and rate of increase.

Recalling that the c_x are proportions summing to unity permits a useful rearrangement of equation (11.18), namely:

$$\frac{1}{b} = \sum_{x=0}^{\infty} (1+r)^{-x} p^x \quad (11.21)$$

Equations (11.18), (11.20) and (11.21) then provide tools for calculating the essential features of population structure. If reproductive and survival rates are known, a trial and error (iterative) solution of equation (11.20) gives the rate of population growth (r) to be expected when the population has reached the stable age-structure given by equation (11.18). Equation (11.21) provides a means for calculating b after equation (11.20) is solved.

The above equations suffer one obvious fault in that a constant rate of survival (p) has been assumed. However, the previously mentioned results of Lotka (1939) and Leslie (1945, 1948) permit replacement of the term p^x above by one that denotes an age-specific survival rate from recruitment to age x

(dated from the time of recruitment). Hence equations (11.18), (11.20) and (11.21) may be rewritten in terms of age-specific survival by replacing p^x by l_x . In practice, one usually estimates annual survival as a series of rates like p_0 = survival from recruitment through the first year, p_1 = survival in the second year, etc., and l_x is then obtained as the product of several such rates.

If reproduction is regarded as a continuous process (as was done in the original development of the equations), then it is necessary to replace the summations above by integral signs, and the geometric rate of increase, $(1+r)^x$, now becomes exponential. However, solutions require the use of the form involving summations, in any case. Hence, we will consider only the summation forms here. These are recapitulated below, using l_x instead of p^x :

$$\frac{1}{b} = \sum_{x=0}^{\infty} (1+r)^{-x} l_x \quad (11.22)$$

$$c_x = b(1+r)^{-x} l_x \quad (11.23)$$

$$1 = \sum_{x=0}^{\infty} (1+r)^{-x} l_x m_x \quad (11.24)$$

The rate, r , is widely known as the "intrinsic rate of increase" (a term coined by A. J. Lotka), and has been subjected to all sorts of attempts at definition in terms of optimal conditions, etc. The best definition is that of equation (11.24), i.e., given constant age-specific reproductive (m_x) and survival rates (l_x) a population will tend to develop a stable age structure, and, if that age structure is attained, the population then increases at rate r . From a practical point of view, one may doubt the likelihood that a particular population has or will actually attain the above conditions, but r is still useful as a quantity defined by equations (11.24) for a given observed set of rates m_x and l_x .

Before considering some simplifications of the above equations it is worth noting that a similar development can be constructed for males by assuming a 1:1 sex ratio at recruitment, and that p_1 represents male survival. Then:

Age structure for males:

$$c_x = \frac{q(1+r)}{1+r} (1+r)^{-x} p_1^x \quad (11.25)$$

$$\text{Total population at year } i = (1+r)^i \left[1 + \frac{q(1+r)}{1+r} \right] \quad (11.26)$$

and, the sex ratio in the population then is:

$$\text{females/males} = \frac{q1+r}{1+r} \quad (11.27)$$

11.15 Net rate of increase

Some variations of the basic equations (11.22 to 11.24) are useful. One is the "net rate of increase" or "net reproduction rate" defined as:

$$R_0 = \sum_{x=0}^{\infty} l_x m_x \quad (11.28)$$

or, in terms of an integral (continuous reproduction):

$$R_0 = \int_0^{\infty} l_x m_x dx \quad (11.29)$$

These equations correspond to equation (11.24) with $r=0$, whereupon $R_0=1$ and the population is "stationary" (neither increasing or decreasing). There is thus the expectation that if $R_0 > 1$ the population should be increasing and vice versa. This is true enough if the age structure is that given by equation (11.19), otherwise it is not necessarily the case, since the equations do not take current age-structure into account.

R_0 has also been used to define something called the "length of a generation". In those species reproducing only once, the length of a generation can be explicitly defined. Familiar examples are insects that reproduce annually and die (or the cicadas, with as much as 17 years of larval development before reproduction and death in a few weeks of adult life), and the Pacific salmon with 3 or 4 year "cycles" between spawning of parental and filial generations. But when a species exhibits repeated and age-specific reproduction the replacement of one generation by another is a gradual process, with no specific time of transition. Thus there is a need to define some measurement if one wants to consider "length of a generation." The necessity for a definition is of major importance in terms of genetic behavior of a population.

Leslie (1966) gives several definitions of length of a generation, with his own recommendation being:

$$\bar{T} = \int_0^{\infty} x e^{-rx} l_x m_x dx \quad (11.30)$$

Leslie pointed out that this is not the same as an earlier definition:

$$T = \frac{\log_e R_0}{r} \quad (11.31)$$

Another definition that has been used is that of the mean age of mothers at the birth of their daughters or "cohort generation time":

$$T_c = \frac{\int_0^{\infty} x l_x m_x dx}{R_0} = \frac{\int_0^{\infty} x l_x m_x dx}{\int_0^{\infty} l_x m_x dx} \quad (11.32)$$

which is equivalent to T when $r=0$. The above equations are written with integrals as would be the case for continuous reproduction, but summation signs need to be substituted for cases where reproduction occurs at discrete time intervals.

Equation (11.31) seems difficult to justify as a useful measurement of "generation time," and for most practical purposes one might elect to use either equation (11.30) or (11.32), with the latter being a little less trouble to calculate. However, students should consult Leslie's paper and Caswell (1989, 2001) if they have occasion to use mean length of a generation.

11.16 Practical applications

With a few exceptions, the actual use of equations (11.22) to (11.24) entails circumstances that permit various simplifications. The exceptions consist largely of human populations (demographic and actuarial data) and a few species that have been extensively studied in the laboratory or in domestication. The main simplification results from the observation that 2 or 3 different survival rate estimates may serve to represent l_x . This largely results as a consequence of constant survival rates for mature individuals. Since most wild animals die violently from predation, accidents, or hunting, and do so at rates great enough to preclude more than a small fraction reaching any sort of old age, it is difficult to determine whether adult age-classes may exhibit age-specific mortality which is simply not observed or whether the adult survival rate is indeed virtually constant. Since very large samples are needed to demonstrate relatively small changes in survival rate, and since year-to-year changes in overall rates do clearly occur, it is not surprising that age-specific rates have seldom been recorded for mature animals.

Whether the younger age-groups are assigned one or two survival rates depends in part on longevity of the species and on quality and quantity of the available data. In most circumstances it seems desirable to use two rates; one for an interval including birth and the relatively short period of high mortality just after birth, and the second for the subsequent interval up to maturity. Very likely early adult life should also be characterized by an additional rate, but only rarely are sufficient data available to do so. The choice of number of survival rates also depends on the choice of age for recruitment to the population being studied. Thus if new recruits are not tallied until they are nearly mature, two survival rates may suffice.

Somewhat similar considerations may apply to reproductive rates. Many avian species appear to have a virtually constant clutch size, so that only one reproductive rate need be used, while the shorter lived mammals may have a smaller first litter, but effectively constant litter size thereafter. Many of the larger mammals show distinct age-specific rates, increasing to the "prime of life", holding relatively constant for some considerable time thereafter, and possibly declining to nearly zero in "old age". In practice, then, age-specific reproductive rates may constitute from two values (zero for younger age-classes and a constant rate for all mature age classes) to perhaps 3 to 5 significantly different rates.

Given a reduced set of l_x and m_x values as suggested in the above paragraphs, one can then simplify the calculations required by equations (11.22) to (11.24) or produce alternative versions to serve as the basis of inferences about population trend, or allowable harvest, etc. Leslie (ibid.) gives an example for a seabird (Uria aalge, the guillemot or murre) which serves to illustrate the procedure. This species first breeds at 3 years of age and it is assumed that clutch size remains constant thereafter. Leslie calculated m_x on the basis of the number of young at fledgling stage, "ready to go to sea", and equal numbers of males and females, so that we start dating events at this time, with $m_0=0$, $m_1=0$, $m_2=0$, $m_3=F=m_x$ ($x > 3$), where F is a constant (since only a single egg is hatched by a breeding pair, F is at most 0.5). Survival for the first year after hatching was defined as P_0 , while that in the second year was P_1 , and the rate in all subsequent years was considered a constant value, P . The essential rates are thus:

Age (x)	0	1	2	3	x(x>3)
Survival	1	p_0	p_1	p	p
l_x	1	p_0	$p_0 p_1$	$p_0 p_1 p$	$p_0 p_1 p^{x-2}$
m_x	0	0	0	F	F

If we let $\lambda = 1+r$, as used by Leslie, then equation (11.15) becomes:

$$\begin{aligned} 1 &= \lambda^{-0}(1)(0) + \lambda^{-1}p_0(0) + \lambda^{-2}p_0p_1(0) + \lambda^{-3}P_0p_1pF + \lambda^{-4}p_0p_1p^2F \\ &= \lambda^{-3}p_0p_1F(1 + \lambda^{-1}p + (\lambda^{-1}p)^2 + \dots) \\ &= \lambda^{-3}p_0p_1F(1 - \lambda^{-1}p)^{-1} \end{aligned}$$

so that:

$$1 - \frac{p}{\lambda} = \lambda^{-3}P_0 p_1 F$$

and:

$$\lambda^3 - p\lambda^2 - p_0p_1pF = 0 \quad (11.33)$$

Hence, given values for P_0 , P_1 , P , and F , one can calculate λ (and thus r) by solving the cubic equation (11.33).

The net rate of increase, R_0 is (equation 11.28):

$$R_0 = p_0 p_1 F (1 + p + p^2 + \dots) = \frac{p_0 p_1 p F}{1 - p} \quad (11.34)$$

so that if $r=0$ (and thus $\lambda=1$) equations (11.33) and (11.34) are identical and represent conditions when the population is of constant size. Thus one can consider the plausibility of the available estimates of P_0 , P_1 , P and F over long periods of time, or conjecture as to the future of the population if survival and/or reproductive rates change.

Using equation (11.24) and $\lambda = 1+r$, we calculate:

$$\begin{aligned} \frac{1}{b} &= 1 + \lambda^{-1} p_0 + \lambda^{-2} p_0 p_1 + \lambda^{-3} p_0 p_1 p + \dots \\ &= 1 + \lambda^{-1} p_0 + \lambda^{-2} p_0 p_1 + \lambda^{-3} p_0 p_1 p (1 - \lambda^{-1} p)^{-1} \end{aligned}$$

which provides an estimate of b from which the stable age-structure can be obtained with equation (11.23). Again, if $r=0$, then $\lambda=1$, simplifying the equation somewhat.

Introducing additional age-specific reproductive rates poses no special problems, beyond including the necessary quantities in a table like the one used above, and keeping track of the necessary algebraic results. This gives results like eq. (11.11). Sometimes, as was done by Leslie in his study, it is convenient to use all adults (3 years old and older in this case) as a base, and calculate proportions of nestlings, 1 and 2 year-olds relative to that base. In practice, one may often not be able to determine the exact age of an "adult", necessitating such an arrangement.

11.17 Exercises

11.17.1 Discrete and continuous rates of increase

Most species show some degree of seasonality in reproduction. There is thus an annual cycle of numbers, peaking at the period of maximum reproductive effort, and reaching a minimum when reproduction is at the lowest annual rate (or not occurring at all in many species). Annual rates of growth should thus be computed from measurements made at the same time each year, and the "compound interest" formula is the appropriate model for population growth:

$$N_{t+1} = N_t (1 + r)^t$$

where t is expressed in years. There are some species that will reproduce continuously under suitable conditions. The continuous time model:

$$N_{t+1} = N_t e^{rt}$$

then provides a close approximation to observations, and t can be measured in time units of varying length.

As a matter of mathematical convenience and custom, many authors use the continuous model exclusively. When r is small, and t not too large, it makes little difference which model is used. Students should try values of $r = 0.01, 0.05, 0.10$ for $t = 5, 10, 20$ in the two equations in order to see how the results differ.

If we utilize λ in equations, it can represent either case, i.e., $\lambda=e^r$ or $\lambda=1+r$. The main difference arises when one reports a rate of increase as r . In one case $r=\log_e\lambda$, and in the other, $r=\lambda-1$. Students should try calculating r from values of λ calculated from the two definitions (i.e., suppose $\lambda=1.25$, and calculate $r=\log_e\lambda$ and $r=\lambda-1$ and compare the results. Compare the series expansions of $\log_e(1+r)$ and e^r for further understanding.

11.17.2 Constructing a stable age distribution

Make a numerical version of the table of Sec. 11.13 using $p=0.9$ and $r=0.1$. Carry it to the fifth generation. Note that each succeeding entry for a given year class is simply p times the previous value, and the last year class in each generation is given by the last term in the equation for total population size given below the table. Check to be sure that your values add to $(1+r)^i$ for the i^{th} generation. Convert the last row into proportions, and calculate the stable age distribution from eq. (11.23) and enter it as a final row in the table. Note that the final year class in each row represents newborns and is consequently c_0 , so that you enter the stable age entries from right to left in the last row.

Since survivors from the zeroth generation are calculated as p^i , it will take quite a long time before that generation is reduced to a level commensurate with the result expected from eq.(11.15). In 20 years, we have $(0.9)^{20}=0.1216$. Compare this with the result you get from eq. (11.23). Hence, with high survival rates, changes in age structure can persist for fairly long time periods.

11.7,3 Calculating rate of increase for fur seals.

Use of equations (11.22) to (11.24) can be illustrated with data on the northern fur seal (*Callorhinus ursinus*). The data appear in Table 11.17.1 below and consist of age-specific reproductive rates (m_x) and annual survival rates (p_x , which are often written as s_x). The m_x entries in the table are one-half of pregnancy rates observed in extensive pelagic collections of fur seals (some 12,000 seals were aged). Survival rates were estimated in various ways. Survival from birth to age 3 comes from estimating pup numbers and the number of 3 year-old males present at the time they were harvested for furs. This three year rate was arbitrarily partitioned over the first 3 years of life. A single adult survival rate was estimated from the rate of decrease of the relative size of age classes from age 3 up to age 16, and then from averaging ratios of pairs of year-classes for animals older than age 16. More details are available in Eberhardt (1981).

Calculations proceed by forming an l_x column from the annual survival rates simply by taking products, with $l_0 = 1$ by definition. Since $m_x = 0$ for the early age-classes, equation (11.20) can be written as:

$$1 = \sum_3^{24} \lambda^{-x} l_x m_x$$

The equation is solved for λ by iteration, i.e., one chooses a value of λ in the neighborhood of unity, calculates the sum, and either increases or decreases λ according as the sum exceeds or is less than unity, continuing until the sum is sufficiently close to unity. Programming a calculator or computer to do the work is desirable if one has more than a few such calculations to carry out.

One iterative method uses linear interpolation. From the data of Table 11.15.1, we can compute that the sum of the $l_x m_x$ is 1.0604. This is equivalent to setting $\lambda = 1.0$. Evidently λ is somewhat larger than unity, so we might try $\lambda = 1.01$. This gives a sum of 0.95787. A simple way to remember the interpolation process is just to note that the two pairs of points computed thus far determine a straight line, $y = a + bx$. It is convenient to let $y = \lambda$, and x represent the sum calculated from the equation above, since we want to predict what value of y results from $x=1.0$. The slope (b) is the rate of change in y resulting from a change in x from 1.0604 to 0.95785. This gives:

$$b = \frac{1.00-1.01}{1.0604-0.95787} = -0.0975324.$$

The intercept, a , is then:

$$a = y - bx = 1.01 + 0.0975324(0.95787) = 1.10342,$$

so the line is thus:

$$y = 1.10342 - 0.0975324x$$

Setting $x=1$ gives $\lambda=1.00587$. Using this value in eq. (11.20) gives a sum of 0.999006. Another iteration can be obtained by following the same procedure, using the new value and that for $\lambda=1$. This gives

$$y=1.10319 - 0.0956119x$$

or $\lambda=1.00578$ which yields a sum of 0.99992. A sketch is always helpful in keeping track of things in the interpolation process.

Linear interpolation formulas given in textbooks and manuals require fewer calculations, but the above procedure should be easy to remember. Other approaches can be used when a calculator or computer is programmed. Perhaps the simplest is to start with $\lambda=1$ and change it by some small increment which is positive or negative depending on whether the resulting sum (eq.(11.20) is greater or less than 1.0. As the sum decreases, then the increment should be made smaller, down to the order of accuracy wanted. As has been noted before here, one can conveniently avoid most of the above by using SOLVER in EXCEL.

When λ has been calculated, eq.(11.22) can be used to calculate the "birth rate per capita" (b) and eq.(11.23) then provides the proportions in various age-classes found in a population having the stable age structure. These are shown for the fur seal data in Table 11.15.1. Note that b is also the first entry in the age structure table (why?).

11.17.4 Rates of increase from Leslie's model

Leslie (1966) gave estimates of $p_0 = 0.7$, $p_1 = 0.8$, $p_2 = 0.9$ and $F = 0.2307$ for a guillemot (murre) population. Solving eq. (11.33) with these values requires finding the positive root of a cubic equation:

$$x^3 - 0.9x^2 - 0.116273 = 0$$

which has the solution $x = 1.01325$ ($=\lambda$). From eq. (11.28), we get $R_0 = 0.116273/0.1 = 1.1627$. Leslie also used $F = 0.375$ and $F = 0.50$ for illustrative purposes. Calculate λ and R_0 corresponding to these two values of F .

11.17.5 Calculating a stable age distribution for Leslie's data

For $F = 0.2307$ and the survival rates given in the exercise above, calculate the stable age distribution for Leslie's guillemots at the time of hatching, assuming that there are 1,000 adults (3 years of age and older). Use the equation following (11.24) to estimate b , and eq.(11.23) to obtain c_x .

11.17.6 Extending Leslie's simplified model

Leslie (1966) assumed that the guillemots began reproduction at age 3. As discussed in Sec. 11.15, he then produced a cubic equation, and solved that for λ . Often, one will not have the data to solve the resulting polynomial, but it may nonetheless be useful to explore possibilities with the data that are on hand. Thus Eberhardt and Siniff (1977) wrote the corresponding equation as:

$$1 = \lambda^{-a} p_0 p_1 p^{a-2} F (1 - p\lambda^{-1})^{-1}$$

where a = age of first reproduction. Students should check to see that this reduces to eq. (11.33) when $a = 3$.

The quantity often missing in field studies is survival from birth to adulthood, so the above equation was rearranged by defining the ratio:

$$K = \frac{p_0 p_1}{p^2}$$

which is the ratio of juvenile survival to that of adults. This then gives:

$$K = \frac{\lambda^{a-1} (\lambda - p)}{F p^a}$$

If we have estimates of F and p , it is then possible to explore the rates of increase likely to result from various ratios of juvenile to adult survival. Eberhardt and Siniff (1977:Fig. 3) plotted some values for convenient inspection of possibilities. Other arrangements of the simplified equations were used to examine other facets of marine mammal population dynamics. An important conclusion was that the age of first reproduction is not as important in determining rates of increase as seems to be assumed in many references.

11.17.7 Calculating length of a generation

Try the equation of Sec. 11.14 on the guillemot data given above. You will need to replace integrals by summations and use sums of series. Make a table showing the three estimates for length of generation for each of the three values of F given in the example.

14	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0
11	12	13	14	15	16	17	18	19	20	21	22	23
0.3097	0.3049	0.2979	0.2876	0.2730	0.2527	0.2251	0.1895	0.1466	0.1001	0.0567	0.0243	0.0243
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0.9795	0	0	0	0	0	0	0	0	0	0	0	0
0	0.9719	0	0	0	0	0	0	0	0	0	0	0
0	0	0.9608	0	0	0	0	0	0	0	0	0	0
0	0	0	0.9445	0	0	0	0	0	0	0	0	0
0	0	0	0	0.9208	0	0	0	0	0	0	0	0
0	0	0	0	0	0.8865	0	0	0	0	0	0	0
0	0	0	0	0	0	0.8377	0	0	0	0	0	0
0	0	0	0	0	0	0	0.77	0	0	0	0	0
0	0	0	0	0	0	0	0	0.6791	0	0	0	0
0	0	0	0	0	0	0	0	0	0.5632	0	0	0
0	0	0	0	0	0	0	0	0	0	0.4261	0	0
0	0	0	0	0	0	0	0	0	0	0	0.2813	0

The first row of the matrix contains the reproductive values for ages listed in the column headings (these are ages one and older as newborns do not appear in the Leslie matrix). The reproductive rates are those of the first column of Table 11.1 multiplied by survival from birth to age 1 (0.6745). The entries in the diagonal starting with age 2 (left column of numbers) are survival rates from age one onwards as shown in Table 11.1. A stable age distribution is calculated as in eq.(11.6) except that one has to recode the age-classes so that age one is now coded as age-class zero. This lets us set up a stable age distribution for 10,000 elk which is used as an age vector (column of ages) along with the MMULT function of EXCEL. To use that function, one enters the range of the Leslie matrix followed by a comma and the range of the initial age structure vector (must have the same number of rows as the Leslie matrix). The MMULT function then produces a new vector which is the population structure one year later. Because we started out with the stable age structure, the same proportions are obtained as in the initial age structure. If one starts with a

different age structure, then repeated projections show a convergence to the stable age structure. The several computations are as follows:

	lambda	1.20133			COHORT		1st	2nd	3rd		
SURVIVAL	Calc. for				1st	2nd	prop	proj	prop	proj	prop
RATES	$l(x)$	$C(x)$	$c(x)$		project	proj					
0	0.995	1	1	0.1798	1798.2	2160.3	0.1798	2595.3	0.1798	3117.8	0.180
1	0.995	0.995	0.828	0.149	1489.4	1789.2	0.149	2149.5	0.149	2582.3	0.149
2	0.9934	0.990	0.686	0.123	1233.6	1481.9	0.123	1780.3	0.123	2138.8	0.123
3	0.994	0.983	0.567	0.102	1020.1	1225.4	0.102	1472.1	0.102	1768.5	0.102
4	0.9936	0.978	0.469	0.084	844.0	1013.9	0.084	1218.1	0.084	1463.3	0.084
5	0.9929	0.971	0.388	0.070	698.1	838.6	0.070	1007.4	0.070	1210.3	0.070
6	0.9918	0.964	0.321	0.058	577.0	693.1	0.058	832.7	0.058	1000.3	0.058
7	0.9903	0.957	0.265	0.048	476.3	572.2	0.048	687.4	0.048	825.8	0.048
8	0.988	0.947	0.218	0.039	392.7	471.7	0.039	566.7	0.039	680.8	0.039
9	0.9845	0.936	0.180	0.032	322.9	387.9	0.032	466.0	0.032	559.9	0.032
10	0.9795	0.921	0.147	0.026	264.6	317.9	0.026	381.9	0.026	458.8	0.026
11	0.9719	0.902	0.120	0.022	215.8	259.2	0.022	311.4	0.022	374.1	0.022
12	0.9608	0.877	0.097	0.017	174.6	209.7	0.017	251.9	0.017	302.7	0.017
13	0.9445	0.843	0.078	0.014	139.6	167.7	0.014	201.5	0.014	242.1	0.014
14	0.9208	0.796	0.061	0.011	109.8	131.9	0.011	158.4	0.011	190.3	0.011
15	0.8865	0.733	0.047	0.008	84.1	101.1	0.008	121.4	0.008	145.9	0.008
16	0.8377	0.650	0.035	0.006	62.1	74.6	0.006	89.6	0.006	107.6	0.006
17	0.77	0.544	0.024	0.004	43.3	52.0	0.004	62.5	0.004	75.1	0.004
18	0.6791	0.419	0.015	0.003	27.7	33.3	0.003	40.0	0.003	48.1	0.003
19	0.5632	0.285	0.009	0.002	15.7	18.8	0.002	22.6	0.002	27.2	0.002
20	0.4261	0.160	0.004	0.001	7.4	8.8	0.001	10.6	0.001	12.7	0.001
21	0.2813	0.068	0.001	0.000	2.6	3.1	0.000	3.8	0.000	4.5	0.000
22		0.019	0.000	0.000	0.6	0.7	0.000	0.9	0.000	1.1	0.000
	sum		5.561	1.000	10000	12013	1.000	14432	1.000	17338	1.000
	B=		0.1798								

Repeat the calculations in order to become familiar with the Leslie matrix. There are many more manipulations possible, and one can calculate lambda directly from the matrix, and so on.

Exercise 11.17.9 Convergence to stable age structure.

The above exercise started out with a stable age structure. Use the same matrix for projections, but start the projections with a different initial age vector. Enter 4287 above the first column of the spreadsheet that will contain your calculations, number a column from 1 to 23 as in the above exercise, but now make the "cohort" entries by multiplying by 0.7 raised to the power of the row numbers, i.e., entries will be calculated from $4287 \cdot 0.7^x$, where x is the row number (1 to 23). The first cohort should be as shown below. Now project out 10 years proceeding as in the previous exercise, and note that the proportions in each age class approach the stable age distribution as in the previous exercise.

	4287
	0.7
COHORT	
1	3000.90
2	2100.63
3	1470.44
4	1029.31
5	720.52
6	504.36
7	353.05
8	247.14
9	173.00
10	121.10
11	84.77
12	59.34
13	41.54
14	29.08
15	20.35
16	14.25
17	9.97
18	6.98
19	4.89
20	3.42
21	2.39
22	1.68
23	1.17
SUM	10000.26

12.0 POPULATION MODELS

12.1 Introduction

A tremendous amount of work on ecological models has been published in the last few decades, including a number of books on the subject, and a journal, *Ecological Modeling*. The focus here will be on models that may be useful in collecting and assessing field data on actual populations. These may range from quite simple equations to complicated computer programs. Almost any kind of analysis depends on a model of some sort. Many scientists use various statistical tests without stopping to consider that each such test depends on a formal model.

The simple linear regression model, $y = a + bx$, serves to assess one possible relationship between paired observations. Fitting a regression model is usually (but not necessarily) done with the least-squares technique. The fitting process depends on the assumption of a model (linear in this case). If least-squares is used in fitting, it brings in the further assumption that the x -values are known exactly, so that the minimizing process involved in fitting deals only with variability in the y -variable (independent variable). For linear regression, this amounts to minimizing the quantity:

$$S = \sum_i [y_i - (a + bx_i)]^2$$

with respect to the coefficients (parameters) a and b .

Going beyond the fitting process to do statistical tests of various kinds requires some further assumptions. For linear regression, the main such assumption is that the model is actually of the form:

$$y_i = \alpha + \beta x_i + e_i \quad (12.1)$$

where the e_i are randomly drawn from a normal distribution with mean zero and a constant variance, σ^2 . We thus have a sequence of events in using models. First one settles on the mathematical form of the model, then on the method of fitting, and finally methods of analysis are considered, i. e., finding out whether the parameters are "significant" and how important they are in the process being studied.

Many population models use time as the independent variable, so that the variation in fitting can reasonably be assumed to be associated solely (or mainly) with the dependent variable. Consequently, least-squares is often an appropriate fitting technique. Unfortunately, many of the models that seem to be useful in ecology are fairly complicated. Consequently, most of the usual statistical analysis techniques can only be validly applied in considering components of the model, and become very doubtful indeed if applied to model outcomes.

This makes the relatively new technique of bootstrapping appear very promising for modelling, as it can be applied to very complex systems. One such application appears in Chapter 11, where bootstrapping and the delta method were applied to a complex model, the Lotka-Leslie model. Because the model has to be solved iteratively, least-squares is no help. Efforts to deal with

the underlying model mathematically depend on the calculus of complex variables and thus become difficult for most biologists. Bootstrapping, however, is quite straightforward and gives results supported by realistic-seeming stochastic models of actual populations.

12.2 Curve-fitting models

A natural extension of the simple linear regression model is to fit more complex curves. Multiple regression offers a simple extension of the linear regression model (eq. 12.1), taking the form:

$$y_i = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + e_i \quad (12.2)$$

where the x_i are several (k) independent variables. Fitting is again readily done by least-squares. Many texts now give the relevant equations in matrix form. Mathematically and computationally speaking, a large number of independent variables can be used, but this is where a lot of biologists get into serious trouble without realizing what they are doing. Many of the available statistical packages will not only fit multiple regressions with many independent variables, but they will also "decide" which of the variables are "statistically significant" via stepwise regression. If fits with the observed variables (x_1, x_2 , etc.) aren't satisfactory, one can try various transformations, e.g., x_1^2 , $\log x_2$, $1/x_3$, etc., looking for a good fit.

The problem with this approach is that it may be difficult to justify even the simplest multiple regression model for biological data. Very often we can be quite sure that linear models are not suitable for biological relationships. Hence experimenting with combinations of variables until one gets a good fit should only be used for predicting a future y -value and then *only* if the model can be proof-tested on an independent data set. Otherwise such predictions may be an exercise in self-delusion.

The most frequently used model in population studies is undoubtedly the exponential function:

$$y = ae^{bx} \quad (12.3)$$

where a and b are again parameters. Usually $a = N_0$, the initial population size, $y = N_t$, current population size, and $x = t$, time of observations. The exponential function is non-linear, i.e., the variables are not related by simple additions and multiplications as in eqs. (12.1) and (12.2).

There are two ways to fit eq. (12.3). One is by non-linear least-squares which requires a computer fitting routine, now found in most statistical packages. Using that approach assumes that the underlying model is:

$$y_i = ae^{bx_i} + e_i \quad (12.4)$$

i.e., that the error terms are additive. The alternative is to take natural logarithms of the y_i , giving the log-linear regression model:

$$\log_e y_i = \log_e a + bx_i \quad (12.5)$$

which is readily fitted by simple linear regression. To justify eq. (12.5), one needs to write the underlying model as:

$$y_i = ae^{bx_i} e_i \quad (12.6)$$

thus assuming that the errors are multiplicative. This appears to be the appropriate model for population data. If one fits an exponential model to a sequence of observations of a population changing at a constant rate, and then examines the deviations of observed points from the fitted curve:

$$\text{Deviations} = y_i - N_0 e^{rt_i} \quad (12.7)$$

it usually turns out that the deviations increase in magnitude with time. Deviations from the log-linear model tend to be independent of time, giving a reasonably constant "variance about regression", and thus conforming approximately to the simple linear regression model requirement of constant variance.

The log-transformed approach using simple linear regression (log-linear regression) is thus to be preferred. This disturbs some workers, who prefer to think in terms of the observed population size. An obvious answer to that complaint is just to present the data in terms of the original measurements, i.e., numbers or counts, rather than the logarithms of those data. Doing so may draw criticism from statisticians, who point out that "transforming back" can introduce bias in estimates. However, the overridingly important result from a log-linear analysis of population trend data is ordinarily the slope, which estimates the rate of change, r , directly (and thus does not need to be transformed back in any case). The other parameter in a fitted equation (N_0) seldom gets much attention, regardless.

In many instances, the basic data may be counts, rather than actual population estimates, so it makes good sense to stay on the logarithmic scale and graph the fitted relationship as a straight line. An advantage, as mentioned above, is that deviations from the fitted regression line tend to be more uniform over time, supporting the notion that log-linear regression gives a useful estimate of the rate of change.

One may thus be led to suppose that it will be acceptable to go ahead and obtain confidence limits for r in the usual manner for linear regression, and this seems to be a rather common practice. Unfortunately, there is reason to doubt that such limits will provide the degree of "confidence" one would expect. In simple linear regression 95% confidence limits on the slope can be interpreted to mean essentially that if we repeat the process generating the observed data many times, then 95% of the resulting estimates of r should fall within the confidence limits obtained in the first place. Hence, it might be supposed that if it were possible to observe many populations growing under conditions identical to the one population studied, about 95% should show values of r within the calculated confidence limits.

The problem is that normal theory confidence limits are based on the assumption that the "errors", e_i , of eqs. (12.1) and (12.5) are randomly and *independently* drawn from a normal distribution. For population growth, this

would mean that each increment of growth would be determined independently of previous increments. However, real populations, like real organisms, do not grow that way. A chance fluctuation early in time influences future population size. Consequently, a population experiencing, by chance, slow growth initially may have a trajectory appreciably below that of a population that happens to "get a good start".

The consequences of this phenomenon appear quite surprising. Eberhardt and Simmons (1992) used stochastic models of population growth for several species of large mammals to study behavior of the confidence limits generated from log-linear regressions on such data. Their results were expressed in terms of "coverage" of calculated confidence limits. That is, confidence limits calculated for log-linear regressions on each of 1,000 runs of a given population model were examined to see whether they included the true rate of population change expected from the population parameters used.

If confidence limits from log-linear regression applied to population growth data behaved as for ordinary linear regression, the anticipated coverage would be 95%. In the simulations, it was about 60%. This raises doubts about confidence limits for the rate of growth determined from log-linear regression. In practice, of course, other factors influence the observed variability, including sampling errors of the measurements of population size and year-to-year fluctuations of the actual rate of growth associated with environmental conditions.

The simulations of Eberhardt and Simmons (1992) did show that an accurate estimate of the rate of change can be obtained from trend data, so that estimating a rate of change from trend data complements estimates from reproductive and survival data very nicely (Fig. 11.12). The main problem is one of how to make comparisons between estimates from the two sources. Some unpublished simulations indicate that coverage from bootstrapping confidence limits based on the Lotka-Leslie model is very close to the expected 95%, so that, if there are no biases in the survival and reproductive estimates used in generating the estimate of λ , the confidence limits from bootstrapping provide a useful tool for further analysis.

It should be noted that there is a theoretical answer to the problem of poor coverage of confidence limits from loglinear regression on population trend data, but one that is impractical for population studies. The approach is that taken in studies of growth of individual organisms, where the same problem exists. In that case, one simply observes growth rates for a number of individuals, and confidence limits for the mean growth rate are based on the rates for the individuals. However, it is seldom feasible to study a number of independent populations under the same conditions, so using "replications" isn't a useful solution.

One further outcome of the simulations of Eberhardt and Simmons (1992) is useful. This is that confidence limits on projections of given populations a few years into the future did give acceptable coverage. It is thus evident that trend data can give useful estimates of rates of change, and provide worthwhile confidence limits on short-term projections of population size. The implication then is that a record of trend data provides useful information on variability for the observed population trajectory, but not for

other trajectories that might arise from the underlying survival and reproductive rates.

Consequently, if we have estimates of rate of change from reproductive and survival data and from trend data, it appears worthwhile to ask whether the estimate from trend data can be considered to be one realization of the many possible population trajectories that might arise by chance from the observed reproductive and survival data. One approach is to do a t-test. A difficulty is that the variance of the trend data estimate can be expected to be different from that obtained from reproductive and survival data. This leads to what statisticians know as the Behrens-Fisher problem. An approach using bootstrapping has been suggested by Efron and Tibishirani (1993).

12.3 Some growth curves

A number of equations other than the exponential have been used to represent the growth of populations, usually for laboratory studies or in the case of introduction into new habitat or after substantial reductions in numbers. Most of the curves are "sigmoid" or S-shaped with the initial stage characterized by nearly exponential growth, followed by a gradual tapering-off towards a constant, or asymptotic level. The curves to be described here are all deterministic, that is, do not take into account chance or stochastic fluctuations in numbers. Stochastic versions are available for several of the curves, but the details are mathematically complex.

All of these curves are also applicable to the study of growth of individual organisms, and several were originally developed for that purpose. Thus we may take the dependent variable, y , as representing either population size or the weight of an individual or average weight of a group of individuals, all presumably being of the same age. Since the curves are all continuous in form, we are effectively supposing the population to be large enough that it doesn't matter whether we use a continuous or discrete representation. This is also implied by the use of a deterministic as opposed to a stochastic model.

Since there are a substantial number of growth curves available, one needs some criteria for choice in particular circumstances. This is a problem that is not resolved. Both the growth of populations and that of individual organisms are sufficiently complex and variable to prevent a solution on purely theoretical grounds, and the choice is often one of convenience for the needs of the moment. Although theoretical bases are available for most of the curves, discussion here will largely be limited to one simple criterion - the rate of change per unit of time. Thus for exponential growth the rate of change is a constant fraction of y , that is, for a small increment of time (Δt) the incremental change in y (Δy) is proportional to y :

$$\frac{\Delta y}{\Delta t} = ry$$

so that, for population growth, this can be described as the difference between births, b (or recruits) per head and loss rate (d), giving

$$\frac{\Delta y}{\Delta t} = by - dy = (b-d)y = ry$$

If a continuous variable is involved, the differential notation is normally used, i.e.,

$$\frac{dy}{dt} = ry$$

It may be more convenient to study a given set of data in terms of a relative rate of change,

$$\frac{dy}{ydt} = r$$

so that one examines the rate of change divided by the current size. In examining a given set of data, one might thus decide on a convenient (but short) time interval, Δt , and determine whether the corresponding changes, Δy , are nearly constant fractions of current size (y). Of course the assumption of exponential growth can be examined much more readily by simply taking logarithms (or plotting on "log-log" paper), but such a simple approach is not available for many of the other growth curves. Such curves can often be fitted by non-linear least-squares (available in a number of commercial computer "packages"). When these routines fail to converge (or give dubious results), it is worthwhile to start over with guesses as to the parameters. Often the first derivatives given below can be used to obtain such starting values.

The remainder of this section will be devoted to a listing of a number of growth curves, the first derivative (dy/dt) and a few remarks. Most of the curves can be written in several forms. Those used here are largely as given by Grosenbaugh (1965). An extensive set of models for analysis of fish growth and survivorship is available in Schnute and Richards (1990). All of the curves described here (excepting, of course, the exponential) have an upper asymptote (denoted by A) which is approached as t becomes very large. All but one are sigmoid, or S-shaped. The exception is sometimes known as the "monomolecular" curve, and has the equation:

$$y = A(1 - e^{-Bt}) \quad (12.8)$$

Replacing e^{-Bt} by the first two terms in the series expansion ($e^{-Bt} \doteq 1 - Bt$) gives an approximation for small values of t :

$$y \doteq Abt$$

which shows that the curve starts out as nearly a straight line. As t becomes large, y gradually approaches A , so we have something like an inverted J, or a curve that is concave downwards. The first derivative (rate of change) can be written as:

$$\frac{dy}{dt} = B(A - y) \quad (12.9)$$

Thus when y is small the rate of change is nearly constant, indicating a straight line as suggested above. As y increases the slope diminishes, and ultimately reaches zero at the asymptote. For a given set of data, an investigator could calculate Δy for some small fixed Δt and expect that a plot of Δy against y would yield approximately a straight line:

$$\Delta y \doteq BA - By$$

and thus suggest the applicability of equation (12.8) as a model. In this case, however, the general shape of the curve should also give a first indication as

to its suitability. This curve has been used to represent the upper portion of growth curves, by simply disregarding the early stages of growth. One way to do this is to write the curve as:

$$y = A(1 - e^{-Bt}) + C$$

so that when $t = 0$, $y = C$ where C is the first value to be considered. Eq. (12.8) is also often useful if one wishes to fit curves to reproductive data as shown in Fig. 11.4. Sometimes it is possible to fit eq.(11.8) directly with non-linear least-squares, but this approach often will fail due to the number of parameters that need to be estimated. This is especially true when there is little data on senescence, as usually is the case. One may then attempt to approximate the curve in sections. Eq. (12.8) represents the left-hand side of the curves (Fig. 11.4), except that a constant (c) appears in eq. (11.8). This constant effectively represents the age at which an appreciable amount of reproduction is first observed (e.g., age 4 in the fur seal data of Fig. 11.4). One can thus obtain a useful guess at c and take a as the reproductive rate observed for prime-age individuals, and then fit the left-hand portion of the curve by non-linear least-squares.

Equation (12.9) indicates that the rate of change depends on y (and the constants A and B) but not on t -- that is growth depends only on the size already achieved, and not on time. A curve that brings in a dependence on time also, is named after the mathematician Gauss, has the following form:

$$y = A(1 - e^{-bt^2}) \quad (12.10)$$

and has the first derivative:

$$\frac{dy}{dt} = 2bt(A - y) \quad (12.11)$$

so that we have the rate of change again decreasing in proportion to size of y (just as in equation (12.9)), but there is an opposite effect due to the value of t .

Equations (12.8) and (12.10) may also be written with another constant (C) as multiplier for the exponential term:

$$y = A(1 - Ce^{-bt})$$

$$y = A(1 - Ce^{-bt^2})$$

which provides more flexibility in fitting data, for which the price paid is less assurance that the particular curve considered is somehow unique or "appropriate". In neither case does C appear in the first derivative.

The logistic (Verhulst) curve has probably been more widely used than any of the others. The equation is:

$$y = \frac{A}{1 + Ce^{-Bt}} \quad (12.12)$$

and the derivative is:

$$\frac{dy}{dt} = y(A - y) \quad (12.13)$$

which shows again a dependence on the approach to the asymptote through $(A - y)$, but now modified by the value of y . A plot of $\Delta y/\Delta t$ against y will now give a curve (a quadratic or second degree curve):

$$\frac{\Delta y}{\Delta t} = By - \frac{B}{A} y^2$$

In fisheries management, an important concept is that of a stock-recruitment curve, which compares the number of recruits to a fishery with the existing stock. These curves can be considered as growth curves. Because recruitment is often an annual or generational event, the underlying equations can best be expressed as difference equations, rather than the differential equations as in the other growth curves described here. Two stock-recruitment curves have been widely used, one being the Beverton and Holt curve (Beverton and Holt 1957) and the second due to W.E. Ricker, who described both curves in detail in his 1975 book (Ricker 1975). It can be shown that the Beverton and Holt curve can be written as a difference equation form of the logistic curve [eq. (12.12)] and that the Ricker curve approaches the Beverton and Holt curve under limiting conditions (Eberhardt 1977c). The Ricker curve is nonetheless different from the logistic and is worth separate listing here, being conveniently written as:

$$N_{t+1} = AN_t e^{-BN_t}$$

where N_t is the population size at time t (often at generation t , as for Pacific salmon where distinct generations are observed) and A and B are parameters.

A rather more complicated curve with an evident similarity to the logistic is the Pearl-Reed curve:

$$y = \frac{A}{1 + Qe^{-(Rt + St^2 + Ct^3)}} \quad (12.14)$$

which has the derivative:

$$\frac{dy}{dt} = \frac{y}{A} (A-y)(R + 2St + 3Ct^2) \quad (12.15)$$

so that the rate of change has the logistic's dependence on y modified by another second-degree equation in t . This is a difficult curve to work with, but is included here to show the range of possibilities generated through a variety of assumptions about the rate of change and all built around the deviation from the asymptote, $y-A$.

A popular curve for representing growth of individuals is named after Gompertz, whose interest was, however, in studying mortality curves for actuarial purposes. The equation is:

$$y = A \exp(-Ce^{-Bt}) \quad (12.16)$$

with derivative:

$$\frac{dy}{dt} = BC ye^{-Bt} \quad (12.17)$$

This is evidently a departure from the rates-of-change thus far considered. An inspection of data may be most convenient in terms of a relative rate of change.

$$\frac{dy}{ydt} = BCe^{-Bt}$$

which might conveniently be plotted on "semi-log" paper (i.e., plot $\Delta y/y$ against t) where it should appear as a straight line. Referring back to equation (12.13) it appears that the relative rate of change of the logistic curve should provide a straight line in arithmetic coordinates:

$$\frac{dy}{ydt} = B - \frac{B}{A} y$$

so that the two curves (logistic and Gompertz) provide a rather distinct contrast.

Growth of individual fish has been represented by a curve developed by Von Bertalanffy, with the equation being:

$$y = A(1 + Ce^{-Bt})^3 \quad (12.18)$$

This curve has the derivative:

$$\frac{dy}{dt} = 3B(A^{1/3} y^{2/3} - y) \quad (12.19)$$

which is not so readily compared with the other forms. However, the original rationale for the curve does give an interesting interpretation. Von Bertalanffy proposed that the rate of growth of a fish depends on the difference between anabolism and catabolism, with anabolism being proportional to surface area and catabolism to body weight. If surface area is taken as proportional to weight over length, an approximation then is weight to the $2/3$ power and we have:

$$\frac{dw}{dt} = k_1 w^{2/3} - k_2 w$$

where the k_i are constants, and we thus have the general form of equation (12.19). An interesting, and useful, sidelight is that the relationship can be approximated in terms of the length of fish, giving an equation equivalent to (12.10). An alternative way to write equation (12.18) is:

$$y = A[1 + e^{-B(t-t_0)}]^3$$

so that t_0 becomes a parameter replacing C . This can also be done for the other equations involving C as a multiplier of the exponential term. Since t_0 is a constant, one can write:

$$e^{-B(t - t_0)} = e^{Bt_0} e^{-Bt} = Ce^{-Bt}$$

where $C = e^{Bt_0}$. The advantage of using t_0 is that it provides an explicit symbol of the fact that this constant lets one shift the curve along the time axis.

Another quite different growth curve is the Johnson-Schumacher equation, which has been used mostly for tree growth:

$$y = A \exp\left(-\frac{B}{t+c}\right) \quad (12.20)$$

The derivative is:

$$\frac{dy}{dt} = \frac{By}{(t+c)^2} \quad (12.21)$$

and this provides another variant on the relationship between rate of change, time, and current level.

12.4 Projection models

Many population studies must deal with the situation where appreciable numbers of individuals are removed from a population, often annually, but also on a less-regular basis. Dealing with such situations accurately requires estimates of *absolute* numbers in the population and of removals. The models used here also require that the removals take place in a relatively short time interval, to avoid the complications of appreciable numbers of deaths from natural causes during the removal period.

Two models may be considered:

$$\text{I} \quad N_t = N_{t-1}\lambda - K_t \quad (12.22)$$

$$\text{II} \quad N_t = (N_{t-1} - K_t)\lambda$$

In essence, model I assumes that removals, K_t , take place just before the census at time t , while the second model assumes that removals take place just after the census at $t-1$. Annual censuses are assumed in both cases. If we let S_1 denote survival from $t-1$ to reproduction, and S_2 denote survival from reproduction, R , to removal, K_t , then Model I can be written as:

$$N_t = (N_{t-1}S_1)RS_2 - K_t$$

so that $\lambda = RS_1S_2$, including reproduction and survival through two periods of natural mortality. However, if there is an appreciable loss between removal and the second census, then another survival rate needs to be introduced, and a bias is introduced.

Hence if we have estimates of absolute population size, and known removals in a short time period, the two models offer a way to estimate λ in the presence of removals. A very convenient way to proceed is to use a ratio estimate. In Model I:

$$\hat{\lambda} = \frac{\Sigma(N_t + K_t)}{\Sigma N_{t-1}} \quad (12.23)$$

Where the summation is over successive observations. An alternative is to use a least-squares approach. A difficulty is that the estimates of population size are serially correlated, i.e., N_{t-1} becomes N_t in the next time period. Eberhardt (1987) examined this problem by testing various approaches on sets of population growth data that did not involve removals. With such data it is possible to estimate the rate of change by log-linear regression, as discussed in Section 12.1. The ratio and least-squares (log-linear regression) estimates gave consistent results, as shown in Fig.12.1.

With such a high correlation between the two methods, it would appear that the ratio method is likely to give useful results. A remaining problem, however, is one of obtaining useful variance estimates, due to the serial correlation issue. One way to approach the problem is to use jackknifing, in which a set of estimates is obtained by dropping each of the n items from which λ is estimated in turn, and forming n estimates of λ from the sets with one item deleted. A variance is then computed from the n estimates. A comparison (Fig. 12.2) between the jackknife variance estimates and that from the least-squares fit suggests that the jackknife method gives comparable estimates, but can't always be depended on.

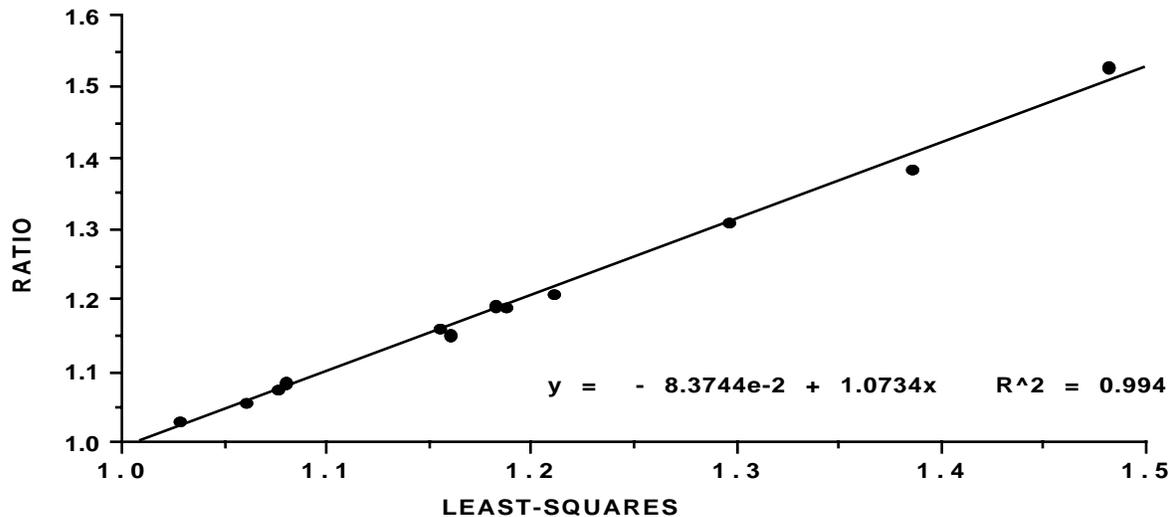


Fig. 12.1. Relationship between estimates of λ from the ratio method and from least-squares. Data from Eberhardt (1987:Table 2). 1:1 line shown.

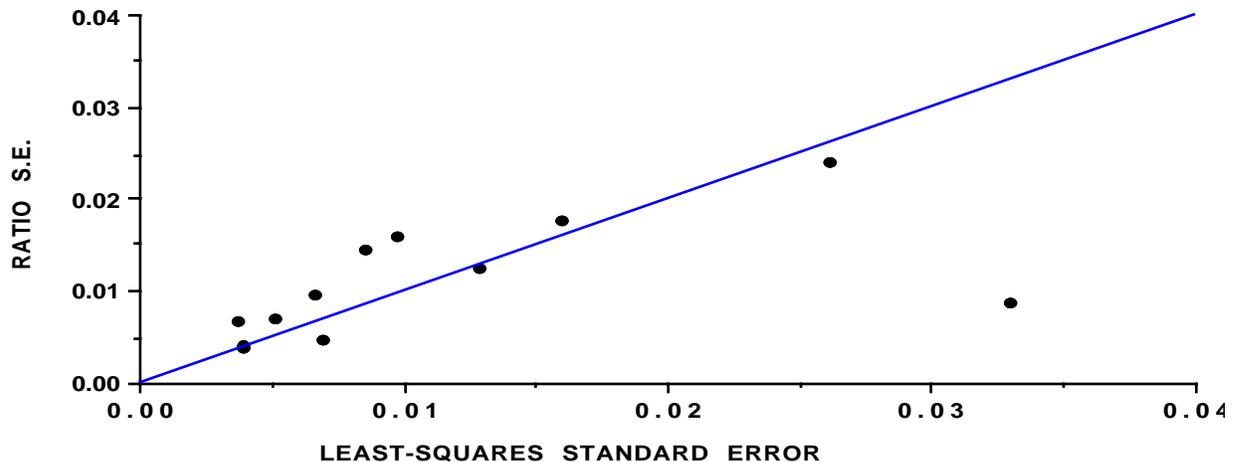


Fig. 12.2. Comparison of estimates of variance for λ obtained from jackknifing (Ratio S.E.) and estimates from a least-squares fit to the data. Data from Eberhardt (1987:Table 2). 1:1 line shown.

From the examples mentioned above, it appears that the projection models may be quite useful in various circumstances. In some instances, reproductive and survival data may also be available, so that an estimate of the

value of λ occurring without removals is also on hand. It will then be important to know whether an estimate of λ obtained from the removal data, as described above, is compatible with that obtained from the Lotka-Leslie approach.

12.5 Predator-prey models

Various simple models for predator-prey interactions have been discussed and explored mathematically for over 70 years. The initial models were two simple differential equations proposed by Lotka (1925) and Volterra (1931). Very little attention has been given to attempts to fit coefficients from actual field data, or to use such data to assess individual components of the models. Some prospects are explored in Section 12.6 by using data on wolves and their prey.

In practice, we usually need to deal with "birth-pulse" populations (Caughley 1977:6) where reproduction occurs in a relatively short period each year. The population then decreases until the next birth period, when it again gets a boost upwards. This sort of behavior may approximately be described by a "step-function", i.e., by a graph looking like a stairway, perhaps with steps of uneven size (cf. Fig. 11.1). One may thus use difference equations rather than differential equations, and the models used here are all computed as difference equations. Much of the recent literature is based on differential equations, which are not appropriate for birth-pulse populations because reproduction does not occur continuously throughout the year.

We first show the relationship of difference equations to differential equations, using the logistic equation (previously given as eq. (12.12) but shown here with the parameters familiar to ecologists):

$$\frac{dV}{dt} = rV\left[1 - \frac{V}{K}\right] \quad (12.24)$$

with K being the asymptotic value. Differential equations are often developed as a limiting expression, where some small increment of time approaches zero. For present purposes, we consider the increment as unity (one year, one day, etc.) and write:

$$\frac{dV}{dt} \approx V_t - V_{t-1} = rV_{t-1}\left[1 - \frac{V_{t-1}}{K}\right]$$

Rearranging gives a difference equation:

$$V_t = V_{t-1} + rV_{t-1}\left[1 - \frac{V_{t-1}}{K}\right] \quad (12.25)$$

Many of the differential equations of interest in ecology have no explicit solutions, but the logistic does:

$$V(t) = \frac{K}{1 + ce^{-rt}} \quad \text{where } c = \frac{K - x_0}{x_0} \quad (12.26)$$

with x_0 = initial value, here taken to be unity.

The logistic model for population growth assumes continuous growth, at an instantaneous rate, r . If we take $r = 0.8$ and $K = 30,000$, then we can plot eq.(12.26) for, say, 30 years. To approximate it with difference equation (12.25), we need to subdivide the year into increments. If we use 20 such increments, then $r_1 = \frac{r}{20} = \frac{0.8}{20} = 0.04$. Fig. 12.3 shows, the difference equation provides a reasonable approximation to the differential equation, with as few as 20 increments per unit of time.

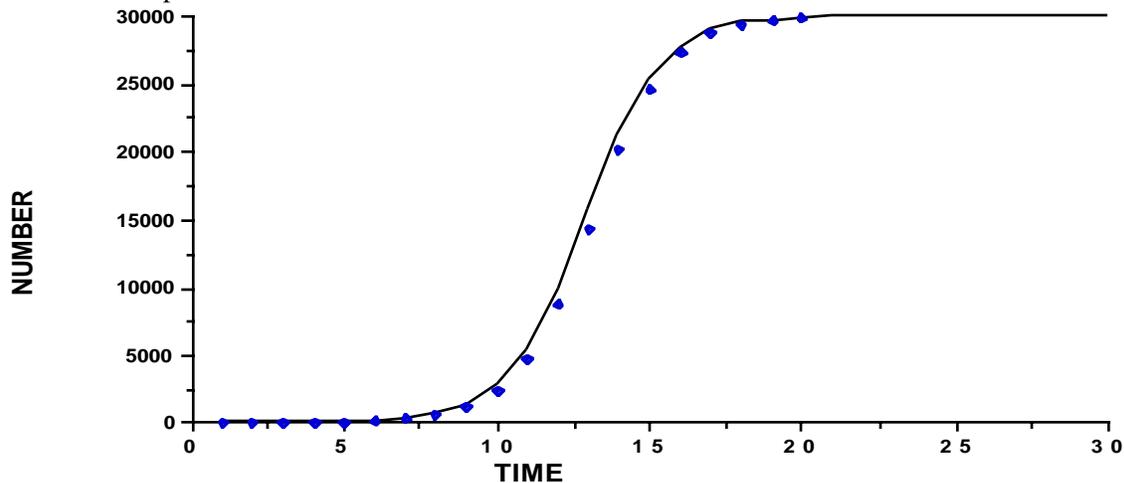


Fig. 12.3. Graph of a logistic equation (line) compared to values (points) from a difference equation version of the underlying differential equation.

Usually, we expect the behavior of the predator population to depend on the abundance of the prey, and it also may be subject to removals by man. The terms "prey" and "predator" often refer to animals, but the models can also serve to represent herbivores and vegetation. Consequently, we use the letters H for herbivores or predators, and V for vegetation or prey. The prey equation usually contains provision for self-limitation or a "density-dependence" term, often the logistic model given above, and then a term representing removals by the predators. This is termed the "functional response" and is denoted below as a generalized function, $F(H,V)$. The predator equation contains a term showing the way in which the predator supposedly responds to prey abundance, and this is termed the "numerical response" and denoted by $G(H,V)$. It is worthwhile to quote the original definition of these terms (Solomon 1949) inasmuch as some discussions in the literature appear to deviate appreciably from that definition. Solomon uses "natural enemy" to cover both predators and parasites:

"to be density-dependent, the enemy must respond to changes in numbers of the host The nature of this response is commonly twofold. First, there must be a functional response to (say) an increase in the host density, because of the increased availability of victims: as host density rises, each enemy will attack more host victims, or it will attack a fixed number more rapidly. A frequent, but not invariable result of this is an increase in the numbers of the enemy (a numerical influence) due to an increased rate of survival or of reproduction, or of both; this may or may not be sufficient to produce an increase in the proportion of enemies to the increasing hosts."

The two generalized equations are:

$$\frac{dV}{dt} = rV\left[1 - \frac{V}{K}\right] - HF(H,V) \quad (12.27)$$

$$\frac{dH}{dt} = HG(H,V) \quad (12.28)$$

Where $F(H,V)$ represents the "functional response" and $G(H,V)$ is the "numerical response". In the early use of these equations, it was assumed that the functional response was proportional to the number of predators present, i.e., that $F(H,V) = \alpha V$ so that the equation becomes:

$$\frac{dV}{dt} = rV\left[1 - \frac{V}{K}\right] - \alpha VH \quad (12.29)$$

with the last term often interpreted as depending on an "encounter rate" and thus the product of the numbers of H and of V . The numerical response has been written as $G(H,V) = -d + \beta V$, and this can be interpreted as a constant mortality rate of predators (d) plus a reproductive rate depending on prey abundance (βV), with the equation then becoming:

$$\frac{dH}{dt} = (\beta V - d)H \quad (12.30)$$

There are a variety of other forms of functional and numerical responses, with a large literature dealing with theoretical interpretation of these terms. Various aspects of the theory appear in May (1981), and some of these are discussed in Section 12.7 below.

With this background, we can take a look at a system variously used to describe vegetation-herbivore interactions (Caughley 1976, 1977, Caughley and Lawton 1981), in which the herbivore has been described as an "ungulate" or as typifying "white-tailed deer colonizing a mosaic of grassland and forest" (Caughley and Lawton 1981). In the several examples, the relevant equations are given as differential equations, which are here translated to difference equations:

$$V_t = V_{t-1} + r_{m1}V_{t-1}\left(1 - \frac{V_{t-1}}{K}\right) - c_1H_{t-1}(1 - e^{-d_1V_{t-1}}) \quad (12.31)$$

$$H_t = H_{t-1}[1 - a_2 + c_2(1 - e^{-d_2V_{t-1}})] \quad (12.32)$$

The exponential terms, $(1 - e^{-d_1V_{t-1}})$, are supposed to adjust the herbivore's intake of vegetation and population growth rate according to density of vegetation.

Plots of vegetation and herbivore density show rather dramatic changes in the first 20 years, and come nearly to equilibrium in 50 years (Caughley 1977:Fig. 9.6, Caughley and Lawton 1981:Fig. 7.3). If we plot the difference equations (12.31) and (12.32) using the constants given by Caughley (1977:129), we get very pronounced and continuing oscillations (Fig. 12.4).

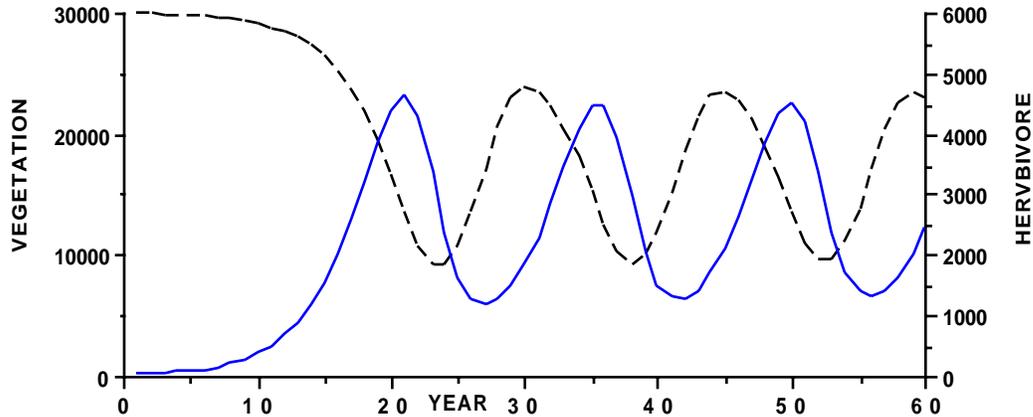


Fig. 12.4. Caughley's (1977) vegetation-herbivore system computed as a difference equation in which reproduction occurs annually.

Why the difference? The answer lies in the way in which the equations interpret reality. The difference equations assume reproduction occurs just once a year for the herbivore and for vegetation. The differential equations assume reproduction goes on throughout the year. Caughley (1977:130) remarked that "the population has been grown in a programmable desk calculator by estimating growth curves twenty times per year and adding on the appropriate increment each time". However, ungulates don't reproduce 20 times a year, behaving instead like the difference equations, producing young once a year. Since vegetation does grow continuously over part of the year, a somewhat different model presumably could be used for vegetation. If we follow Caughley's prescription cited above, using difference equations and rates divided by 20, then the difference equation model reflects his Fig. 9.6 reasonably well, but the time scale is now multiplied by 20. In effect, if reproduction occurs only once a year, then the curve shown by Caughley takes something like 500 years, not 50 years.

The main point to be made here is that one needs to be sure that the equations used do reflect the biology of the situation at hand. Some other problems with the system discussed above are described by Eberhardt (1988). Another example concerns an effort to consider the role of interactions between species in the management of multispecies fisheries. May et al. (1979) used differential equation models to illustrate the possible impacts of commercial harvesting of the main food supply (krill) of baleen whales in Antarctic waters. They stated that "A crude Lotka-Volterra form of predation is assumed, with prey being consumed at a rate proportional to their density, aN_1 , per predator". Written as a difference equation, this gives:

$$V_t = V_{t-1} + r_1 V_{t-1} \left(1 - \frac{V_{t-1}}{K}\right) - c_1 H_{t-1} V_{t-1} \quad (12.33)$$

while their predator equation is:

$$H_t = H_{t-1} + r_2 H_{t-1} \left[1 - \frac{H_{t-1}}{aV_{t-1}}\right] \quad (12.34)$$

The problem is that baleen whales are very unlikely to have evolved to each take a proportional share of the available prey. Instead they quite clearly take as much prey as needed to supply the individual whale, so that the last term in eq. (12.33) becomes $c_1 H_{t-1}$. Such a change makes an appreciable difference in trend of the predator and prey curves, and conclusions about equilibrium points, relative yields, and the like, as indicated in Eberhardt (1988).

12.6 A wolf-ungulate model

As noted above in connection with eq. (12.33), it is unlikely that an effective large predator will take a proportional share of the available prey. A further modification of eq.(12.33) is available in the form of the generalized logistic in which the growth rate may not begin to decline appreciably until the population approaches its asymptotic value (K). With these modifications, eq. (12.33) becomes:

$$V_t = V_{t-1} + r_1 V_{t-1} \left[1 - \left(\frac{V_{t-1}}{K} \right)^z \right] - c H_{t-1} \quad (12.35)$$

The general form of eq.(12.34) has been recently popular under the label of "ratio dependence" (Matson and Berryman 1992). Eberhardt (1997) used data on moose and caribou from the literature to show that the functional relationship very likely can be reduced to the constant, c , of eq.(12.35). It is worth noting here that the equilibrium values (obtained by setting $H_t = H_{t-1}$ and $V_t = V_{t-1}$ are:

$$V = K \left[1 - \frac{ca}{r_1} \right]^{1/z} \quad \text{and} \quad H = aV \quad (12.36)$$

The best available data are those for moose. Eberhardt (1997,2000) used data from the literature to estimate c in eq. (12.35) as 2 moose killed per wolf per 100 days in winter, with an annual rate of a little less than 7 moose per wolf. Eberhardt (1998) calculated a maximum rate of increase for moose as $\lambda = 1.38$. The parameter z is not well established, but was used as $z = 5$. Carrying capacity (K) depends on the specific population considered and thus may be selected arbitrarily here. For eq. (12.34) the maximum rate of increase for wolves was estimated as $\lambda = 1.48$ (Eberhardt 1998) and the ratio-dependence constant as $a = 1/20$ wolves/moose (Eberhardt and Peterson 1999). Using these constants, and starting with a moose population of 30,000 (at carrying capacity, K) and a small initial wolf population (20) gives the results of Fig. 12.5, which can be compared to Fig. 12.4. A small initial fluctuation soon disappears and both moose and wolves settle down to steady-state numbers at the assumed ratio of 20 moose per wolf.

An obvious feature of Fig. 12.5 is the absence of the oscillations that were induced in the model of Fig. 12.4 by the choice of mathematical forms for the functional and numerical responses. There is a good deal of evidence that such oscillations do in fact occur in insect populations with their generally high rates of increase, but wolves and their prey have much lower rates of increase so that the models derived for insects and other species with high rates of increase ought not to be arbitrarily assumed to be appropriate as seems to have been done in the literature.

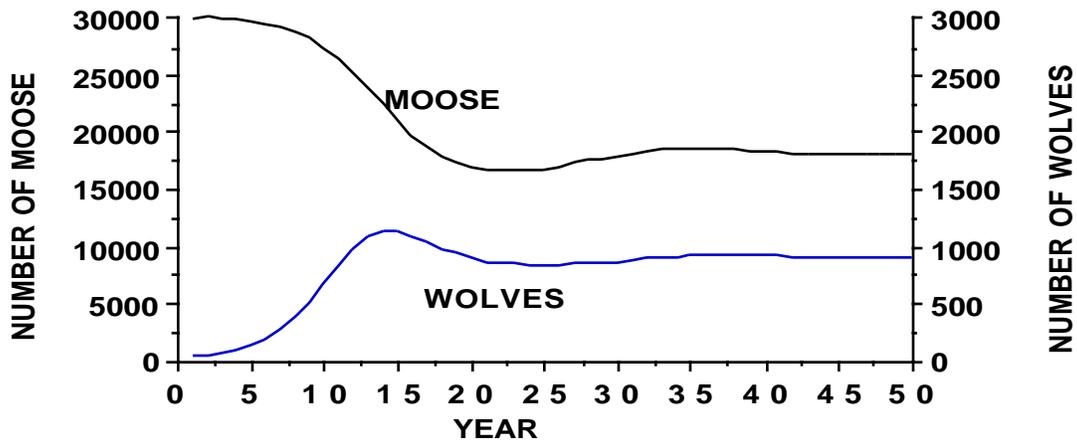


Fig. 12.5 Trend of moose and wolf populations using Eq. (12.34) and (12.35) with the parameters given above.

Another carryover from entomological studies is the notion of a “total response” model in which the predator population is assumed to respond instantaneously to changes in prey abundance. Using such a model one can conveniently depict predator abundance as a function of prey abundance and produce a variety of models along the lines of Messier (1994). However, wolves do not respond instantaneously to fluctuations in prey numbers, and hunting, trapping, and “control” actions have resulted in non-equilibrium numbers of predator and prey in most real-world situations. A realistic approach thus needs to depend on equations like (12.34) and (12.35) that show non-equilibrium conditions. The “total response” model traces back to Holling (1959) who warned about the consequences of assuming that model as follows: “The method is an over-simplification, since predator density is portrayed as being directly related to prey density. Animal populations, however, cannot respond *immediately* to changes in prey density, so that there must be a delay of the numerical response” and “the total response obtained when prey or hosts are steadily increasing will be different than when they are steadily decreasing. The amount of difference will depend on the magnitude and amount of delay of the numerical response, for the functional response has no element of delay”.

An interesting aspect of the model of eq. (12.34) and eq. (12.35) is that it can be fit directly to actual data, using multiple regression. An attempt to do so for data on Isle Royale moose and wolves appears in Eberhardt (1998:Fig. 2). Unfortunately, there are few data with enough years of observations on both prey and predators to make such an approach widely applicable and both moose and wolves on Isle Royale appear to have been subjected to influences not considered in the simple model. This illustrates the major dilemma in model-building in ecological studies. Usually only a few parameters can be unambiguously estimated from actual data, forcing the use of simple models that cannot accommodate unexpected changes due to environmental or other factors. A further possible difficulty with the simple model of eq. (12.34) is the assumption that wolf abundance is limited solely by moose abundance through the ratio-dependence parameter (a). In reality, wolf packs defend more or less

exclusive territories, a factor that may limit their abundance in the presence of high prey densities.

12.7 Assessing differential equation models

This Section is provided to supply an indication of how differential equation models for predator-prey studies can be evaluated. The basic problem is that most such equations cannot be “solved”, i.e., integrated. Section 12.5 used the logistic equation as an example of a differential equation that can be solved directly and showed that the difference equation analog can be used to provide a good approximation to the behavior of the differential equation. This is the simplest approach for studying differential equations when direct solutions are not possible. Quite a bit can be learned about the equations by plotting trend of the populations on a diagram showing the “isoclines”. We can illustrate the basics by considering a modification of the original predator-prey equations due to Lotka (1925) and Volterra (1931). The original equations are:

$$\begin{aligned}\frac{dV}{dt} &= r_1V - bVH \\ \frac{dH}{dt} &= cVH - dH\end{aligned}$$

where V denotes prey and H denotes predator as before. According to May (1981) “This system has pathological dynamical properties...”, and we will not consider the equations in the above form further here. A major problem from the biologist’s point of view is that, when no predators are present ($H = 0$), the prey population will grow continuously, without limit. Hence early workers made the modification given by eq. (12.29) which introduces the logistic equation as a control on the prey rate of increase if no herbivores are present, i.e.,

$$\frac{dV}{dt} = r_1V\left[1 - \frac{V}{K}\right] - bVH$$

One approach to studying differential equations of this type is to plot the trend of equilibrium solutions, i.e., let $dV/dt=0$, and thus:

$$H = \frac{r_1}{b}\left[1 - \frac{V}{K}\right]$$

so that, given values of the constants, one can plot H against V as a straight line. The solution of the predator equation is just $V = d/c$, a constant, and thus a vertical line on the plot of H against V. The intersection of the two lines (known as isoclines) provides an “equilibrium point” or the joint solution of the two equations. This is the point at which the predator and prey populations settle down to constant values (equations are known that cycle endlessly, and one example will be given later in this Section).

To go further, we need to convert the differential equations to difference equations. This is done by replacing dV/dt by $\Delta V/\Delta t$ where Δ

represents a small increment. In the case of approximating the logistic equation in Section 12.5, we could make this increment as small as we like, but the predators and prey mainly considered here reproduce only once a year, so the time increment needs to be one year, i.e., $\Delta t=1$ and $\Delta V = V_t - V_{t-1}$, or the size of the prey population at time t minus its size at time $t-1$. The difference equation for the prey population then becomes that of eq. (12.33) and the predator equation is:

$$H_t = H_{t-1}[1 - d + c_2 V_{t-1}] \quad (12.37)$$

This is the same as equation (12.30) but now is written as a difference equation. These equations are readily plotted in EXCEL, using values supplied here. For convenience in making comparisons and for discussion purposes, parameters used in the following examples are standardized so as to give roughly equivalent equilibrium values and to approximate values for moose and wolves used by Eberhardt (1997, 1998) and Eberhardt and Peterson (1999). The equilibrium values are approximately 4000 moose and 200 wolves, while initial values are 8000 moose and 50 wolves. Where an asymptotic prey value is needed, it is set at $K = 10,000$. Maximum rate of increase for prey is $r_1 = 0.38$ and $r_2 = 0.48$ for predators. For eq.(12.33) the take by wolves is assumed proportional to number of prey present. This is not a very realistic assumption as wolves generally are likely to be capable of taking what they need as suggested in Section 12.6, in connection with eq. (12.35). In that equation, a constant rate is assumed and is set at $c = 5$ below. For comparability we thus assume that $c_1 V = 5$ in eq. (12.33) where V is the equilibrium moose population, so that $c_1 = 5/4000 = 0.00125$. In the prey equation [eq. (12.37)] d is regarded as an annual wolf mortality rate and set at $d = 0.40$, while c_2 denotes a kind of reproductive rate, being the gain realized from consuming moose. At equilibrium we have $c_2 V = d$, i.e., the gain to the wolf population just offsets loss (d), so we set $c_2 = 0.40/4000 = 0.0001$.

The various equations are readily computed and plotted in EXCEL, using the values supplied here. One starts out with initial conditions for predator and prey, $V_0 = 8000$ and $H_0 = 50$, and uses the parameters given above. Thus in eq.(12.33) the observation for prey at time 1 is:

$$V_1 = V_0 + r_1 V_0 \left[1 - \frac{V_0}{K}\right] - c_1 V_0 H_0 = 8000 + 0.38(8000) \left(1 - \frac{8000}{10000}\right) - 0.00125(8000)50 = 8108$$

while the predator number [eq.(12.37)] is;

$$H_1 = 50[1 - 0.4 + 0.0001(8000)] = 70$$

and at time 2 the prey number is:

$$V_2 = V_1 + r_1 V_1 \left[1 - \frac{V_1}{K}\right] - c_1 V_1 H_1 = 8108 + 0.38 \left[1 - \frac{8108}{10000}\right] - 0.00125(8108)70 = 7981.5$$

Subsequent terms are computed in the same manner, and a plot of predator and prey can be obtained as in Fig. 12.6, which shows the course of the populations over time. We can also plot predator numbers against prey numbers along with the isoclines $H = 304[1 - 304V/10000]$ and $V = 0.4/0.0001 = 4000$ getting Fig. 12.7. In this example, the two populations oscillate over more

than 100 years and ultimately settle down at the intersection of the isoclines, i.e., 4000 prey and about 180 predators.

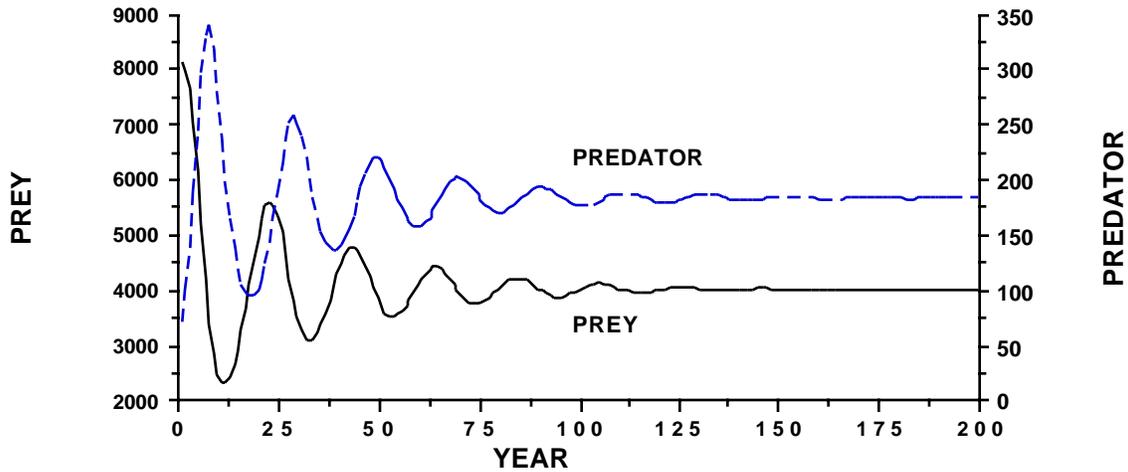


Fig. 12.6 Course of predator and prey populations calculated from equations (12.33) and (12.37).

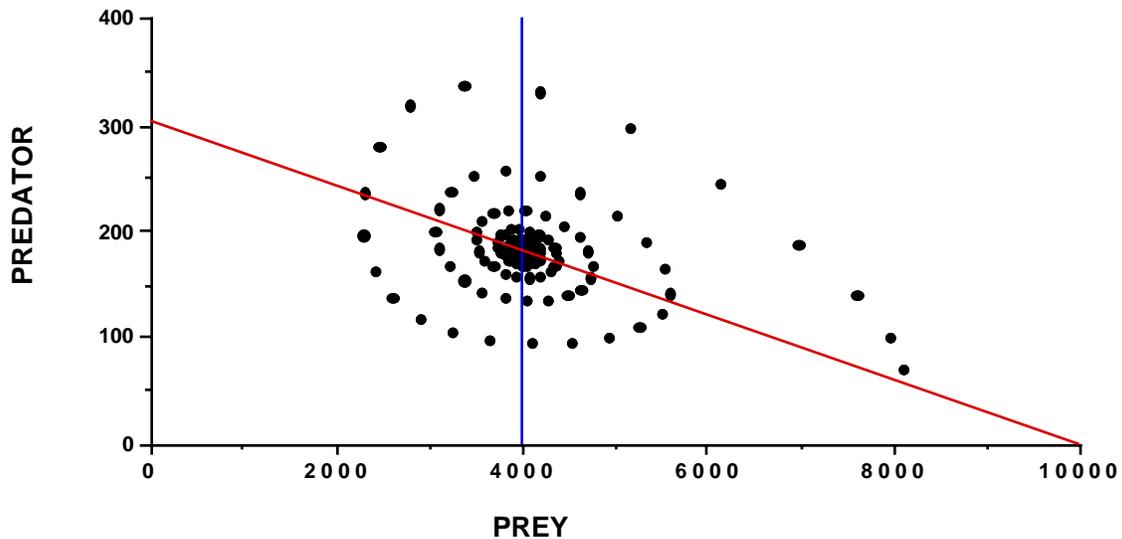


Fig. 12.7 Diagram showing isoclines (lines). Course of populations is shown by the spiral of points starting near initial values (50 predators, 8000 prey) and ending near junction of isocline lines (equilibrium values).

The dramatic fluctuations generated by the above equations have been observed for insect populations, and might well be invoked for some vertebrate populations that show cycles of varying lengths, but it is not established as yet that such observed cycles are necessarily due to predator-prey interactions. It is thus useful to look at a few equations that do not vary so dramatically. One set is that given by eqs.(12.33) and (12.34) and used by May et al. (1979) to model Antarctic food chains. A diagram (Fig. 12.8) showing isoclines for these equations shows only one simple curve to the equilibrium

point from the initial numbers (8000 prey and 50 predators), and the population plots (Fig. 12.9) show essentially one cycle before settling down.

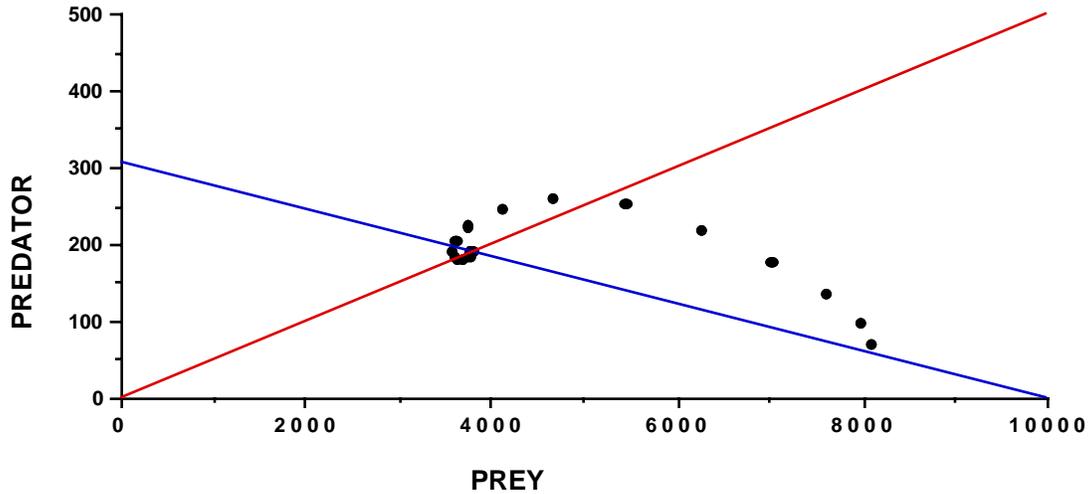


Fig. 12.8 Isoclines and trace of points for equations (12.33) and (12.34).

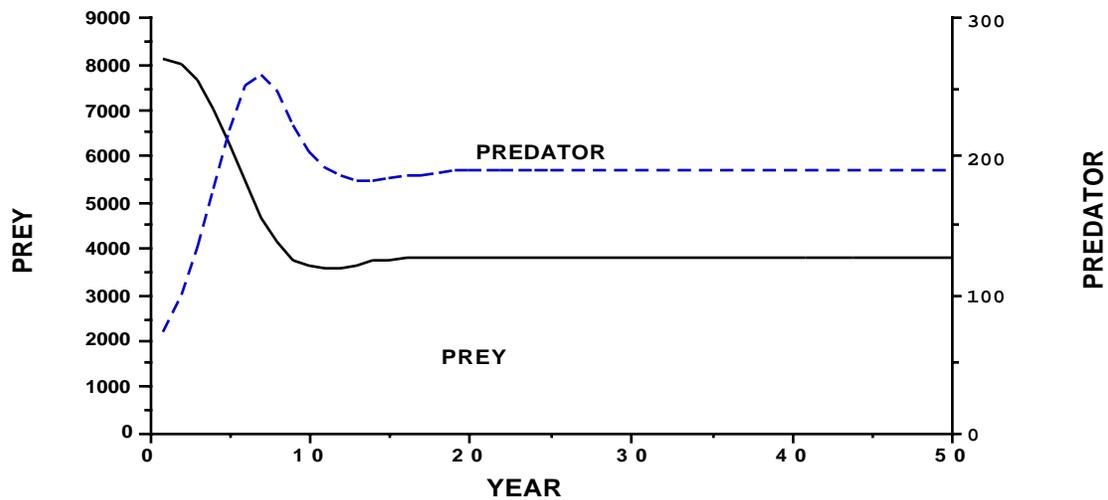


Fig. 12.9 Population trace for equations (12.33) and (12.34).

In connection with equations (12.33) and (12.34) it was remarked (Section 12.5) that the last term in eq.(12.33) very likely should not include the prey, i.e., that large predators like baleen whales and wolves most likely take what prey they need rather than taking prey in proportion to its abundance. Thus the modified equation was used as eq. (12.35) in the wolf-ungulate model of Section 12.6, with one further modification, which was to change the logistic term in the prey equations above to a “generalized logistic”:

$$\left[1 - \left(\frac{V_{t-1}}{K}\right)^z\right]$$

where the exponent, z , is equal to or greater than unity (set at unity it gives the usual logistic term). This modification is essential in that it gives prey

populations a much more nearly constant rate of increase until they approach the asymptote. If the rate of increase is permitted to decrease linearly as in the usual logistic model, then prey populations cannot stand much in the way of predation. The population course for this set of equations [(12.34) and (12.35)] is very much the same as in Fig. 12.9, but the isocline lines now differ (Fig. 12.10), in that the equilibrium solution for the prey equation produces a curve which is a second degree polynomial for $z=2$, but would be a third degree polynomial if $z=3$, and so on. Note that using $z=2$ has increased the equilibrium values substantially, due to use of the generalized logistic rather than the ordinary logistic which was used in producing Fig. 12.8

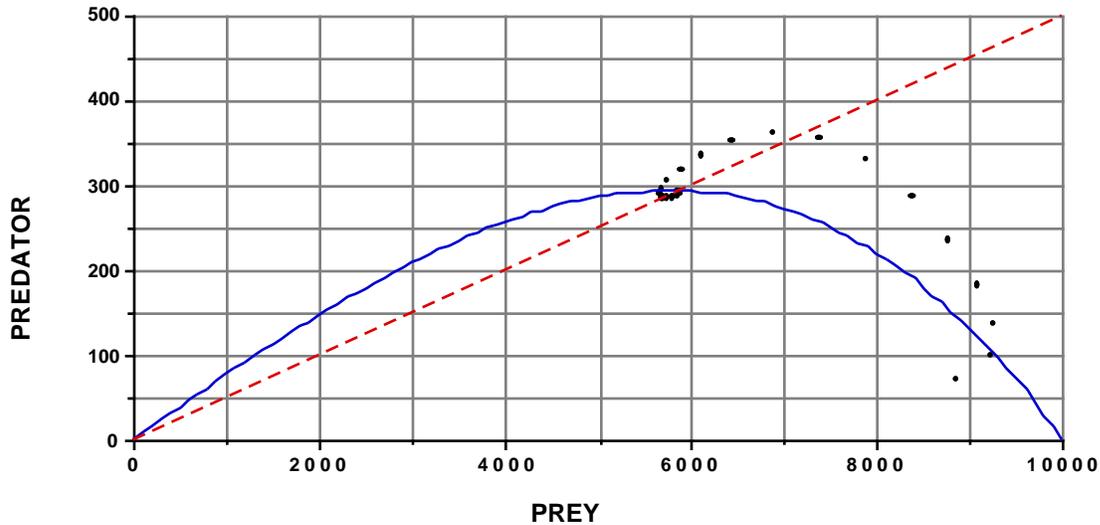


Fig. 12.10 Isocline lines and population plot for equations (12.34) and (12.35).

Another version of the prey equation may be considered at this point. This incorporates the widely-used functional relationship originated by Holling (1959), and usually described as his Type II curve. This gives rise to the following prey equation:

$$V_t = V_{t-1} + r_1 V_{t-1} \left[1 - \frac{V_{t-1}}{K} \right] - \left[\frac{m V_{t-1}}{w + V_{t-1}} \right] H_{t-1} \quad (12.38)$$

The functional relationship contains two parameters, m and w , and gives rise to a curvilinear relationship replacing the very simple constant, c , of eq. (12.35). The parameter m was set at 5 and $w = 1000$. The population trend is again a simple curve (Fig. 12.11) and the isoclines are much like those of Fig. 12.10. The Type II curve is not very much different from an exponential curve, as can be shown by series expansions:

$$\frac{mV}{w+V} = m \left(1 + \frac{w}{V} \right)^{-1} = m \left[1 - \frac{w}{V} + \left(\frac{w}{V} \right)^2 - \left(\frac{w}{V} \right)^3 + \dots \right]$$

$$m e^{-\frac{w}{V}} = m \left[1 - \frac{w}{V} + \frac{\left(\frac{w}{V} \right)^2}{2!} - \frac{\left(\frac{w}{V} \right)^3}{3!} + \dots \right]$$

When V is large, as it is for the most part here, these two expressions give very similar curves. Using $m=5$ and $w=1000$ as used here, the Type II curve is much the same as an exponential with the same parameters (Fig. 12.12). The agreement is worth noting mostly by way of suggesting that one not place too much faith in the derivations of the Type II curve found in the literature.

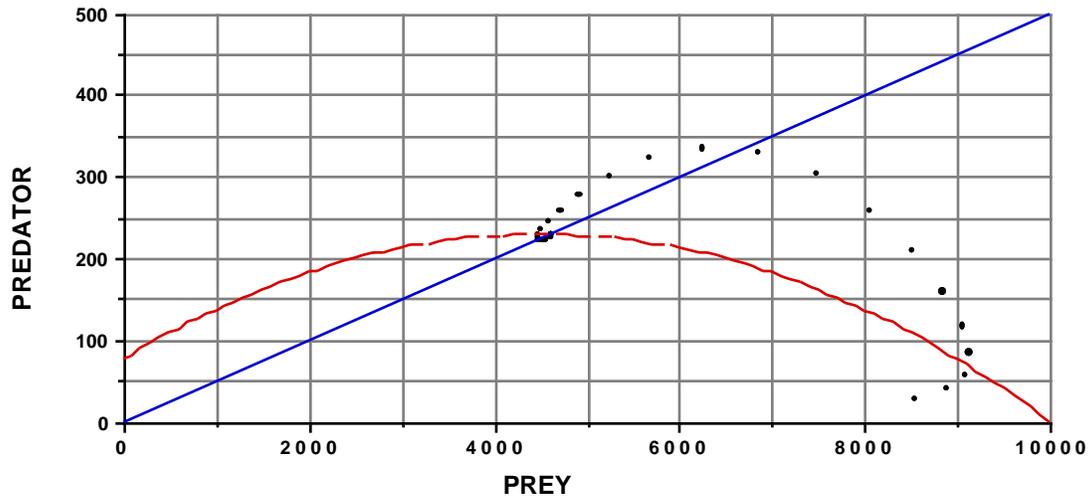


Fig. 12.11 Isoclines and population trend for equations (12.38) and (12.34).

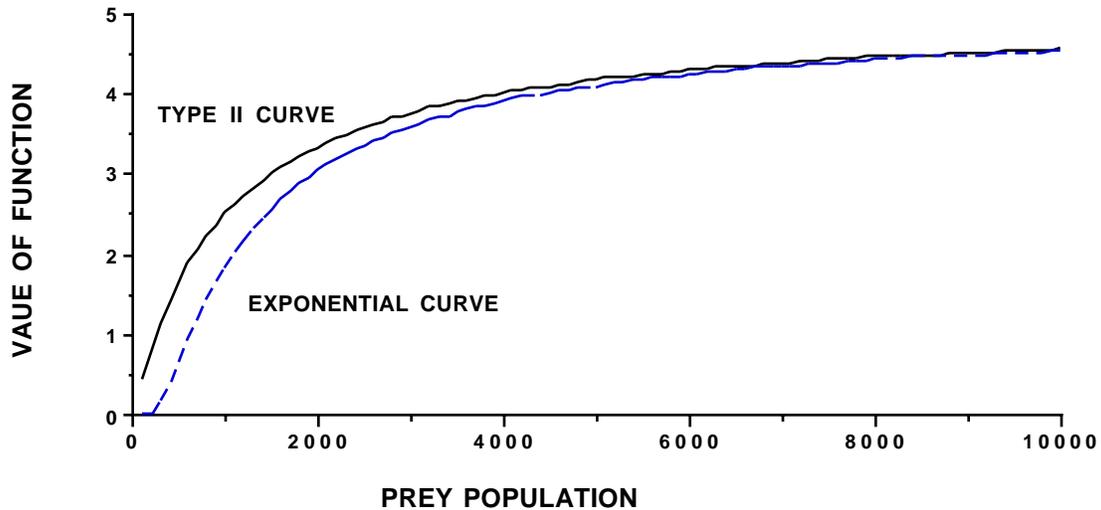


Fig. 12.12 Comparison of Hollings Type II curve and an exponential function with the same parameters.

Thus, although we have used a different model for prey, the general results are much the same as for 3 earlier cases. However, if we follow what seems to be the current trend in the literature and use the same curve for functional and numerical response, then a very different result is obtained, i.e., sustained oscillations of the form exhibited in Fig. 12.4. The prey equation now is:

$$H_t = H_{t-1} \left[1 + \frac{c_3 m V_{t-1}}{w + V_{t-1}} - d \right] \quad (12.39)$$

The combination of eqs.(12.38) and (12.39) yields dramatic and continuing oscillations and does not appear to converge on equilibrium values (Fig. 12.13). Also, different starting values of predator and prey give different results. Hence these curves are not considered further here. Readers interested in the theoretical basis for such curves should consult May (1981) and the current literature on predator-prey models. Their practical utility remains to be demonstrated.

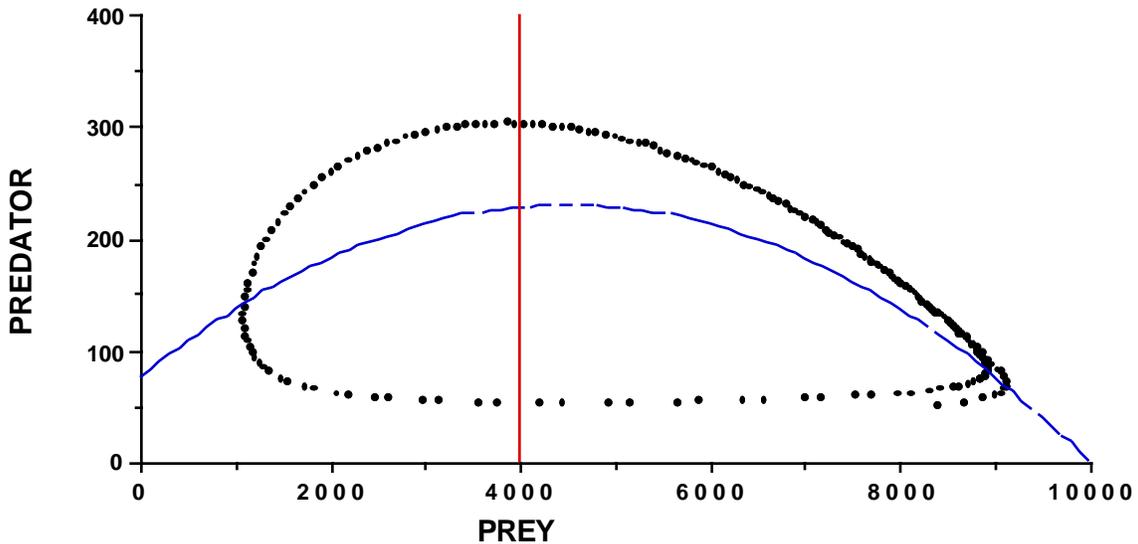


Fig. 12.13. Trace of points generated by eqs. (12.38) and (12.39).

From the results considered thus far, it appears that the ratio dependence model for predators [eq.(12.34)] yields rather stable results when combined with various models for prey abundance, and that the prey equation used for moose and wolves [eq.(12.35)] may be presently most useful for actual data on large vertebrate predator-prey studies. However, most ungulate populations are preyed on by hunters, as well as by wolves. Consequently, it is worthwhile to consider a modification of eq.(12.35) that brings in removals by hunters:

$$V_t = V_{t-1} + r_1 V_{t-1} \left[1 - \left(\frac{V_{t-1}}{K} \right)^z \right] - c H_{t-1} - R \quad (12.40)$$

Here, we assume a constant annual harvest of, say, moose by hunting, set for convenience at 290 individuals per year. Utilizing eq.(12.34) for predators, population trends are qualitatively much the same as before with the populations approaching equilibrium in about 20 years. The isocline diagram (Fig. 12.14) differs, however, in several respects from Fig. 12.10, which represents the situation without hunting removals. The equilibrium values are substantially lower with both prey and predator numbers reduced, and the isocline lines just intersecting rather than crossing. Perhaps the most interesting and instructive result comes if we increase the hunter harvest just

slightly—from 290 per year to 294 per year. This results in a crash of both populations (Fig. 12.15), after a long, rather slow decline. Of course, if the harvest is set a little larger, the crash occurs sooner. These equations should not be taken as useful direct models of actual populations, inasmuch as they are very sensitive to small changes in parameters. In reality, wolves will most likely “switch” to alternate prey if possible or leave the area entirely. The general picture is well-illustrated by the fate of the Nelchina caribou herd in Alaska, which became very large during wolf control, but ultimately crashed when wolves regained high numbers and a largely fixed hunting harvest was maintained. Some details of this event were reported by Eberhardt and Pitcher (1992). A very considerable controversy about likely causes is discussed in references cited in that paper.

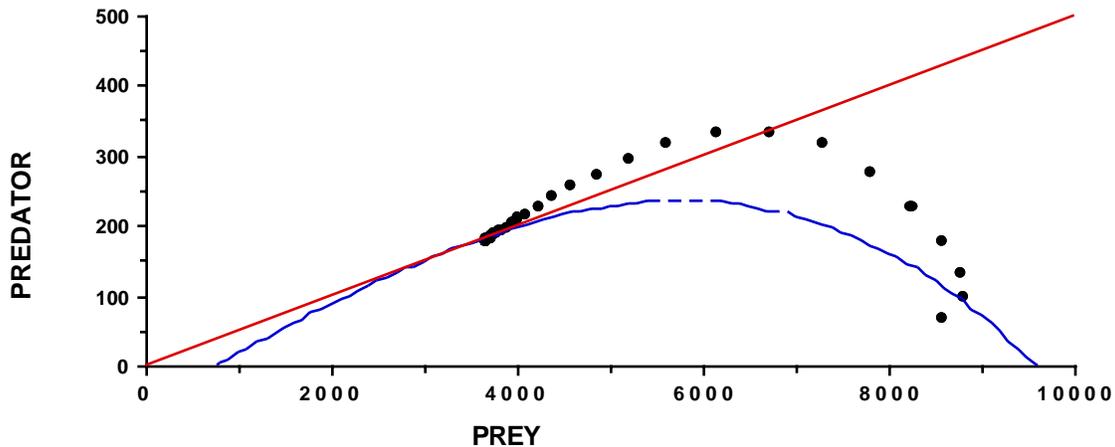


Fig. 12.14. Isocline lines and population plot for equations (12.34) and (12.40).

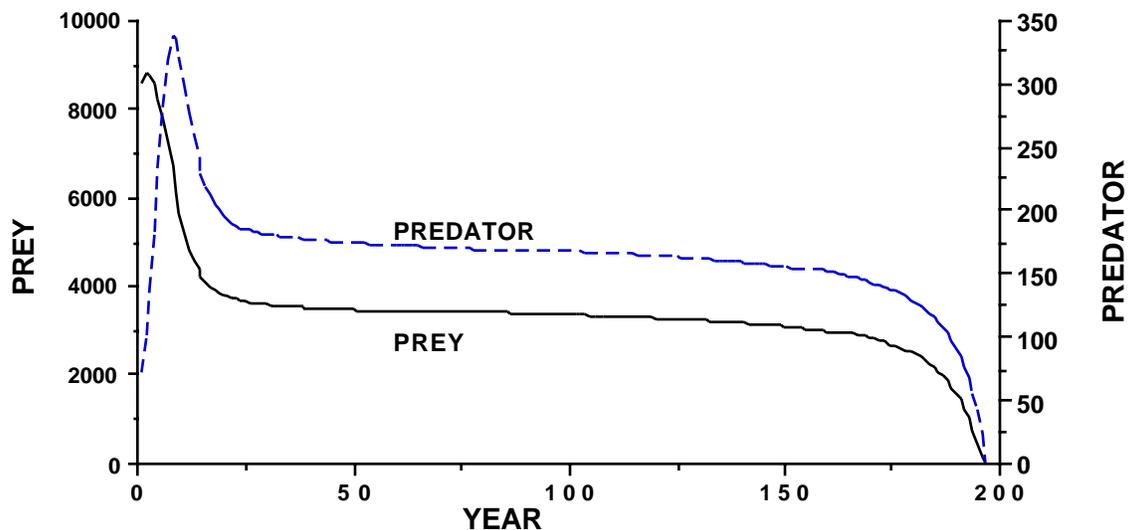


Fig. 12.15 Trend of predator and prey populations when hunter harvest is high enough to ultimately result in a crash of both populations.

Up to this point, we have considered some problems in predator-prey models (Section 12.5), looked briefly at a wolf-ungulate model (Section 12.6),

and then delved deeper into the various differential equation models in the present Section. A tentative conclusion from the review here is that the equations of Section 12.6 seem to “behave” reasonably well in comparison with some popular versions. It may thus be desirable to provide more detail on the evidence supporting that model. For convenience the underlying difference equations are repeated here.

The prey equation is:

$$V_t = V_{t-1} + r_1 V_{t-1} \left[1 - \left(\frac{V_{t-1}}{K} \right)^z \right] - c H_{t-1} \quad (12.35)$$

and the predator model is:

$$H_t = H_{t-1} + r_2 H_{t-1} \left[1 - \frac{H_{t-1}}{a V_{t-1}} \right] \quad (12.34)$$

while the equilibrium conditions are:

$$V = K \left[1 - \frac{c a}{r_1} \right]^{1/z} \quad \text{and} \quad H = a V \quad (12.36)$$

Definitions and some likely parameter values were given in Section 12.6. Here we look briefly at the data supporting these models and the parameters.

Evidence for the constant, c , in eq.(12.35) has been collected from a number of sources (Fig. 12.16) that suggest the moose kill/wolf/100 days to be nearly constant over a wide range of moose densities.

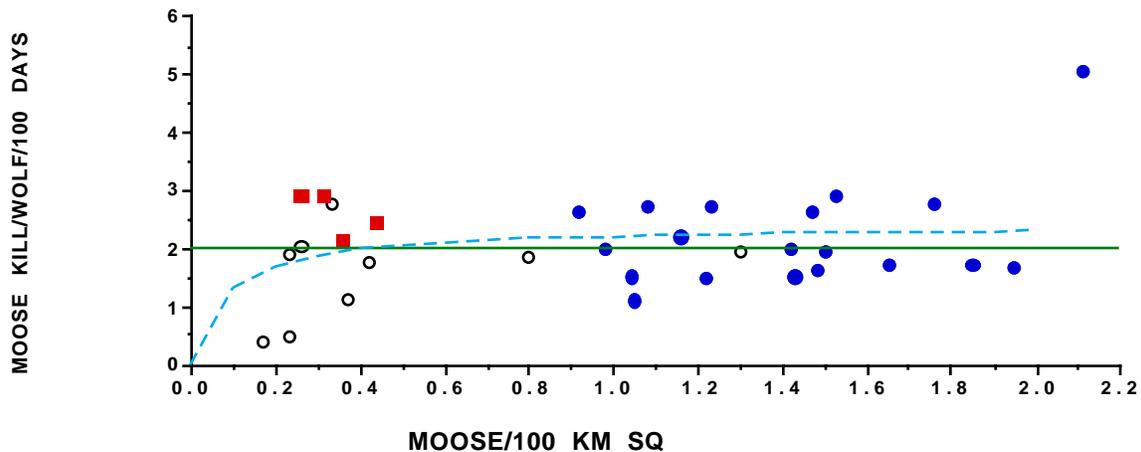


Fig. 12.16. Data on moose kill per wolf per 100 days in winter from Eberhardt (2000) with data (solid square symbols) from Hayes and Harestad (2000) added. Solid circles represent individual year data from Isle Royale and open circles show data of Messier (1994: Table 2). Solid line shows constant rate assumed here, from an average that does not include two very low and one very high point. Broken line shows fit of Type II curve.

Some lower kill rates at low moose densities might be taken as evidence for the Type II model used in eq. (12.38). That model has been fitted to the data with non-linear least-squares as shown in Fig. 12.16. As was noted in connection

with Fig. 12.11, introducing the Type II curve does not appear to change behavior of the equations significantly. Arbitrary values of the parameters m and w were used in eq. (12.38) to give results compatible with most of the other equations. Using actual data now gives quite different values, namely $m = 2.38$ and $w = 0.082$.

The other important evidence for the model suggested here has to do with the ratio dependence constant, a , in eq. (12.34). The arguments for this model were described in more detail by Eberhardt and Peterson (1999). A possible improvement offered here is that the underlying relationship between rate of increase for wolves (λ) and the wolf/moose ratio may be curvilinear, with an intercept at a somewhat lower wolf/moose ratio, changing the equilibrium value from about 20 moose/wolf to 24 moose/wolf (Fig. 12.17).

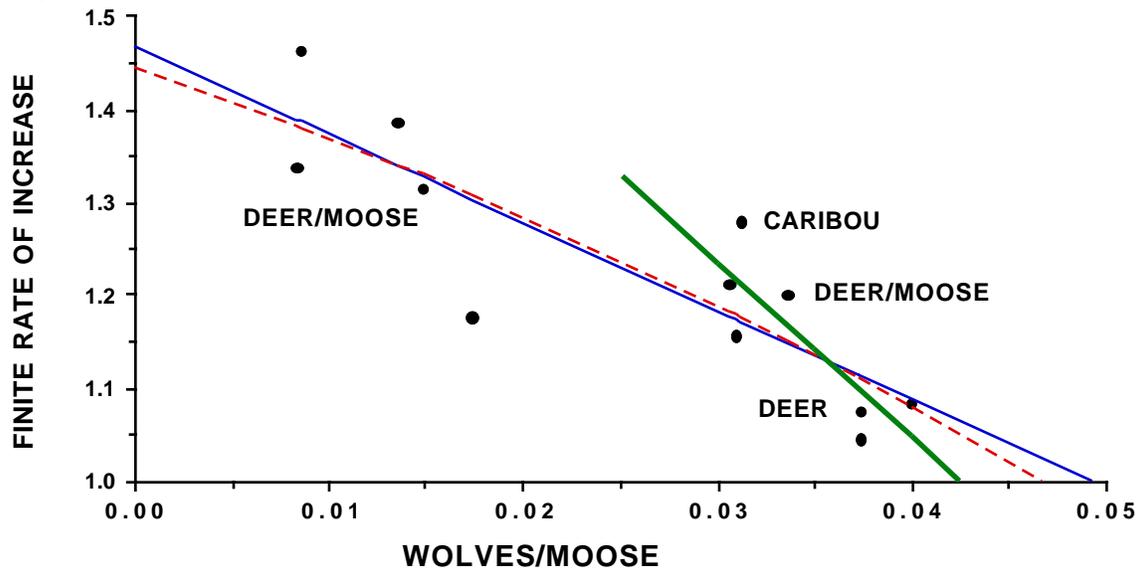


Fig. 12.17. Finite rate of increase (λ) of wolf populations vs. observed wolf-moose ratios. The broken line depicts a second degree polynomial fitted to the data, and the short regression line has been fitted to the data from the 7 highest moose-wolf ratios to approximate what may be an underlying strongly curved relationship. Sources of the data appear in Eberhardt (1998) and Eberhardt and Peterson (1999).

The best support for eq. (12.35) comes from the trend of the moose population on Isle Royale, Michigan (Fig. 12.18). In the early stages of the study it appeared that the wolf population was controlling moose abundance, with the two populations approaching equilibrium by about 1979 (Eberhardt and Peterson 1999). However, the wolf population “crashed” in 1980, very likely in consequence of the arrival of a virus (parvovirus) in the population. Moose numbers then increased steadily and the moose population ultimately crashed in a severe winter. An epizootic of winter ticks (Delgiudice et al. 1997) interrupted the steady growth of the moose population in 1988, so eq.(12.35) has only been fitted up to that point.

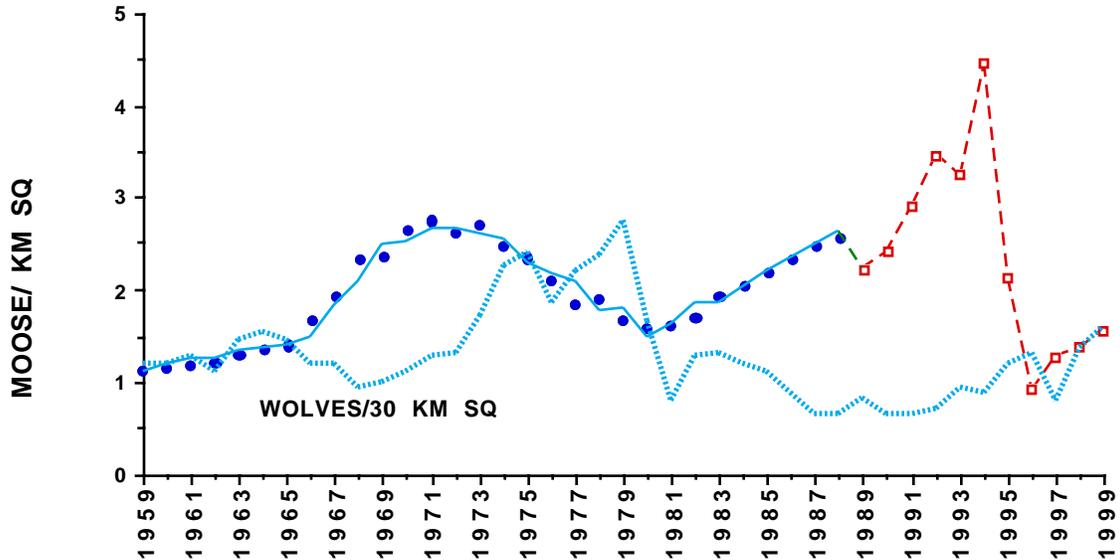


Fig. 12.18. Trend of the Isle Royale moose population. Solid points and a solid line show moose abundance through 1988 and the fit of eq.(12.35) to the data. Open symbols show estimated moose numbers from 1989 onwards and the estimates are connected by a dashed line to show the recent trend. Wolf density is shown by a broken line.

12.8 EXERCISES

12.8.1 Non-linear least-squares. The following are the data from Fig. 9.3 (muskox population growth curve). Fit the loglinear regression model of eq. (12.5) using the analysis toolpak of EXCEL. Then explore non-linear least-squares fitting of the same data by making a table bordered by values of the two parameters and calculating the sums of squares for trial values of the parameters. Calculate the sums of squares from:

$$S = \sum [y_i - (a \exp(bx_i))]^2$$

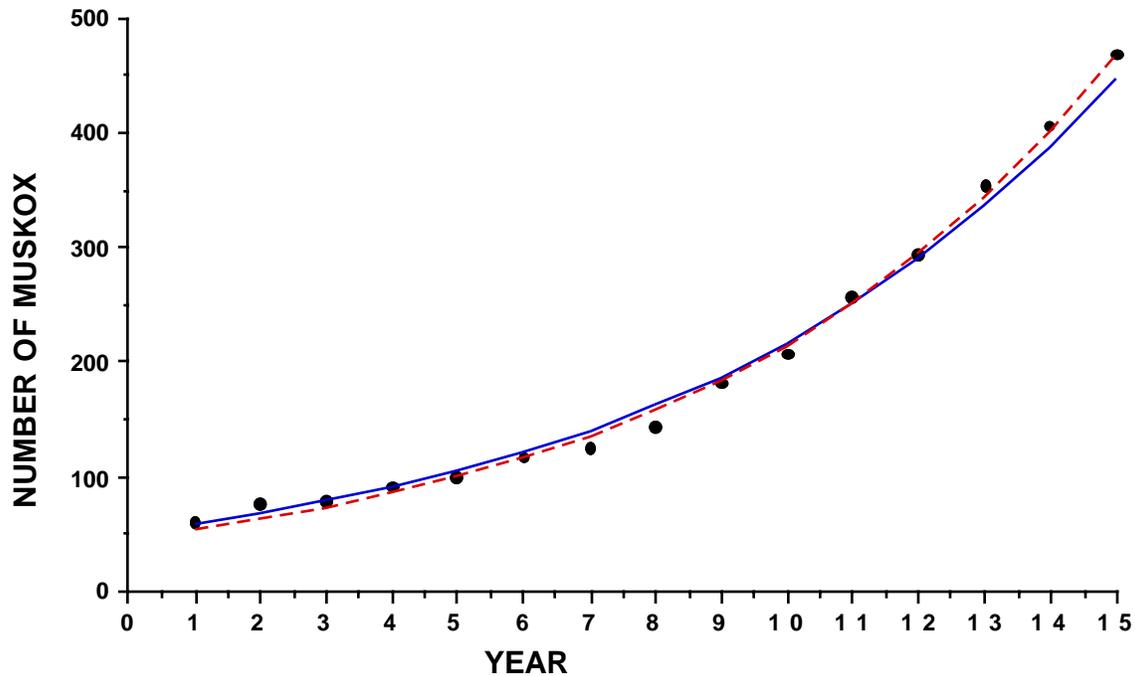
You will need trial values of a and b , and these can be taken from the loglinear fit. It can be a tedious undertaking unless you start out with trial values, put the sum of squares in a table and border these with sums of squares from the adjacent values of a and b . Use 3 decimal places for b and 2 digits for a (don't forget that the loglinear fit gives $\log a$ as the intercept, so you have to transform back). You can map the sums of squares and trace the minimal sums of squares until you find one that is the lowest in its region. Sometimes this "brute force" approach is useful if you don't have ready access to a computer program that provides nonlinear least-squares estimates. If you do have access to such a program, use it to check your results.

The data are:

YR	NUMBER
1	61
2	76
3	77
4	90

5	100
6	116
7	126
8	143
9	181
10	206
11	256
12	293
13	353
14	406
15	467

A plot of expected values from loglinear regression and from the nonlinear least-squares fit is:



The solid line is the fit from loglinear regression. The nonlinear fit (dashed line in the figure) does seem better for the higher counts but experience with a variety of population growth data suggests using the loglinear fit if, as is usually the case, one is mainly interested in studying rates of growth.

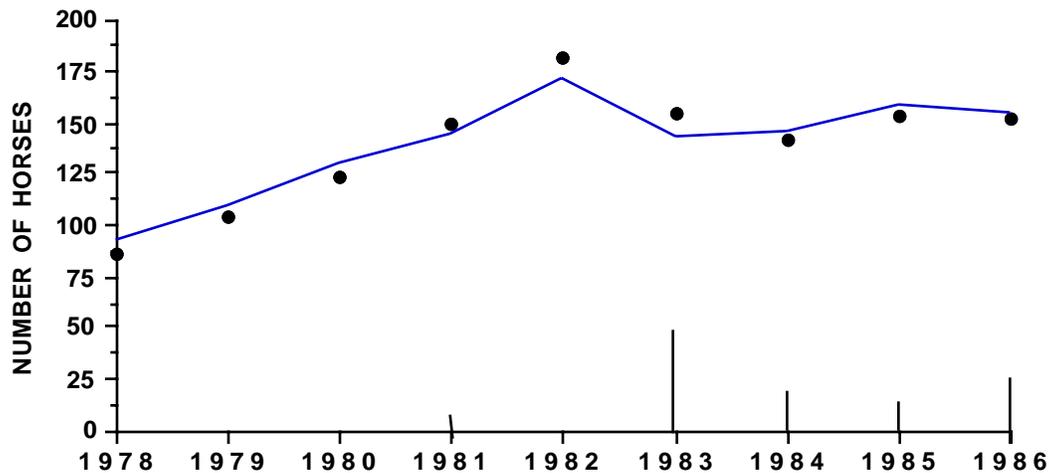
12.8.2 Use parametric regression bootstrapping (Ch. 2) to calculate approximate 95% confidence limits on the population growth rate obtained by loglinear regression in Exercise 12.8.1 and compare them with the limits obtained from the regression program in EXCEL.

12.8.3 Plot the rate of change for the logistic equation [Eqs.(12.12) and (12.26)] using the approximation given after eq.(12.26) which was used to make the plot of Fig. 12.3. This is a serious limitation for use of the logistic equation for the large vertebrates, as their rates of change do not seem to behave this way.

12.8.4 An alternative to the logistic is the “generalized logistic” used in eq. (12.35). Note that it was used with $z=1$ to approximate the ordinary logistic for Fig. 12.3. Repeat the plot of Fig. 12.3 (in EXCEL), and make companion plots with $z=2,5$, and 11. Plot the rate of change and compare it with the rate of change obtained in Exercise 12.8.3. It is this nearly constant rate of change over much of the range of population growth that typifies the large vertebrates.

12.8.5 A set of data on feral horses (Garrott and Taylor 1990) is given below. Use Model I (Eqs. 12.22) and the ratio estimator of eq. (12.23) to estimate λ using EXCEL. Compare your results with the following plot.

YEAR	OBSERVED	REMOVALS
1978	86	0
1979	104	0
1980	123	0
1981	150	7
1982	181	0
1983	155	50
1984	142	20
1985	153	13
1986	152	27



Plot of feral horse data showing removals and fitted curve for Model I.

12.8.6 Using EXCEL and the constants given in Section 12.6 reproduce the results of Fig. 12.5 with eqs. (12.34) and (12.35). You can use this approach to reproduce many of the models given in the literature.

13.0 DEALING WITH DENSITY-DEPENDENCE

13.1 Introduction

"Density-dependence" is a convenient label for the notion that populations cannot increase in size indefinitely. Eventually some resource becomes limiting and the population tends to level off. In long-lived vertebrates, early survival is often the first parameter to change with increasing abundance. The best known model for a density-dependent response is the logistic equation (eq. 12.12 and 12.26) described in Section 12.3 and 12.5. The resulting curve (Fig. 12.3) shows a smooth approach to an asymptote (usually denoted by K). A major problem in applying the logistic equation is that the rate of increase (r) declines continuously throughout the growth of the population, approaching zero as the population size approaches K . This behavior is evident from the underlying differential equation, but is perhaps more conveniently exhibited in the analogous difference equation (eq. 12.25) where it can be seen that

$$r = r_1 \left(1 - \frac{N_t}{K}\right)$$

so that when population size is small relative to K , r is nearly at the maximum value, r_1 , which is sometimes appropriately labeled r_{\max} .

The logistic curve became popular after Raymond Pearl (1926) used it to describe the growth of yeast populations. While it may be useful as a model for growth of populations of some organisms, the logistic is generally unsatisfactory as a model for growth of populations of large vertebrates. This is because, in practice, the rate of increase tends to be relatively constant over much of the range of population size and then begins to decrease as the asymptotic value is approached. For this reason, a "generalized logistic" curve provides a much more satisfactory model for large vertebrates. For this model, r holds nearly constant over much of the range of population growth, declining quite steeply as the asymptotic level (K) is approached. The generalized logistic is written as a difference equation:

$$N_{t+1} = N_t + r_1 N_t \left[1 - \left(\frac{N_t}{K}\right)^z\right] \quad (13.1)$$

Note the change in notation from $t-1$ and t used in Chapter 12 to t and $t+1$. This is done to conform to the notation used in references cited below. When $z=1$ the rate of increase is as given above, i.e., the rate for the ordinary logistic, written as a difference equation. A comparison (Fig. 13.1) of the two rates of increase (for $z=1$ and $z=5$) shows why the generalized logistic yields a population growing at a relatively constant rate up until the "carrying capacity" level (K) is approached, and is thus a better model for population growth of the long-lived vertebrates, which generally appear to have a relatively constant growth rate at levels below carrying capacity.

In considering density dependence it is important to have in mind some characteristics of different groups of species. Much of the current literature

deals with invertebrates, where very high rates of increase can lead to chaotic-seeming behavior of populations. At the other extreme are the long-lived vertebrates with relatively low growth rates that usually result in less erratic behavior. Many (but not all) species of fish can achieve high growth rates and may thus follow patterns closer to those of invertebrates. Many of the smaller vertebrates have short generation times and thus lack the stabilizing qualities of age-structured populations and can also show substantial year-to-year fluctuations in abundance. Consequently, some of the complexities in the current literature may not be relevant when considering the long-lived vertebrates, and attempts to apply models derived from studies of invertebrates to vertebrate populations may be quite unsatisfactory.

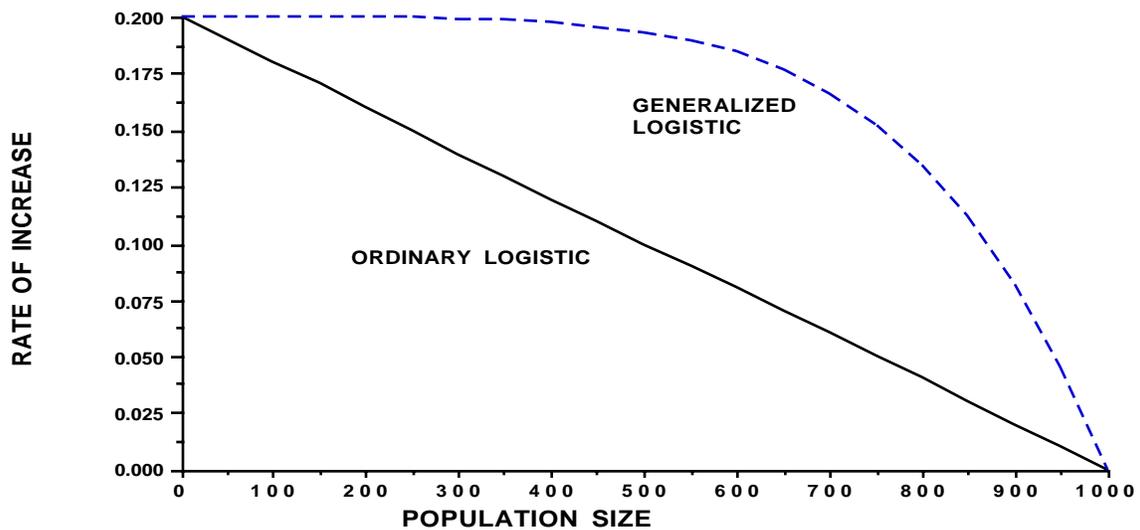


Fig. 13.1 Comparison of rate of increase for the ordinary and generalized logistic growth curves.

Unfortunately, there isn't much in the way of accurate and precise data on population growth in large vertebrates over a wide range of population sizes, making it difficult to demonstrate directly that population growth is more like the generalized logistic curve than the ordinary logistic. However, if one considers Fig. 13.1 and notes that many vertebrate populations do appear to grow exponentially over much of the observed range, the ordinary logistic is much less palatable. Experience with harvested populations also shows that the ordinary logistic will not support observed rates of removals. It is difficult to settle on a value of z , the parameter governing the rate of decrease in r with increasing population size. Often one must make an arbitrary choice. In developing guidance for managing cetaceans, the International Whaling Commission considered various values of z . The value used in various examples here is $z=5.0$, which gives the maximum numerical increment of growth at 70% of K . Because the details of density-dependent responses and "population regulation" are so poorly known for large vertebrates, about all that can be done here is to try to supply some guidance for practical approaches to the subject. The basic assumption is made that populations do not grow indefinitely, and thus are somehow limited.

13.2 Some historical aspects

Hairston, Smith, and Slobodkin (1960) produced an all-embracing analysis, starting with the premise that energy is not accumulated significantly in organic matter, and working up from there to conclusions about major trophic levels. Two critiques of their paper then appeared (Ehrlich and Birch 1967, and Murdoch 1966), and were rebutted by the original authors (Slobodkin, Smith, and Hairston 1967). The rebuttal extends the original thesis somewhat and makes for interesting reading. All of these papers depend on qualitative assessment of limiting cases and are thus not very helpful in deciding what controls a specific population. One might, for example, agree that predators as a class are food-limited, without abandoning the notion that particular predatory species seldom increase beyond bounds imposed by behavioral (territorial) constraints.

Andrewartha and Birch (1954) offered the strongest objections to the idea of density-dependence. Various critics of their views emphasized their primary involvement with insects and that many of their conclusions depend on events observed near the geographic limits of particular species, where "limitation by catastrophe" probably does control populations. That this is an over-simplification of their views was vigorously pointed out by both authors, individually and jointly. Some amelioration of the original stand seems evident in later papers by Andrewartha and Birch. In any case, their work has provided useful evidence that many factors may interact to limit populations, not the least of which is the effect of essentially random "shocks" from weather conditions. Murdoch (1994:284) claimed that the "decades-old controversy about regulation has been resolved in recent years", suggesting that the "Nicholson school was right that regulation via stabilizing density-dependent processes is essential to account for species persistence" but that "the local randomness and spatially out-of-phase dynamics emphasized by the Andrewarthan school can create the necessary stabilizing density-dependence".

V.C.Wynne-Edwards promulgated what might be said to be a universal solution -- that social factors, frequently expressed as territorialism, control population size. His thesis closely parallels the wildlife management principle of carrying-capacity, and fits nicely with a readily observed aspect of the behavior of many species. There are, however, too many situations in which territorial behavior does not seem to be either pronounced or effective, perhaps most notable among the larger herbivores. Wynne-Edwards' philosophy is expounded at length in a large book (1962), and compactly in a later paper. Several authors have considered that a weak point in his theme is an invocation of group-selection as a necessary evolutionary device to insure success of regulation by territorialism.

A specific physiological mechanism for population regulation was proposed by J.J.Christian, and then tied to social aspects (Christian and Davis 1964). Objections to the "adrenal function" theory have centered around apparent absence of enlargement of the adrenal glands under situations where it would seem to be called for by the current status or ultimate fate of a particular population. D.Chitty has been particularly explicit on this point (see the discussion and references in Christian and Davis 1964). Chitty's own

attitude on the matter of regulation appeared mostly to have been one of a rather reluctant invocation of a genetic mechanism, at least to handle the problem of cyclic crashes of lemmings and voles. The most detailed exposition of his "polymorphic behavior" mechanism was by one of his students (Krebs 1964) who contrasted that theory with Christian's stress hypothesis and a food-supply hypothesis offered by Pitelka (cf. Pitelka 1958).

One of the apparent reasons for Chitty's belief in a change in quality in vole populations during a decline is the evident necessity for some deleterious factor that persists through more than one generation in order to bring the population down to the low levels observed in the field. In a long-term study of the vole, *Microtus californicus*, Pearson (1966) came to the conclusion that such a factor is supplied by predation. He contended that some other (unspecified) cause is responsible for the initial decline, but that predation then acts to reduce the population to very low levels, whereupon the predators then die off (in the Arctic) or shift to other prey (in temperate regions).

A feature not sufficiently stressed in many analyses of population regulation is that of evolutionary forces. A valuable review of the importance of studying evolutionary ecology is that of Lack (1965). One of his more telling points is that few real opportunities remain to study ecological processes in a sufficiently natural state that one might hope to appraise selective forces free of changes due to human influence.

Even a short excursion into the literature is sufficient to show that there are many views of population regulation. It seems reasonable to suppose that, just as a variety of life forms has evolved, so may have a range of controlling mechanisms. The actual structuring of a population model may, however, call for a specific choice of mechanism, particularly if relatively long spans of time are to be dealt with. On the other hand, if we choose a mathematical relationship to represent population regulation without good evidence that it represents a mechanism actually existing in the population, then we are really deciding the outcomes of the modelling exercise in advance, so that it may have little relevance to the actual population.

13.3 Testing for density-dependence

If a population appears to level-off, it may be desirable to have a statistical test to demonstrate that the apparent tendency is not due to chance fluctuations alone. There is a very large literature on testing for density-dependence, but the results are as yet often ambiguous and controversial for large vertebrates. This is perhaps less true for smaller organisms, particularly insects. Due to the very different life histories involved, suitable tests for density-dependence for insects should not necessarily hold for larger species. In particular, most insect populations are capable of very high rates of increase, as are some fish populations. A density-related response may thus occur over a short time span and be readily evident in the data.

One supposed test for density-dependence has been repeatedly shown to be of dubious value over the last 30 years or so, but still turns up in the literature. No doubt this is because the approach is simple and seems natural. It consists of estimating λ from the ratio of two successive population estimates, $\lambda_t = N_t/N_{t-1}$, and then calculating the correlation coefficient of λ_t and N_{t-1} on

a series of such estimates. If λ_t is negatively correlated with N_{t-1} , then one might assume density-dependence is involved as λ_t apparently decreases with increasing population size. The difficulty, however, is that values of N_t appear successively in the numerator and denominator of the ratio determining λ_t and this induces a spurious correlation. This prospect was discussed for insects by Watt (1964), and some additional examples were noted by Eberhardt (1970).

There has been considerable interest in testing for density dependence for the last 30-40 years, but it remains uncertain whether density dependence can be reliably detected from a sequence of data on population size or trend alone. Many of the slower-growing populations appear to increase fairly smoothly from low levels, and ultimately show signs of leveling-off. Those with high growth rates may behave quite erratically, making it difficult to discern any pattern. Most of the efforts to devise tests appear to have been inspired by studies of populations showing rather erratic growth patterns, particularly insect populations. Nearly all of the recent efforts proceed along the lines of statistical hypothesis-testing. Most of the available statistical methodology for conducting such tests is based on linear models, and this limits the range of models that can be considered. Most of the tests depend on taking logarithms (base e, usually) of population size, which generally appears to yield symmetrical distributions of deviations from the model, and thus encourages assuming normally distributed errors (a necessary assumption in many of the tests). Statistical hypothesis testing depends on setting up a "null" model and attempting to reject that model in favor of an alternative model. Rejecting the null model does not demonstrate that the alternative model is correct, a fact that sometimes appears to be overlooked in the ecological literature. Because the null model generally needs to be linear and fairly simple in form, the process more or less restricts attention to two models, both linear and quite simple. In some cases, "nested" models may be considered, giving a little more flexibility.

A number of authors have used the following model for density dependence:

$$N_{t+1} = N_t^b K^{(1-b)} \quad (13.2)$$

Where $0 < b < 1$, and K is the asymptotic value. An example of the resulting curve is shown in Fig. 13.2. Taking logarithms gives the model commonly used in studying density dependence:

$$X_{t+1} = bX_t + (1-b)\log_e K \quad (13.3)$$

Here X denotes the logarithm of N and b and K are as given above. In producing statistical tests, an error term (ϵ), assumed to be normally distributed with some variance is added. Bulmer (1975) wrote the above equation in terms of deviations from $\mu = \log K$:

$$X_{t+1} - \mu = b(X_t - \mu) + \epsilon_t$$

Bulmer's is one of the more widely quoted papers on testing for density dependence. He used the null model (described as a "random walk" model):

$$X_{t+1} = X_t + \varepsilon_t$$

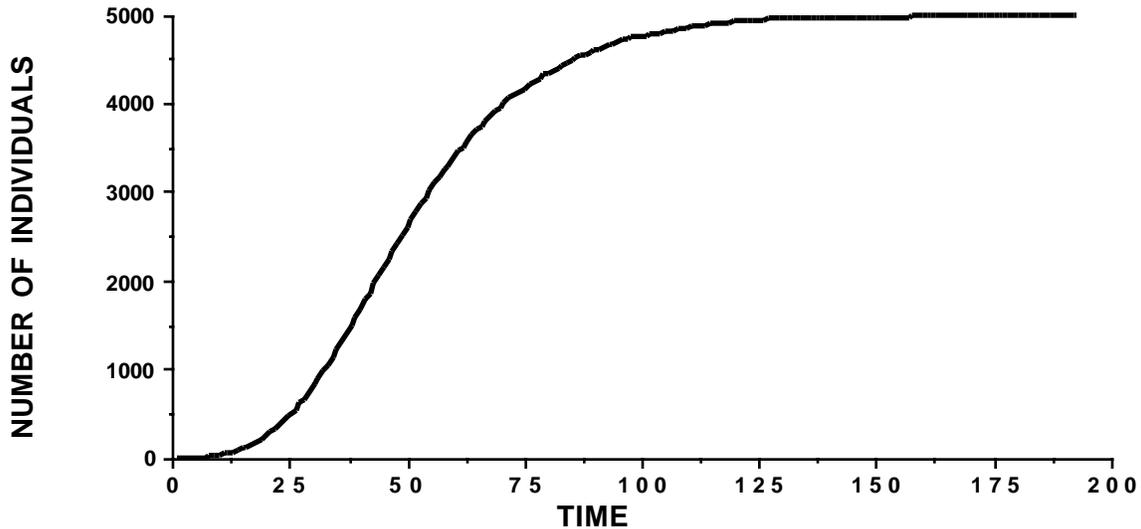


Fig. 13.2 An example of curves represented by Eq. (13.2). $K=5,000$, $b=0.95$.

Pollard et al. (1987) used the same models, but added another null model, described as a "random walk with drift":

$$X_{t+1} = X_t + r + \varepsilon_t \quad (13.4)$$

It is important to note that this is the logarithmic form of the usual model for exponential growth (over one unit of time) with a multiplicative error term ($\varepsilon_t = \log_e e_t$):

$$N_{t+1} = N_t \exp(r)e_t$$

Pollard et al. use the three models [one, eq. (13.3) representing density dependence and the two null models above] in "nested" form, so the tests can presumably be used to choose among the 3 models. Examples of the random walk and random walk with drift are shown in Fig. 13.3 and 13.4.

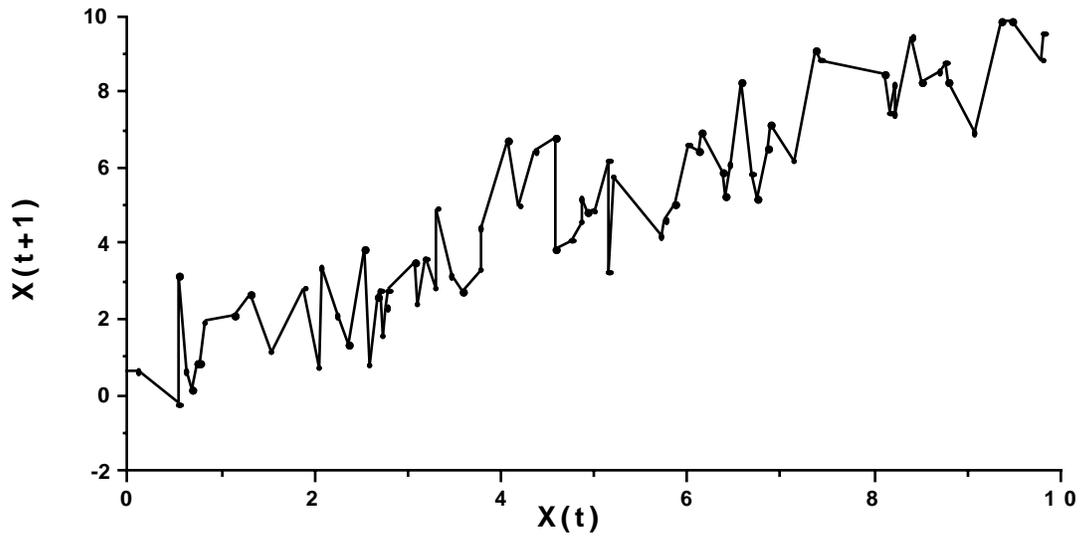


Fig. 13.3 An example of the random walk model used as a null model in various tests for density dependence.

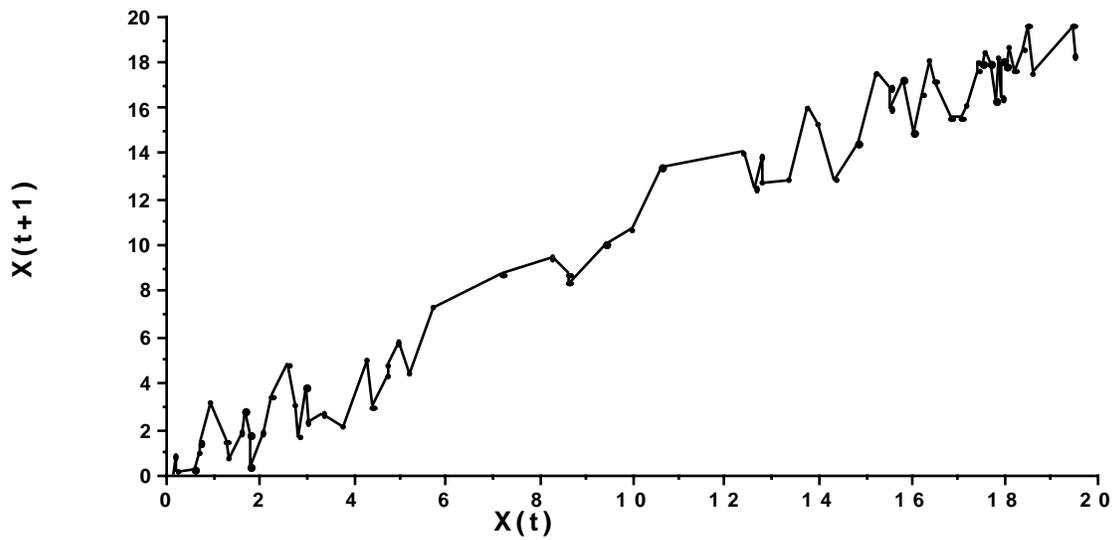


Fig. 13.4 An example of a random walk with drift null model (Eq. (13.4)).

An example of the linear (log scale form) of eq. (13.3) with an additive error term appears in Fig. 13.5.

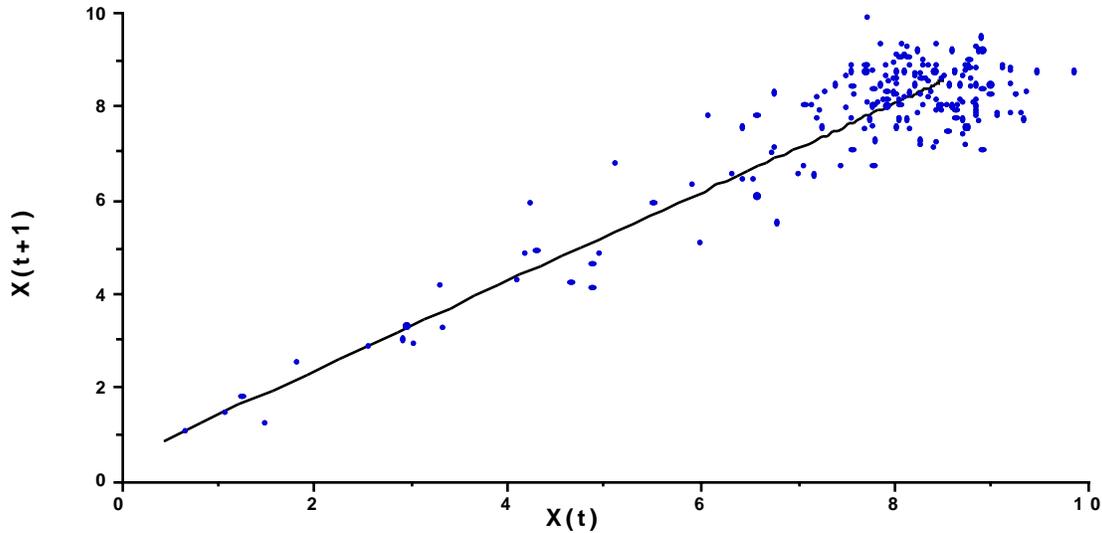


Fig. 13.5 . An example of the linear model of eq.(13.3) with an additive error term

The sizable cluster of points on the right results from the substantial section of the underlying curve (Fig. 13.2) close to the asymptotic value.

Some problems with the approach to testing density dependence with the models illustrated above can be seen by comparing the basic model (eq. 13.2) and Fig. 13.2) with the generalized logistic of eq. (13.1). Making such comparisons requires a comparable set of parameters. This can be obtained by choosing parameter values that give similar values of λ for the models of eqs.(13.1) and (13.2). Values of λ for long-lived vertebrates may range up to about $\lambda = 1.5$ but are often close to 1.2. Rearranging eq.(13.2) gives:

$$\lambda_t = \frac{N_t}{N_{t+1}} = N_t^{b-1} K^{1-b} = \left(\frac{K}{N_t}\right)^{1-b}$$

and values (Fig. 13.6) of $b=0.95$ and $b=0.98$ fall in this range (up to $\lambda = 1.5$). Comparing the generalized logistic with these two curves (Fig. 13.7) shows the sharp difference between eq.(13.1) and eq. (13.2). Using smaller values of b shifts plots of eq. (13.2) to the left but yields the much larger values of λ characterizing insects and some species of fish.

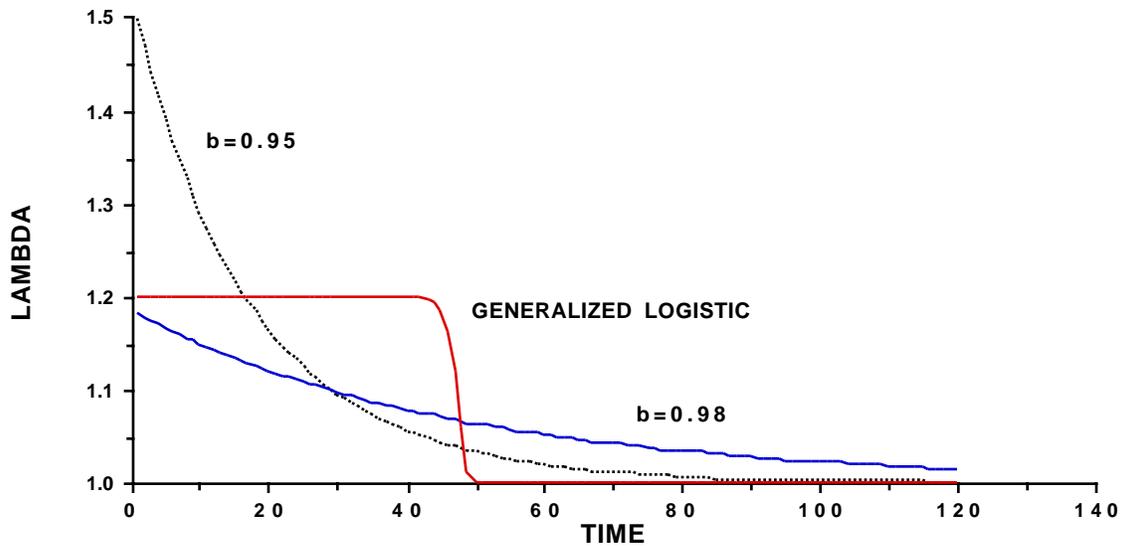


Fig. 13.6 Lambda for density dependence model used by Bulmer (1975) and Pollard et al. (1987) compared with lambda for a generalized logistic curve. Values of b used here are much larger than those encountered in studies of invertebrates in order to obtain values of λ in the range expected for long-lived vertebrates.

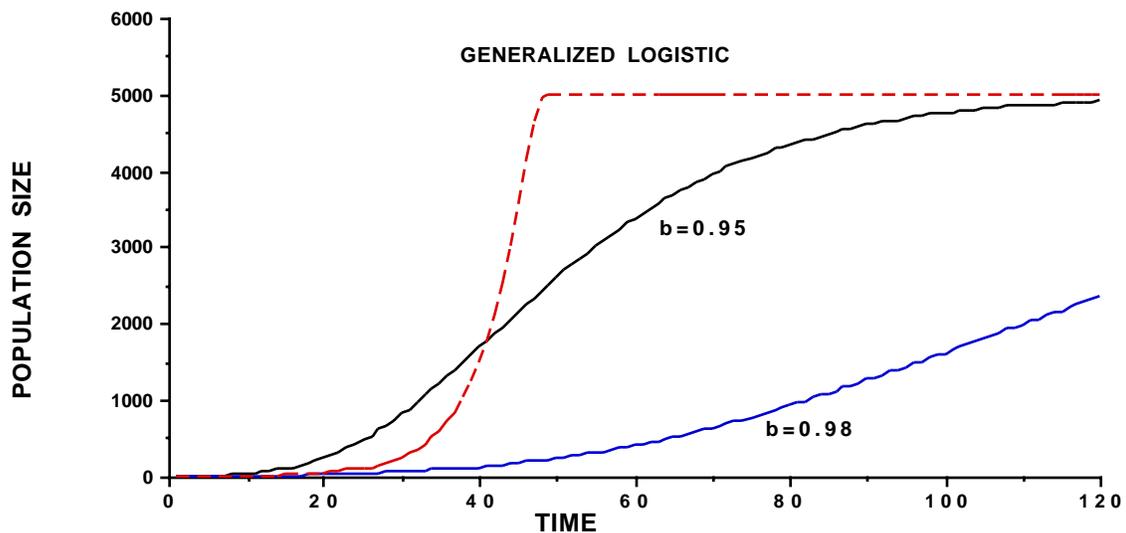


Fig. 13.7. Comparisons of curves from eq. (13.1) and a plot of the generalized logistic (broken line above). Generalized logistic curve based on lambda 1.2 and $z = 5.0$.

A basic issue thus is that the curves are quite different, so that there may be a question about efficiency of the test if the alternative model really should be more like the generalized logistic than like eq.(13.1). Another important issue is that the generalized logistic curve behaves essentially as an exponential curve for the lower part of its range (see Fig. 13.1), so one would expect that the null model of eq. 13.4 would not be rejected for a population growing in the lower part of its range. Thus the test would have essentially zero power of detecting density dependence in an important portion of the

range of growth. This is not necessarily true for eq. (13.2) as λ is large and changes dramatically in the early growth phase for smaller values of b . Hence the tests may be effective for insects and fail for long-lived vertebrates.

Dennis and Taper (1994) proposed tests for a more complex model of the following form:

$$N_{t+1} = N_t \exp(a + bN_t + \sigma Z_t) \quad (13.5)$$

where $b < 0$ and Z_t is assumed normally distributed with zero mean and unit variance, while σ is a constant that lets one choose the variance of the error term. Taking logarithms gives:

$$X_{t+1} = X_t + a + b e^{X_t} + \sigma Z_t \quad (13.6)$$

This model is thus nonlinear (because of the exponential term), making for “intractable” distributions of test statistics. Dennis and Taper got around the problem by using bootstrapping and some simulations to check their results. The model of eq.(13.5) is equivalent to one used by W.E. Ricker (1975) to study stock-recruitment in fish [essentially the relationship between size of a parental generation (stock) and the succeeding generation (recruits)]. Without the error term, eq.(13.5) can be written as:

$$N_{t+1} = N_t e^{r(1 - \frac{N_t}{K})}$$

where r is the growth rate and K is the asymptotic value. If r is in the range of growth for most long-lived vertebrates (say $0 < r < 0.5$) then a good approximation to the above curve can be obtained by using the first two terms of the Taylor series expansion of the exponential term above, obtaining:

$$N_{t+1} = N_t + rN_t(1 - \frac{N_t}{K})$$

which is the discrete form of the logistic growth curve (more details are available in Eberhardt 1977). This brings the model of Dennis and Taper (1994) closer to the kind of growth curve one might expect would hold for long-lived vertebrates. However, they also use nested models with the same random walk and random walk with drift (eq. 13.4) used above. Hence, if the generalized logistic actually is the best model for a given population, then the power of the test for populations at the lower end of the growth range is again essentially zero, as discussed above.

Further difficulties with the recent approaches to testing for density dependence were proposed by Shenk et al. (1998) who used computer simulations designed to mimic the incorporation of sampling error in the data. In most assessments, the error term is assumed to be due to “changes in climate and other environmental factors” (Bulmer 1975), although Bulmer did consider sampling error in one of his tests. Shenk et al. concluded that

“because these tests have been shown to be either invalid when only sampling variance occurs in the population abundances (Bulmer’s R , Pollard et al.’s and Dennis and Taper’s tests) or lack power (Bulmer’s R^* test), little justification exists for use of such tests to support or refute the hypothesis of density dependence”. It seems likely that their results will not be the last word on the issue, inasmuch as it is possible to estimate the contribution of sampling error in at least some circumstances, and also because they chose to use a model of a large mammal population to simulate density dependence in their analyses, while most of the methodology described above seems most appropriate for insect populations.

Only a brief sketch of apparent problems in testing for density dependence has been attempted here. Clearly the suggested difficulties need to be examined in more detail. This calls for assessments of the actual tests and thus brings in the subject of likelihood ratio testing.

13.4 Population regulation

A broader view of the issues involved in testing for density dependence was taken by Murdoch (1994). He discussed population regulation, “which arises as a result of potentially stabilizing density-dependent processes, even when brought about by ‘non-equilibrium’ mechanisms”. He defined regulation very broadly, stating “Population regulation seems best defined by defining non-regulation, which is random walk dynamics”. He defined random walk in the form described above as random walk with drift” (eq.13.4) and thus espoused the null model of exponential growth. Such a population is unbounded (i.e., increases indefinitely) and thus Murdoch stated that “we can identify regulation with boundedness”. Because nearly all populations (excepting perhaps humans) appear to be bounded, practical-minded observers may see little reason to study density dependence. However, Murdoch suggested several possibilities for modes of population change that may transcend the simpler concepts of density dependence. He noted that regulated populations “thus include not only those with a stable equilibrium but also those with a stable attractor, i.e., cyclic or chaotic populations”. A cyclic model was described here by eq.(12.33) and (12.34) and illustrated in the upper panel of Fig. 12.4. Another very interesting concept for regulation without apparent density dependence can be developed by considering a metapopulation in which isolated subpopulations follow a random walk (without “drift”) and ultimately go extinct but with the area they occupied later recolonized by immigration from another distinct population. Murdoch (1994:275) commented that “little can be inferred from a statistical analysis of the time series of a single population. Thus if we are interested in exploring regulation in a particular population we need to investigate the mechanisms directly.”

Murdoch (1994) reported the results from assessments of a large number of sets of data on small bird populations that found the populations to be remarkably stable. He commented that “The remarkable invariance of these bird populations makes it difficult to believe that they are not in fact well regulated. If they are, we clearly have still not developed adequate means of detection”. From a statistical point of view, there is little or no prospect of detecting tight regulation from an analysis of the time series of observations alone. This is essentially proposing to prove a null hypothesis holds, when we can only reject a hypothesis through statistical analysis. Another example of

this problem concerning an elk population is discussed in Section 14.6 . In that example there is good evidence of a strong climatic effect that results in dramatic changes in first-year survival, yet the population trend is far less variable than would be expected. Our tentative conclusion is that high survival of adults “buffers” the effect of variable first-year survival, and that there is a strict upper limit on first-year survival enforced by the availability of winter thermal refugia. Examination of population trends can only show that there is much less population variability than one would expect on the basis of environmental data (30 years of snow depth data are available) so that the main prospect for assessing density effects is by direct measurement of population parameters. Detecting regulation then requires that the populations somehow be perturbed and essential parameters monitored during the return to an equilibrium condition.

13.5 Testing for a maximum net productivity level.

Thus far we have seen that the prospects for detecting density dependence from a series of measurements of population trend are presently doubtful. In the present section, we assume that density dependence does prevail in a population and seek a means to determine how it affects the current trend. An important practical problem is to determine status of a population with respect to the maximum net productivity level (MNPL). This is the point where the slope of the s-shaped curve of population growth shows a change in rate of change of the slope (the inflection point of the curve). Determining status of a population with respect to this level is important for several reasons. One is that the maximum sustainable yield (MSY) of a population can be taken at this level, a second is that this point figures in legislation that requires managing populations at or above that level (for example, the Marine Mammal Protection Act of 1972), and a third is the evidence (suggested in Section 13.3) that there likely is little prospect of detecting density dependence using the approaches presently in the literature when a population is below the inflection point in the growth curve. It is important to stress that the test suggested here depends on assuming that an s-shaped growth curve does govern the trend of a particular population, i.e., that we assume density dependence exists.

The basis for the test is to attempt to ascertain whether the recent trend of the population traces out a curve that is concave upwards (below MNPL) or concave downwards (above MNPL). This approach thus assumes an s-shaped growth curve and the existence of density dependence. The test has two stages and depends on the approximate test for curvilinearity given in Section 1.7. In the first stage, a straight line and second degree polynomial are fit to untransformed data. If the test for curvature is significant, we conclude that the population is below MNPL. If the test is not significant on data in the arithmetic scale, then we transform to the log scale and repeat the test. Significance suggests the population is above MNPL. The test was studied with simulations based on the assumption that the generalized logistic applies to first-year survival according to eq. 13.7.

$$s_1 = s_0 \left[1 - \left(\frac{N_t}{K} \right)^\varepsilon \right] \quad (13.7)$$

Here N , K , and z are as before and s_0 is a constant while s_1 denotes first-year survival. This general approach was later used in the simulations of Shenk et al. (1998). More details on the test and simulation are available in Eberhardt (1992). The approach essentially depends on the fact that the usual growth curve is concave upwards below MNPL and concave downwards above MNPL. However, transformation to the logarithmic scale yields a curve that is nearly a straight line below MNPL but remains concave downwards above that point (cf. Eberhardt 1992:Fig. 3), and this seemed to improve the power of the test (Eberhardt 1992).

13.6 Components of density dependent responses

If, as seems to be the case generally, one cannot reliably detect density dependence from a sequence of measurements of population trend data, it will be necessary to examine the behavior of population parameters as the population changes. Thus far there are relatively few sets of data that provide the necessary details. Some preliminary results can, however, be suggested. It appears to be a generally accepted principle that adult female survival is the key to well-being of at least the populations of the long-lived vertebrates, and quite possibly to most of the sexually reproducing species. Over 20 years ago it was proposed that there appears to be a sequence of changes in vital rates as population density increases towards maximal levels (Eberhardt 1977). Later studies support this sequence (Fowler 1981,1987, Gaillard et al. 1998). The proposed sequence is given in Fig. 13.8.

With the advent of effective mark-and-recapture methods, particularly those employing radiotelemetry, it has become possible to obtain data on this sequence for a number of large mammal populations, and to estimate the rate of change (λ) from such data. Often the resulting estimate of λ can be checked by direct estimates of population density over time, or through measures of relative abundance (indices of abundance). Inasmuch as most estimates from field data are subject to a variety of potential biases, it is very important to have estimates from both reproductive and survival data and from independent direct measures (census or index).

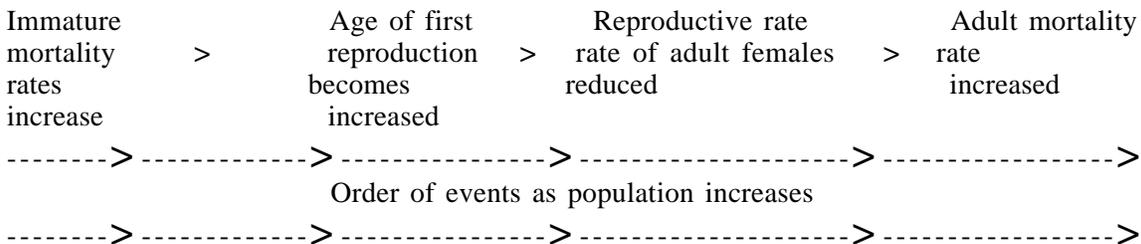


Fig. 13.8 Sequence of events contributing to regulation as a population of long-lived vertebrates increases in abundance.

For populations with little impact of human activities, adult female survival appears likely to be at least 0.94 and likely higher (Fig. 13.9). Many of the examples used here were subjected to impacts that likely reduced survival below a feasible maximum. Consequently, it seems very likely that adult female

survival will be at least 94% and may be 99% or higher in the prime age classes under truly undisturbed conditions. Such a high annual survival is necessarily accompanied by reductions in survival due to senescence which may not always be taken into account in reported estimates of adult female survival.

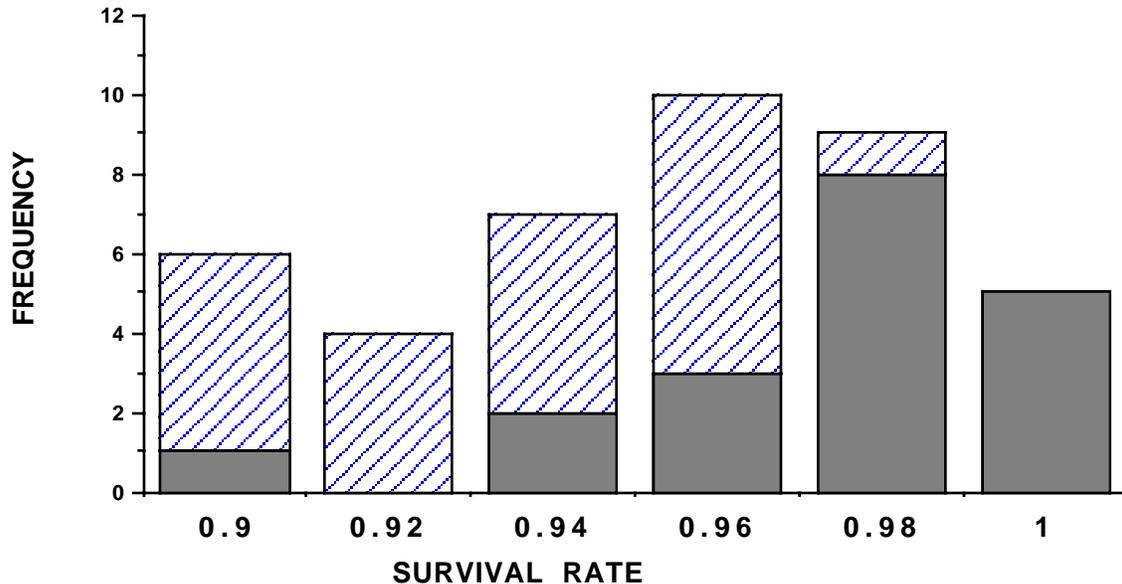


Fig. 13.9. Frequency of adult female survival rates. Most of the examples are from large mammal data, but a few are from long-lived birds. Dark bars represent species where there is reason to believe human influence might have a minor influence, while those represented by the lighter bars likely were affected by modest human influences.

Under largely undisturbed conditions, it seems reasonably well-established that a density-dependent response in long-lived vertebrates will initially be evident in first-year survival. Poor first-year survival normally results in poorer physical condition which will, in turn, result in a delay in ages of first reproduction. It may well be that these two factors will be sufficient to control population growth without invoking the subsequent stages of Fig. 13.8, but extreme conditions may well result in an impact reflected in all four stages. The Hawaiian monk seals of French Frigate Shoals provide one example of this result (Section 14.5 and Example 13.1 below).

In Chapter 11 (Section 11.9) we considered a method for obtaining approximate variance estimates for the rate of population change (λ). An essential component is estimating the partial derivatives of λ with respect to survival and reproductive rates (eq. 11.13). These provide a measure of the relative importance of such rates in determining population growth. They indicate that adult female survival is most important, with reproductive rate next, and survival to reproductive age apparently least influential (Fig. 13.10). We can thus speculate that, because long-lived vertebrates have relatively slow rates of population growth, early survival might be the best mechanism for at least initiating population regulation.

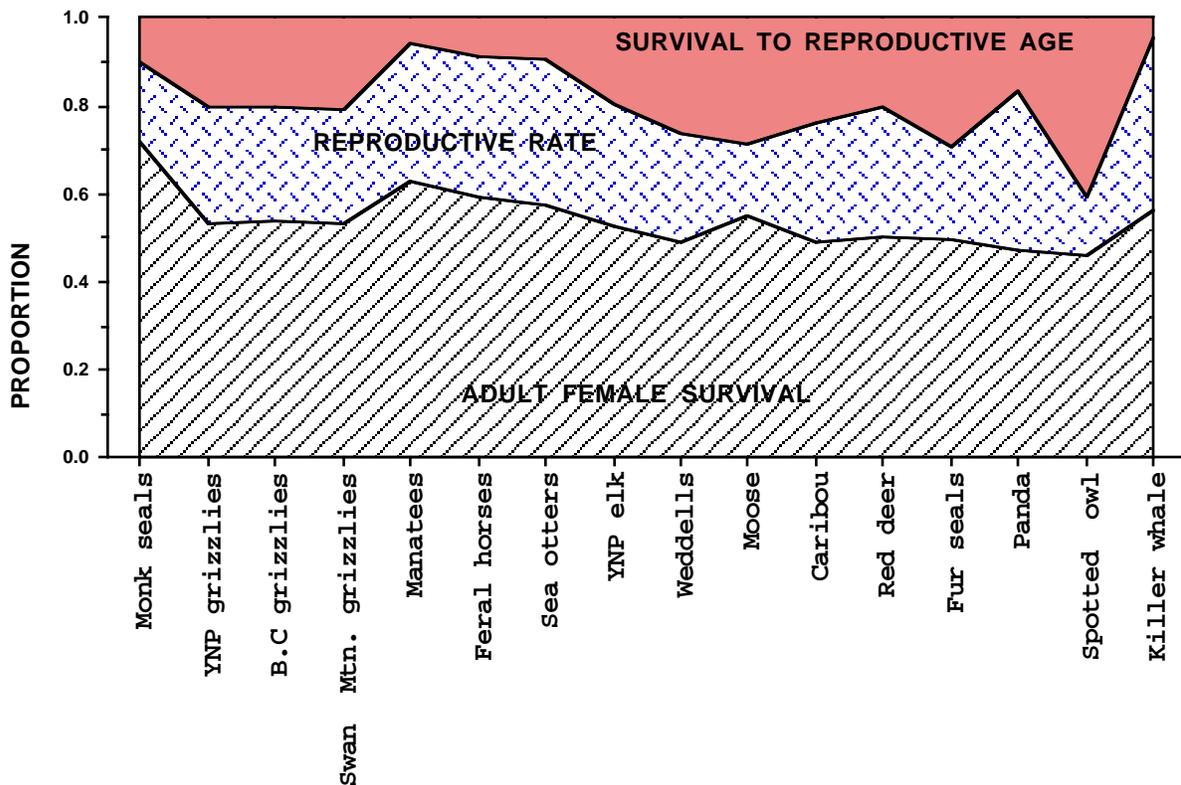


Figure 13.10 Relative importance of components of rate of population change for a number of species of long-lived vertebrates.

The evidence for reduced juvenile survival as the first parameter to change with increasing population density is now quite substantial. Gaillard et al. (1998:Table 1) reported that juvenile survival was implicated in 7 of 8 species for which there was evidence of density dependence. They noted that Fowler's (1987) review showed juvenile survival involved in 15 of 21 species for which density dependence was reported. A major problem in assessing data from the literature is that the chronological sequence in which various parameters change often cannot be determined from the reports, but it does seem clear that changes in the age of first reproduction largely follow as a consequence of the reduced resource availability leading to reduced early survival.

Gaillard et al. (1998) presented data that showed much greater relative variability (coefficient of variation) in juvenile survival (their Table 3) as contrasted to that of adult female survival (their Table 2). The likely cause of the higher variability in juvenile survival was illustrated by Gaillard et al. (1998:Fig. 1) as the sequence in time of adult female and juvenile survival

rates. As density increased, juvenile survival decreased sharply, thus substantially increasing the coefficient of variation. Gaillard et al. (1998) noted that "only two island populations (Soay sheep and red deer) exhibited density dependence in adult survival" and that "Even for these two species, survival of adult females varies much less than juvenile survival." Discussing red deer, Saether (1997) pointed out that "the major density dependent mortality losses occur among calves during their first winter", and that "This pattern was found in three different studies of the species in both Europe and North America". Gaillard et al. (2000) reviewed a substantially larger set of data that confirmed the results of Gaillard et al. (1998) and ranked variability from highest in juvenile survival to lowest for adult survival.

Clutton-Brock et al. (1991) reported that "The persistent instability of the St. Kilda [Soay] sheep population is probably caused by their consistently high fecundity associated with a super-abundance of food in summer. In conjunction these may permit the population to pass through winter at a level close to carrying capacity; to increase as much as 50% during the subsequent summer; and to enter the next winter at a level substantially higher than the island can support".

A major problem in assessing the sequence proposed here is the uncertainty as to the forces that regulated undisturbed populations before human interference became all-pervasive. The relative magnitudes of the numerical values of the essential parameters (Fig. 13.10) indicate that "sensitivity" of λ to parameter changes is least in early survival and greatest in adult survival, with reproductive rate in an intermediate position. It might thus be supposed that this sequence may have some significance in an evolutionary sense. The notion of "r and K selection" has lost some of its earlier preeminence, but nonetheless provides a convenient shorthand classification of species, with r-selected groups able to recoup losses in a very short period, and K-selected species requiring many years to recover from reductions in numbers. Consequently, it seems logical that selection in K-selected species might favor a regulatory process that tends to maintain equilibrium numbers by small, rather than large, changes in λ , hence by modifying early survival. This argument is supported by the results of Gaillard et al. (2000) who evaluated temporal variation in "fitness components" of large herbivores. These authors found that the coefficient of variation for early survival was highest (0.64) in large herbivores and lowest for adult survival (0.09). They suggested that "the immature stage ... may be the critical component of population dynamics for large herbivores", indicating that "our review supports Eberhardt's hypothesis and generalizes it to all sources of temporal variation".

Three major forces can be proposed as influencing the large vertebrates under "natural" conditions: weather, resources and predation. The preponderance of evidence from demographic studies indicates that early survival responds first to both severe weather and to resource limitations. As a population begins to outstrip its resources, inclement weather may have a much more pronounced effect, and may result in rather sharp fluctuations. A key question is the possible regulatory role of predation. In most natural systems human interference severely limits opportunities to evaluate the role of predation (Section 12.5). Wolf (*Canis lupus*) predation on ungulates has been observed to dominantly affect early survival and survival of senescent individuals (Peterson 1977, Peterson et al. 1984:Fig. 5, Ballard et al. 1987:Fig. 16,

Mech et al. 1998). To serve as a regulatory function, predation should exhibit a direct relationship with prey density, increasing at high prey densities and decreasing at lower prey numbers. The notion of ratio dependence (Matson and Berryman 1992) may fulfill this requirement. Evidence has accumulated to favor the ratio dependence model in wolf-ungulate interactions (Eberhardt 1997, 1998). The crucial question may then be whether predation tends to hold ungulate populations below levels where weather and resource limitations become important. If so, then an evolutionary pattern may be implicated.

Example 13.1 Sequence of events in density dependence.

The illustration of the original development (Eberhardt 1977:Fig. 2) of the sequence of events indicated in Fig. 13.8 used data from a number of sources. The Hawaiian monk seal population of French Frigate Shoals (Section 14.5) illustrates the full sequence in a single population. Trend of the French Frigate Shoals population through 1993 appears in Fig. 1.18, and was contrasted with the other sites by Gilmartin and Eberhardt (1995). The population increased steadily until sometime in the 1980's, and then began a dramatic decline. The decline was initially evidenced by finding malnourished pups, and a decreasing first-year survival (Fig. 13.11).

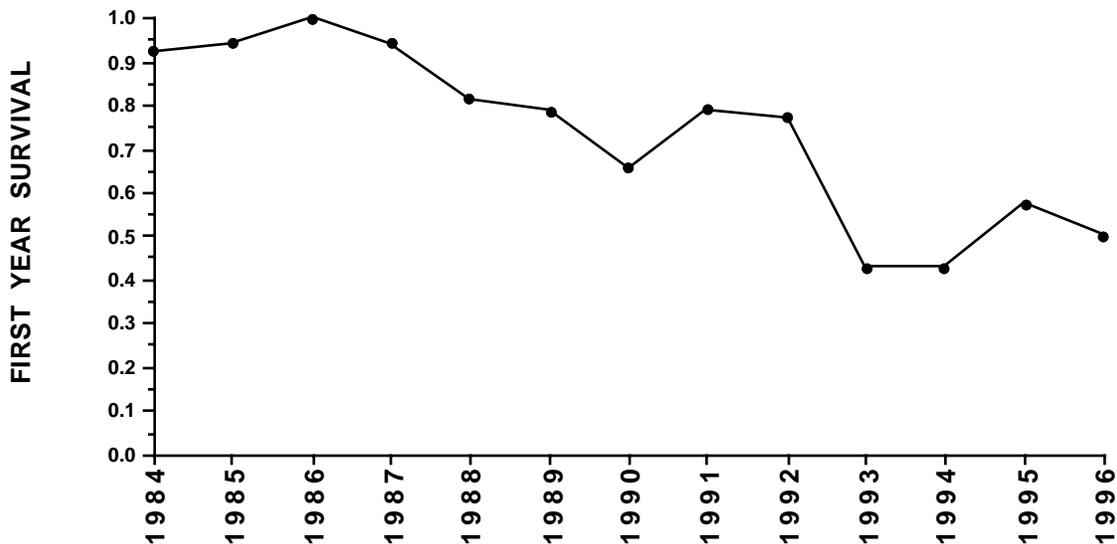


Fig. 13.11 First year survival of monk seals on French Frigate Shoals.

High adult survival was recorded during the phase of population increase (Gilmartin et al. 1993), but soon began a steady decrease. The population is not a large one, so the number of weaned pups tagged each year (a major effort is made to tag all pups at weaning) has averaged around 40 individuals. Consequently sample size has an appreciable effect on the numbers of older animals located (less than 10 individuals in the data used here), as does the stochastic (chance) variation in survival. Nonetheless, the pattern of decreasing survival is very evident in logarithmic plots of numbers observed over the years (Fig. 13.12).

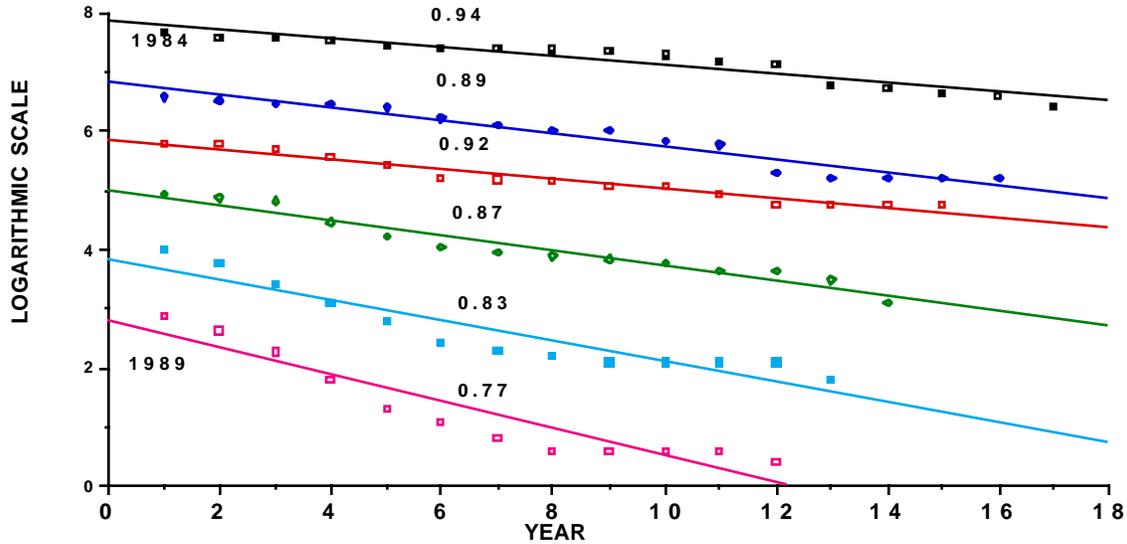


Fig. 13.12. Logarithms of numbers of tagged seals recorded at French Frigate Shoals, Northwest Hawaiian Islands, 1984 through 1989. Estimates of survival rates are shown for each year-class. Because the numbers initially tagged (as weaned pups) are roughly the same, the vertical scale has been adjusted so that the points do not overlap.

The survival rates given in Fig. 13.12 were obtained from the slopes of logarithms of numbers of individuals returning to the Atoll each year. More details on survival estimation appear in Gilmartin and Eberhardt (1995). It appears that there has been a steady decrease in adult survival rates, and this is, of course, evident in the population trend (Fig. 1.18). It is particularly interesting that the first cohort tagged (1984) apparently continues to have relatively good survival. The decreasing trend in survival rates has continued, so that the cohort tagged in 1996 exhibits an annual survival rate on the order of 40%. Evidently the available food resource for this population has continued to decline, and "carrying capacity" may not yet have been reached.

Due to relatively small sample sizes, it has not been feasible to determine whether the sequence of events of Fig. 13.8 has occurred in this population in precisely the order suggested, but clearly the prescribed series of events has transpired. Age of first reproduction clearly has been reduced, and it seems reasonably certain that the overall reproductive success of adult females has decreased (Fig. 13.13). The oldest tagged individuals in the overall population are now about 18 years of age, so it is as yet to be determined when senescence becomes important. The largest sample of reproductive data comes from another site (Laysan Island, see map in Section 14.5), and suggests that monk seals do not reach a reproductive peak until appreciably later than other pinnipeds.

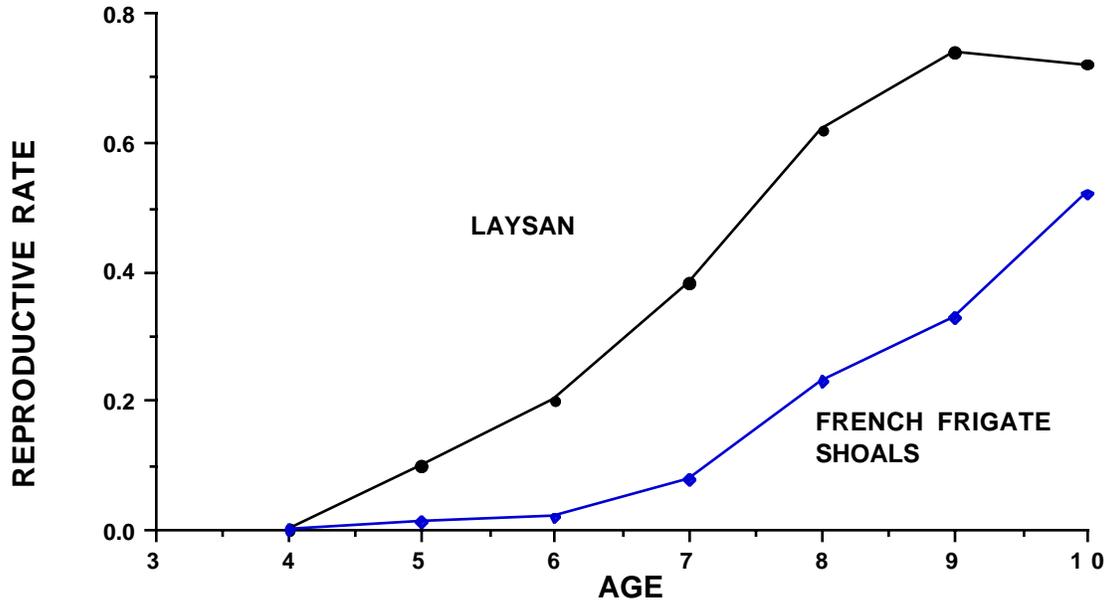


Fig. 13.13 Reproductive rates for monk seals at Laysan Island and French Frigate Shoals.

It should be noted that this example very likely is not typical, in the sense that there is not sufficient information to determine what the "typical" approach to an asymptotic population level might be. As noted previously in this Chapter, it may be that a reduction in first year survival and the (likely) concomitant delay in first reproduction may be sufficient to cause a population to stabilize in the neighborhood of a carrying capacity value. This is one of the issues that characterize the limitations of the present state of knowledge about long-lived vertebrates.

13.7 Exercises

13.7.1 Make plots of the rate of increase used in Eq. (13.1) and Fig. 13.1 for $r(1)=0.2$, $z=1,2,5,11$, and $K=5000$.

13.7.2 A model for density dependence.

Plot eq.(13.3) with $b=0.95$ and $K=5000$ and the log form of the generalized logistic [eq.(13.1)] with $r(1)=0.2$, $z=5$, and $K=5000$. Note that you will need to rearrange eq.(13.1) to put it in a log form, and that you will need to use $\exp(\ln X(t))$ to represent $N(t)$ in the portion of eq.(13.1) in brackets, inasmuch as $\exp(\ln(x))=x$.

13.7.3 Generating normal distributions.

A convenient way to generate normally distributed errors for studying density dependence is given in the following equations:

$$x_1 = (-2 \log_e r_1)^{1/2} \cos 2\pi r_2$$

$$x_2 = (-2 \log_e r_1)^{1/2} \sin 2\pi r_2$$

where the x 's are random normally distributed values with mean 0 and variance 1 and the r 's are uniform random variables (range 0-1). You can obtain uniform random variables with the function `RAND()` in EXCEL (note that you need the parentheses to generate uniform random variables but don't

need to enter a value in them). The above equations generate 2 random normal variables from 2 uniformly distributed values. In doing largish simulations its convenient to use both of the random normal numbers thus generated. For purposes of illustration, we need only use one. Set up a table of random normal numbers using the above equations and try the random walk of Fig. 13.3, using 100 values. Create a graph on the same EXCEL sheet and change the values to see how the random walk changes. Note that the populations go extinct fairly regularly, whereas Fig. 13.3 shows an increasing population. EXCEL has a command that changes the random numbers to make a new graph. on occasion it is desirable to make the recalculation manual so as not to have the graph change while you are working on it.

Exercise 13.7.4 “Random walk with drift”.

Several of the papers referenced in this chapter discuss “random walk with drift” using eq.(13.4). It was noted in connection with eq.(13.4) that this really simulates an exponentially increasing population. Add another column to the random normal numbers generated above which multiplies one of the random normal numbers by a constant so that you get an error term like that used in eqs.(13.5) and (13.6). Then add a constant to the random walk model above to construct results according to eq.(13.4), as shown in Fig. 13.4. Make a graph of the random walk model on the same EXCEL sheet. Vary the constant multiplying the normal random variable from, say, 0.1 to 1 and note how it affects the graph of population size. With a small multiplier the graph should essentially follow an exponential plot (but on a log scale). With a larger multiplier it becomes quite erratic.

Exercise 13.7.5 Rapid rates of increase.

It was noted in the Chapter that many of the papers on testing for density dependence appear to be based on data on insects. To examine this prospect, plot eq.(13.2) as shown in Fig. 13.2 with $K=5000$, but set it up with 3 columns and make a graph so you can compare different values of b . Try $b=.95$, and $b=.98$ which should replicate Fig. 13.2 and part of Fig. 13.7. Now try values of $b=0.5$ to 0.8 to see how different the growth curves are.

Exercise 13.7.6 Lambda for rapid rates of increase.

Add 3 columns to the worksheet for Exercise 13.7.5 to show lambda for each of the population plots. This can be done with the equation given in the chapter, i.e.

$$\lambda_t = \frac{N_t}{N_{t+1}} = N_t^{b-1} K^{1-b} = \left(\frac{K}{N_t}\right)^{1-b}$$

This shows why the growth curves of exercise 13.7.5 increase so rapidly.

Exercise 13.7.7 The generalized logistic.

Calculate the generalized logistic (eq.13.1) for $K = 5,000$, $z=5$, $r=0.2$, 0.4 , and $K = 5,000$, $z=2$, $r=0.2$, 0.4 , and make a graph of the results on the same worksheet. Note the more rapid increase generated by $r=0.4$. One characteristic of the generalized logistic is shown by the “overshoot” and oscillation for higher rates of increase and larger values of z .

Exercise 13.7.8 Bulmer's tests

Some of the difficulties in testing for density dependence can be illustrated by considering Bulmer's (1975) tests. The first test is based on the fact that the squared differences between successive observations of population size can be used to approximate the overall variance of a set of observations, if there is no trend or pattern in the data. The test is essentially the Durbin-Watson test of eq. (9.13), except that Bulmer used the reciprocal of the ratio, i.e., the reciprocal of eq.(9.13). The table below gives the first 20 observations from the generalized logistic used in Exercise 13.7.2, 20 observations bracketing the inflection point, and 20 observations above the inflection point, along with logarithms of 20 observations from an exponential growth curve with $N(0)=10$, $r=0.20$.

First 20 obsns generalized logistic				
	5 OBSNS 0.2 BRACKETING 5000 INFLECTION 10 POINT	OBSNS ABOVE INFLECTION POINT	EXPONENTIAL MODEL N(0)=10 R=1.2	
X(t)	X(t)		X(t)	
1	2.30259	7.04288	8.51718	2.30259
2	2.48491	7.22509	8.51719	2.48491
3	2.66723	7.40715	8.51719	2.66723
4	2.84955	7.58883	8.51719	2.84955
5	3.03187	7.76954	8.51719	3.03187
6	3.21419	7.94789	8.51719	3.21419
7	3.39651	8.12049	8.51719	3.39651
8	3.57884	8.27961	8.51719	3.57884
9	3.76116	8.40979	8.51719	3.76116
10	3.94348	8.48962	8.51719	3.94348
11	4.12580	8.51505	8.51719	4.12580
12	4.30812	8.51718	8.51719	4.30812
13	4.49044	8.51719	8.51719	4.49044
14	4.67277	8.51719	8.51719	4.67277
15	4.85509	8.51719	8.51719	4.85509
16	5.03741	8.51719	8.51719	5.03741
17	5.21973	8.51719	8.51719	5.21973
18	5.40205	8.51719	8.51719	5.40205
19	5.58437	8.51719	8.51719	5.58437
20	5.76669	8.51719	8.51719	5.76669

Bulmer's first test is $R=V/U$ where:

$$U = \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 \quad \text{and} \quad V = \sum_{i=1}^n (x_i - \bar{x})^2$$

He gives a way to calculate significance levels, which gives a value of 0.696 for the 1 % point and 0.909 for the 5% point, where the test is significant if the R is less than the stated significance point. Calculate Bulmer's R test for the first three columns above and comment. Also comment on the meaning of a comparison of the first and last columns (note that the data are available in your results for Exercise 13.7.2 so you don't have to copy the numbers above).

Exercise 13.7.9. Bulmer's second test.

Bulmer (1975) recognized that his R-test would not behave satisfactorily in the presence of sampling error so devised a second test for that purpose, which is $R^* = W/V$ where:

$$W = \sum_{i=1}^{n-2} (x_{i+2} - x_{i+1})(x_i - \bar{x})$$

He gives approximate criteria for significance as:

$$R_{0.05}^* = -\frac{13.7}{n} + \frac{139}{n^2} - \frac{613}{n^3}$$

$$R_{0.01}^* = -\frac{20.1}{n} + \frac{258}{n^2} - \frac{1279}{n^3}$$

and indicates significance if R^* is less than the calculated significance level. Try this on the first 20 observations for the generalized logistic as given in Exercise 13.7.8 and comment on your results.

14.0 SOME CASE HISTORIES

14.1 Introduction

A substantial number of quantitative techniques have been presented in the preceding chapters. Because these methods are mainly supported by mathematical equations of various kinds, the presentation tends to suggest an exact methodology. All the user has to do is supply numbers for the equations and accurate results with appropriate confidence limits should become available. However, the real world is not like that. The techniques presented here must largely be applied in field settings and in circumstances where various uncertainties arise. Furthermore, work with the populations inhabiting large areas is always expensive and time-consuming, and various sources of bias may exist and may not be identified or corrected. Consequently, this chapter is devoted to reviewing some "case histories" to illustrate some of the problems encountered in real-world settings. Also, data from a number of the cases discussed below were used to illustrate various methods in the previous chapters, so the following sections provide some background information as to how and where the data were collected.

14.1 The Florida manatee (*Trichechus manatus latirostris*)

Manatees are slow-moving warm-water herbivores that produce a single calf at two-year intervals under good conditions, with part of the population giving birth at three-year intervals. Reproduction does not occur until age 4 or 5. Aerial surveys produced an estimate of about 1,000 manatees in Florida in the 1970's and early 1980's. Most adults carry characteristic scars from encounters with speedboats, which appear to be the principal source of mortality. A continuing increase in boating use exacerbated the situation. Concern about the accidental death rate led to initiation of a program to recover and examine carcasses in the mid-1970's. When the number recovered reached 200 individuals per year, it was widely assumed that the manatee population might soon be extirpated as the population clearly could not sustain an annual loss rate of even 10% let alone 20%. This "crisis" situation continued for another 15 years or so, with annual announcements of the imminent demise of the species. Efforts to recover carcasses were accelerated, and the aerial surveys were continued. Aerial counts are most effective during unusually cold periods, when manatees concentrate in warm water areas and are thus more readily seen. By concentrating on unusual cold-snaps and using a sizable number of aircraft, the total count was increased to approximately 2,000. However, the carcass recovery program also turned up more deaths over time, resulting in continued concern about prospects for the population. Efforts to determine survival rates and to better understand the available trend data eventually led to the conclusion that the population had, in fact, most likely actually been increasing throughout the period (Eberhardt and O'Shea 1996).

The problem was that both sources of data (aerial surveys and carcass recoveries) only accounted for some unknown fraction of the numbers present and the annual losses. At present, there is no really suitable estimate of the "efficiency" of either survey nor of the total number of manatees present in Florida waters. With the realization that the population actually was not in any real danger, it is quite possible that a relaxation of the controls on boating in manatee areas might eventually lead to a sharp reduction in the

population without that event being detected by the two major surveys. Continued close attention to survival rates is thus indicated. This is not a simple matter due to the fact that scar patterns had to be used for marking. We thus have a situation where two surveys (one for population size and one for mortalities) are biased (underestimate) and there is as yet no suitable means for correcting the results. Given a suitable technique for marking, it seems likely that improved survival estimates can be obtained and will offer some protection against the very real risk that the population may decrease substantially without that fact being detected until too late.

14.2 The Pacific walrus (*Odobenus rosmarus divergens*).

The population of the Pacific walrus occupying the Bering and Chukchi seas was severely depleted by whalers starting in about 1848, and continued to be overharvested beyond the turn of the century, resulting in deaths by starvation of an appreciable part of the coastal Eskimo populations. A partial recovery was again heavily harvested for commercial purposes, starting in the 1930's. The Soviet Union prohibited taking of females by their hunters in 1957, and the population recovered to what likely was an asymptotic level by about 1980. Reproduction appeared to be substantially reduced about that time, and harvests for subsistence purposes and for ivory may well have again started an overall decline, and quite likely have reduced the male segment of the population rather sharply.

A number of efforts have been made to estimate total abundance of the walrus population through aerial surveys, but it has not been possible to demonstrate that any of these surveys have approximated the total number of walrus present due to the very bad weather often present at the times when walrus are concentrated and available for counting. The harvests are very selective, so that it is extremely difficult to determine composition of the population, and the harvest estimates depend very much on reports by hunters. Walrus ivory is valuable so there is a risk of overharvesting. Reproductive rates are roughly comparable to those for manatees, so a reduced population may take a long time to recover. Thus, while we know that there are large numbers of walrus, the prospects for ascertaining total abundance or trend are not encouraging, and the population very likely can not be adequately understood without very large expenditures of time and money.

There is evidence that walrus depend heavily on clams and that the prey species may be about as long-lived as the predator. It is thus quite possible that the extensive commercial utilization of walrus and great reductions in the population in the past may have resulted in a build-up of prey to the point where the recovering walrus population may well have "overshot" the steady-state carrying capacity of the environment in the 1980's, resulting in the observed sharp reduction in walrus reproductive rates. A very substantial fluctuation in relative biomass in the Bering-Chukchi region may well have occurred without anyone being aware of details, and the cycles might well persist over long time periods. There may be an appreciable potential for eventual human harvests of the molluscan resources of the region, leading to the necessity for the very difficult job of assessing walrus populations in detail. Thus far there are no good indices of trend, and no way to reliably estimate survival rates.

We thus have an example of a very large population [perhaps as many as 250,000, Fay et al. (1997)] that most likely has fluctuated substantially since the 1850's, with very little reliable information as to current size or trends, and few immediate prospects for improving the situation.

14.3 Bowhead whales (*Balaena mysticetus*)

At each of its meetings from 1977 to 1982, the Scientific Committee of the International Whaling Commission recommended that no bowhead whales be removed from the stock inhabiting the Bering, Chukchi, and Beaufort seas. Commercial whaling began on this stock in 1848 and continued to 1914, greatly reducing the population. From 1914 to 1970 an average of about 12 whales per year was landed by Alaskan Eskimos, so that cessation of these harvests was considered undesirable for cultural and subsistence reasons. After 1970, however, the number of whales landed increased to about 30 per year, with an increase also of whales "struck and lost", i. e., harpooned but not landed, so that as many as 100 whales may have been killed in 1977. A significant increase in the rate of removal on an already depleted population indicated a need to be concerned about possible extinction. A series of counts of northward migrating whales has been conducted in recent years, and utilized with catch data and a backcalculation model. These counts gave erratic results due to weather and ice conditions and various improvements in technique. Hence, only a single current population estimate was utilized in the model.

The bowhead calculations (Breiwick, Eberhardt and Braham 1984) provide an example of the complexities involved in attempting to assess a population trajectory with a minimum of information on reproductive and survival rates. Likely ranges of parameters had to be adopted, and many runs of the model used to explore the effects of various parameter combinations. The major concern at the time was whether or not the population would continue to recover under current and projected levels of harvest. Consequently, the analysis mainly had to evaluate the prospects for a continued decline, considering the history of harvests and current population size.

A reliable method for determining age was not available for bowhead whales, so that the available population data consisted of a classification as calves, immatures, and mature individuals. Sex ratios in the harvest and in the population appeared to be about unity so that particular complication was neglected. To utilize the information on fraction of mature animals in the population, it was necessary to use an age-structured model. Initially, it seemed realistic to combine the mature animals in a single class, thus using a reduced Leslie matrix. However, it soon became evident that high adult survival rates would bias the outcome unless lowered survival due to senescence was introduced. Hence a full matrix was used, truncated to approximate the effect of senescence. Consequently, we have the necessary but anomalous-seeming situation in which an age-structured model is used without knowledge of ages. Inasmuch a commercial harvesting had ended about 1914, and the population about 1970 was known to be at least 4,000 individuals, runs of the model with parameters in likely ranges (deduced both from the available biological data on bowheads, and on similar species) made it evident that a continuing decline was improbable under the current harvesting regime.

The important aspect of the bowhead situation for present purposes, is that the initial discussions of the problem revolved around the notion that there might only have been a few hundred whales present. The first serious efforts at a census resulted in estimates of about 1,000. Continued efforts at improving techniques ultimately pushed estimates up to nearly 8,000 whales. These efforts required about 20 years to accomplish, illustrating the fact that there is no simple way to deal with population problems in remote areas.

14.4 Grizzly bears (*Ursus arctos horribilis*)

The uncertain future of the grizzly bear in the coterminous United States resulted in a "Threatened" status under the Endangered Species Act. It is thus essential to monitor the population as closely as possible. The remaining subpopulations occupy forested habitats, are highly mobile, often secretive, and have very large home ranges. The subpopulation considered here is that of the Greater Yellowstone area, containing about 20,000 sq km, and centered on Yellowstone National Park. Direct determination of population size has thus far been very difficult, so that it has been necessary to depend on an assessment of reproductive and survival rates (obtained largely through telemetry) to assess rate of change. Reproduction occurs at about 3 year intervals, with an average litter size of about 2 cubs.

The major concern in the study is one of maintaining the isolated population in the Greater Yellowstone area. Open garbage dumps had provided a supplementary food supply since the 1920's, and likely served to concentrate the population seasonally. Closure of the dumps in the early 1970's resulted in very extensive mortalities, associated with interactions with humans in the course of seeking new food supplies. Radiotelemetry studies were initiated in 1975, and indicated a long-term decrease of about 2% per year (Knight et al, 1985), based on calculation of λ from reproductive and survival data. Estimation of survival rates has identified the key factor in maintaining the population. Adult female survival rates have been about 92%, whereas a population free of human interference should have a rate on the order of 99%. Hence protection of adult females was stressed, with a policy that a recorded loss of more than two adult females per year would be likely to jeopardize the future of the population.

Radiotelemetry data build up very slowly, so that in the initial analysis it was necessary to combine males and females in calculating subadult survival rates. As more data accumulated, it became evident that subadult male survival was appreciably lower than that of subadult females, and a revised calculation (Eberhardt et al. 1994, Eberhardt 1995) indicated that the population may in fact have been increasing. The reproductive and survival data yield an estimate of λ of 1.05, with confidence limits of about 1.00 to 1.09. Data for an index of relative abundance for Yellowstone grizzlies come from records kept of distinct family groups (females with cubs-of-the-year). Females with cubs are more readily seen than other bears, and the presence of cubs provides various clues useful in distinguishing one family group from another (number in the litter, size, coat color, radiotransmitters, etc.). Distance between sightings and various obstacles (mainly highways and the Grand Canyon of the Yellowstone River) are also important in distinguishing families. Details of the method appear in Knight et al. (1995). A major problem with using the index is simply that visibility of individual bears varies appreciably from year to year. In wet years with lush vegetation, bears are

able to find adequate food supplies in relatively heavy cover and are thus not seen so readily. In dry years, they are forced to range more widely, and thus are seen more frequently.

The question of total size of the Yellowstone population has been a recurring problem. Research on Yellowstone bears in the late 1960's was concentrated on the population utilizing garbage dumps where bears were readily accessible. Marking and tagging was used to produce an estimate of population size, which was reported (Craighead et al. 1974) to be about 230 bears. Later reports (McCullough 1981) used an estimate of about 300 bears, but this number has never been documented, apparently having arisen in unpublished correspondence after the study by Craighead et al. ended. Closure of the garbage dumps in the 1970's resulted in a great deal of controversy and termination of the Craighead study. Little or no field work was done until the mid-1970's, when the population was dispersed and very much changed in behavior and general character. Estimates from tagging and recaptures or resightings were not feasible, so that, as noted above, trend had to be approximated from reproductive and survival data from the radiotelemetry work, supported in a limited way by the index of relative abundance based on tallies of females with cubs of the year.

Attempts to estimate total numbers during most of the recent study depended on projections of the population size estimates from the work by the Craigheads, and the heavy losses recorded during the period of dump closures made it seem that the population must have reached a low level. Administrative and public requests for an estimate ("How many are there?") led to construction of minimum estimates in the recent study, starting with the tally of adult females with cubs. Although these estimates were described as minimums, many public reports treated them as estimates of actual numbers. Ultimately it became possible to utilize the accumulated data on sightings to produce a direct estimate of recent numbers (Eberhardt and Knight 1996). This data suggested that the current population may be on the order of 400 bears. Although the variable nature of the available data resulted in wide confidence limits, it does seem likely that the estimates of the late 1960's were biased downwards by the use of data collected mainly at the garbage dumps, and that the population was then larger than supposed. Mortalities associated with closure of the dumps thus very likely did not result in as abrupt a decrease in numbers as seemed apparent at the time, and the recovery may have been much more rapid than generally assumed. Focussing on total numbers thus seems to have greatly confused the issues.

The intense controversy over dump closure resulted in the abrupt termination of the Craighead study with little or no monitoring of marked bears. The subsequent study did not begin marking until 1975, and then initially only outside of Yellowstone National Park. It thus took nearly 10 years to build up sufficient information to approximate a population analysis (Knight and Eberhardt 1985). Had the early study been continued throughout the period of dump closure, the long interval of uncertainty about status of the population might have been reduced appreciably. However, the Craighead study depended largely on visual resightings of marked bears at the dumps, while the recent study has utilized radiotelemetry of bears trapped at various places throughout the Park. Whether continuation of the field studies during the dump closure would have reduced the confusion is thus not certain. It does seem evident that the emphasis on population size led to continued confusion

and uncertainty. To estimate the actual numbers of bears in the Greater Yellowstone area by a capture-recapture technique would be extremely expensive and might not succeed. Live-trapping would be required which presumably calls for moving traps by helicopter (not permitted in Wilderness Areas, which are an important component of the overall bear habitat). Because female survival is crucial, it may well be that this is the essential parameter to monitor.

14.5 Hawaiian monk seals (*Monarchus schauinslandi*)

The Hawaiian monk seal is known to be sensitive to human intrusions, and has been sufficiently reduced in numbers to become classified as "Endangered" under the Endangered Species Act. An as yet poorly understood series of events resulted in redistribution of the population over the several atolls of the Northwestern Hawaiian Islands which constitute the sole habitat of the species). Abundance was sharply reduced in the western part of the chain, leaving a severely imbalanced sex ratio, with adult males substantially outnumbering females in at least two sites. The excess of males has resulted in a breeding phenomenon known as "mobbing" behavior in which females may be killed outright or injured sufficiently to lead to death from shark predation. Population trends are available as "beach counts" and survival and reproductive data are available from data on marked seals.

Six primary sites contain the majority of the population. Midway Atoll was subjected to intense military activity in World War II, and was used extensively by the U. S. Navy until recent years. This intense human activity essentially extirpated the monk seal population, with mean beach counts of about 50 seals observed in 1956 and 1957. Apparently these were very old seals, and they were mostly gone by the late 1960's. A few seals moved in from the adjacent sites (Kure Atoll and Pearl and Hermes Reef), so that mean beach counts on the order of 10 seals were observed recently. All of the other major sites, with the exception of the easternmost, French Frigate Shoals, exhibited a sharp decline beginning in the 1960's. It is likely that there was a common agent external to the individual populations that caused the decline, but a specific cause has not been identified. Occupation of Kure for operation of a Loran station is known to have contributed to the decline there, and efforts were made to rebuild the population by introducing young seals rescued from French Frigate Shoals and "rehabilitated" in captivity for a year. A disease of unknown origin turned up in seals kept in captivity for use in the rehabilitation effort, resulting in cessation of that activity.

A report on status of the population (Gilmartin and Eberhardt 1995) illustrates the sharp decline in the early years (1960's). Currently, the smaller, western populations (Kure, Midway, Pearl and Hermes Reef) appear to be increasing, while the two intermediate populations (Laysan and Lisianski) are static or decreasing slowly. The French Frigate Shoals population increased while the other populations were decreasing in the 1960's (tagging shows only minor movements between individual populations), peaked in 1988 or 1989, and then dropped dramatically in consequence of persistent starvation of weaned pups.

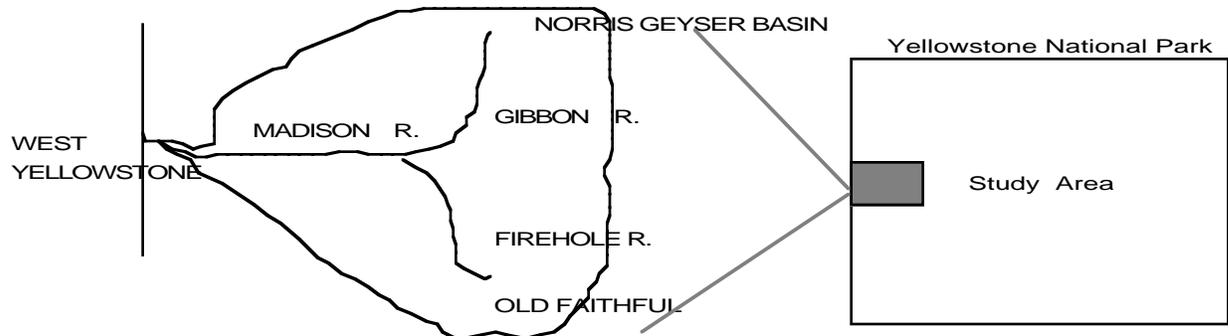
The monk seal data are particularly interesting because there is no clear evidence of causes of the initial major decline. Because adult sex ratios were dramatically shifted (about 3 adult males per adult female) and subsequent data show the usual large mammal pattern of lower male survival, one can only conclude that the decline resulted from some human activity, but this conclusion is unsupported by data (and largely not discussed in the literature or in reports on monk seals). The essentially static population on Lisianski and Laysan is characterized by reduced adult female survival, but again no cause has been identified. The "crash" at French Frigate Shoals is clearly due to resource limitation, but again the cause is unknown. Quite recently monk seals have begun to move into the main Hawaiian Islands. They have not previously been known to occupy these islands, but may simply have been extirpated by the earliest Polynesian inhabitants, without any archeological or other record becoming available as yet.

14.6 The Madison-Firehole elk herd

The study and data described below are included to provide an example of circumstances where a natural population has remained remarkably constant in the face of severe and variable winter conditions.

The study system

The Madison-Gibbon-Firehole drainage (general location sketched below) consists of a rugged mountainous landscape with elevations ranging from 2150-2800 meters. The pronounced topography results in variable aspects and slopes supporting a diversity of plant communities, including mature conifer forests, sedge-grass meadows, and aquatic communities. Numerous geyser basins warm low-elevation meadows and rivers, thereby reducing snow cover in these areas and enabling some unique plant associations to continue photosynthesis throughout the winter. Large scale fires during 1988 burned 55% of the drainage, creating a complex mosaic of burned and unburned forests at different stages of succession. Winter climatic factors are an important abiotic component of the ecosystem, as snow depths frequently exceed 1.2 m, drastically reducing food availability and producing severe energetic bottlenecks for herbivores during the winter.



Sketch map of the study area, which includes the Madison, Gibbon, and Firehole River drainages.

The approximately 500-600 elk residing in the area are nonmigratory, remaining within the borders of the Park throughout the In the 10 years of study to date, no radio-marked elk has left the study area, nor have any of these elk moved outside the 3 river drainages constituting the study area (based on approximately 7,000 radio-locations). Neither the landscape nor the elk that inhabit this drainage have been manipulated to any significant extent since European settlement, hence the elk population is regulated solely by natural processes. This is a particularly unusual feature of this study system since nearly all large herbivore populations in North America are harvested by man, including the previously studied Northern Range elk population of Yellowstone National Park (Houston 1982). Human harvest regimes dramatically influence nearly all demographic attributes and processes within a population. The only anthropogenic influence that may have appreciably impacted the natural dynamics of the Madison-Firehole elk population is the extirpation of wolves from the Park in the early part of this century. Reintroduction of wolves was initiated in 1995, with the first releases of wolves into the Madison-Firehole drainages occurring in spring 1996. Elk and bison are the only ungulates with substantial numbers in the area. A few mule deer and an occasional moose may be seen, but do not constitute a significant prey base.

Population data

In the fall of 1991, 5 years prior to initiation of wolf reintroductions in Yellowstone National Park, intensive telemetry-based investigations of the landscape-use patterns and demographic characteristics of the elk population were initiated. The sedentary nature of these animals, combined with the protection afforded by the Park and frequent contact with humans, has resulted in animals throughout this drainage becoming relatively tolerant of humans. This tolerance greatly facilitates opportunities to economically capture and mark animals using ground-based delivery of immobilizing drugs via dart rifle. Once marked, animals can be frequently monitored with hand-held telemetry systems and observed for quantification of survival, cause of death, habitat use and distribution patterns, feeding behavior, and the collection of snow-urine and fecal samples for indexing nutritional and reproductive status. These ongoing studies represent the most current and extensive databases on elk in Yellowstone National Park.

A sample of 30-40 females, ranging from 1 to 15 years of age and instrumented with radio collars has been maintained throughout the study.

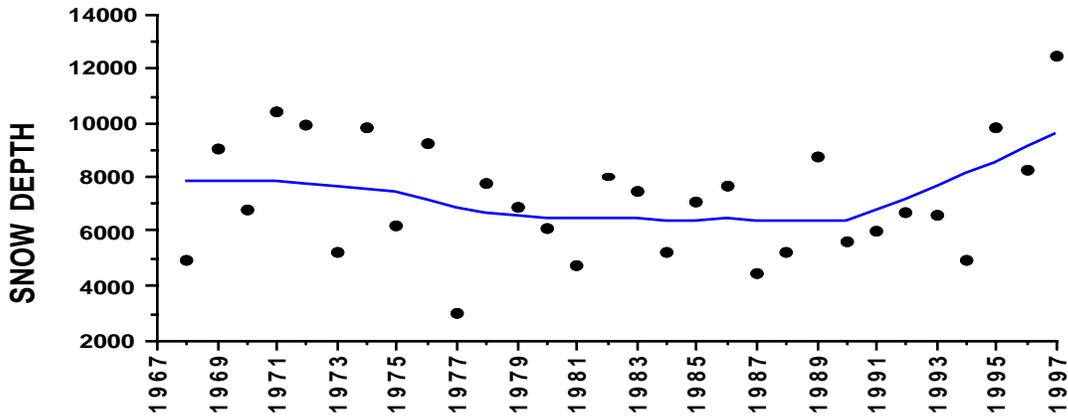
Intensive field studies, involving one or more graduate students and 1-2 technicians have been conducted annually from December thru April, with low-level monitoring of all animals occurring through the rest of the year. Data on survival beyond age 1 has been accumulated along with reproductive performance. Very high pregnancy rates have been observed, along with high survival beyond age 1 for the first 5 years of the study. Transmitter performance has been nearly perfect, and no mortalities were observed in the first 3 years of the study. These data provide an annual survival estimate for prime-age females (1-11 years of age) of 97%. Survival rates for prime age classes in the study were much the same as those reported by Houston (1982) for the northern Yellowstone elk herd, but dropped off sharply after age 11, while Houston's data suggest high survival continuing out to age 16 or beyond. Preliminary findings from the Madison-Firehole study indicate that high fluoride and silicate concentrations in area vegetation may induce accelerated tooth wear and thus early senescence.

Pregnancy rates of all instrumented elk were also assessed each year using serum and fecal steroid assays providing a pregnancy rate estimate of 92% for animals >2 years of age. As expected, yearling pregnancy was lower, estimated at 40%. Both survival and pregnancy rates were reduced during the 6th year of the study (1996-97), which was the severest winter recorded in the past 60 years. Reproductive rates observed in the Madison-Firehole area also closely approximated those for the northern herd.

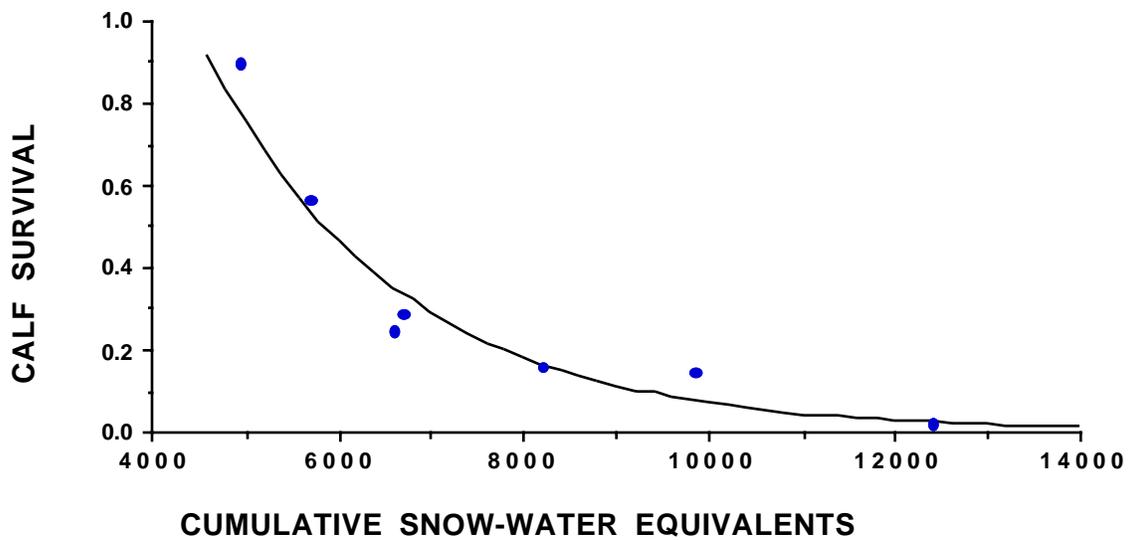
Population size was estimated from aerial surveys conducted during the first 4 years of the study. These surveys utilized the radiocollared elk and Petersen estimates. In April of 1997, estimates using ground counts were begun. These counts were conducted just after the meadows became snow-free, and elk were concentrated at these sites. Petersen estimates were obtained from 10 such ground surveys conducted in the spring of each year, and used to estimate population size after winter mortality had ended. Because survival over the summer and early fall is virtually 100%, similar counts in October were used to estimate recruitment of calves born in the summer, and thus approximate a total population size.

The estimates of recruitment and early survival depend on large samples (1,000 to 1,500 observations each year) of calf-cow ratios, taken during the months of December through April. In all years calf-cow ratios decreased from about 42-48 calves per 100 cows in December to April ratios ranging from 38 to <1 calves per 100 cows. The April ratios are closely related to cumulative snow depth data, with the highest observed ratio associated with the year (1993-94) of lowest snow depth. Because adult cow survivals were very high, the drop in calf-cow ratios is evidently mainly a function of over-winter calf mortality.

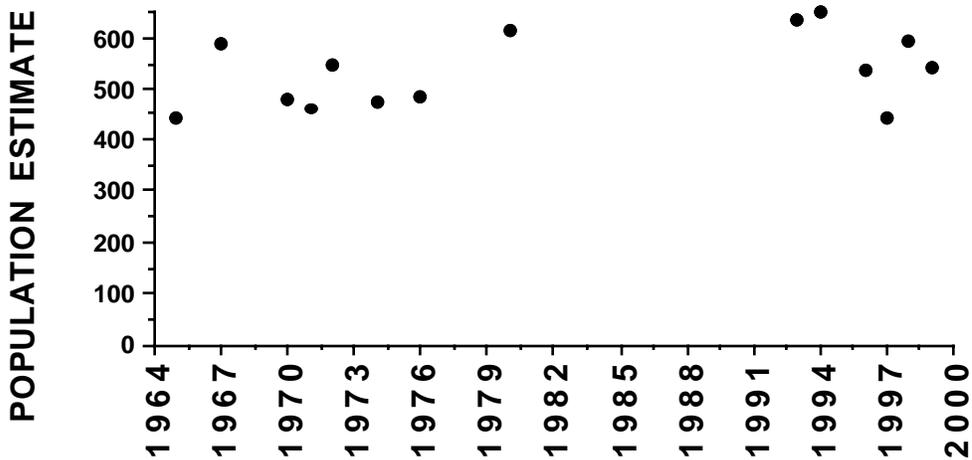
Snow depth records (an automated weighing system is used) are available for the last 30 years (see graph below). These records can be used to demonstrate that the calf-cow ratio decreases markedly with increasing spring snow depths. The mild winter of 1994 was associated with a high spring calf-cow ratio (38), while the severe winter of 1996-97 resulted in the lowest such ratio reported (<1 calf per cow in the spring).



Snow depths recorded over the last 30 years in the Madison-Firehole study area.



Calf survival and spring snow depths, measured as snow-water equivalents. An exponential curve was fitted to the data.



Population trend of Madison-Firehole elk population.

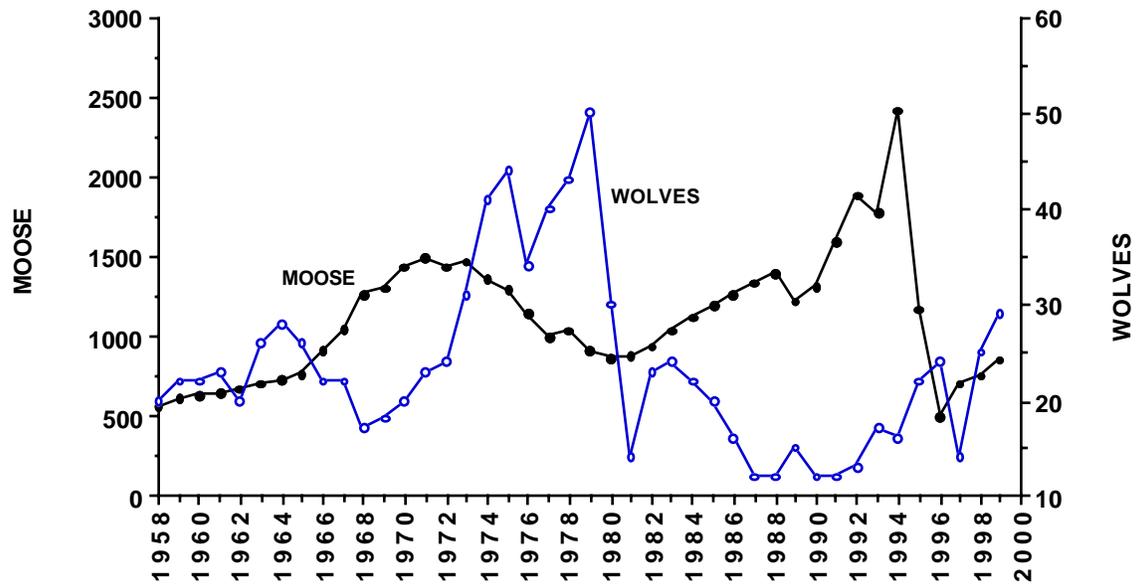
In spite of the dramatic effect of over-winter snow depths on calf survival, the population has remained remarkably constant due, we suspect, to a density-dependent mechanism which most likely results from a strict limitation on numbers of individuals that can be accommodated in the thermal refugia. This impact appears to fall principally on calves, inasmuch as adult cow survival has remained high. Wolves have begun to take significant numbers of calves, and some adult elk, so it will be particularly interesting to see how they influence this relatively stable population.

14.7 Isle Royale wolves and moose.

An appraisal of the Isle Royale moose and wolf data using eqs. (12.34) and (12.35) suggests the possible utility of these equations for further wolf-ungulate studies (Eberhardt 1998, Eberhardt and Peterson 1999). Further examination of the data suggests that eq. (12.35) seems to be roughly satisfactory as a model for the moose data, at least up to the point where a tick epizootic influenced the population for several years (starting in 1989), but the ratio-dependent model for wolves does not give a very good fit to the trend of the Isle Royale wolf population. Hence, the issues of how to test such a fit need to be further examined with the data. The concept of ratio-dependence has been criticized in the literature (as well as defended), and thus needs further attention. It is also quite possible that the Isle Royale wolves may be influenced by the limitations of an island (544 km. sq.) habitat, and be responding to that limitation to the point where it limits utility of the ratio-dependent model.

Several papers (Messier and Crete 1985; Messier 1994, Gasaway et al. 1992) have proposed models for wolf-ungulate interactions having two equilibrium states, one (the upper) induced by resources available to the ungulates, and the second (lower) presumably induced by predation. The data thus far analyzed do not support these models (Eberhardt 2000), and need to be further assessed, along with the prospects for a "predator-pit" type of lower equilibrium. The analysis will necessarily be somewhat speculative for lack of data on multiple-prey situations, but it does seem that there is enough data to raise questions about the possible existence of such an equilibrium in wolf-ungulate interactions.

For purposes of the present assessment, however, it may be instructive to compare the trends of Isle Royale moose and wolf populations (shown below) with the theoretical model of Fig. 12.5. That model predicts an equilibrium would be reached after about 20 years, starting with a large moose population and a few wolves. On Isle Royale there has been little indication of an equilibrium state as yet. However, parvovirus was reported in the wolf population in about 1980 (Dr. R. Peterson, personal communication) and apparently drove it to low levels which have persisted until quite recently. The moose population then increased steadily and ultimately crashed in a very severe winter. If the "predator pit" models are correct, one might expect the moose population to continue to be suppressed by wolves. It will be interesting, and hopefully instructive, to see how the trend of Isle Royale moose and wolf populations compare with those of Yellowstone elk and wolves.



Trend of moose and wolf populations on Isle Royale.

LITERATURE CITED

1

- Abegglen, C. E., A. Y. Roppel, and F. Wilke. 1958. Alaska fur seal investigations, k Pribilof Island, Alaska. Report of U. S. Fish and Wildlife Service Bureau Commercial Fisheries, Seattle WA 187 pp.
- Allee, W. C., A. E. Emerson, O. Park, T. Park, and K. P. Schmidt. 1949. Principles of Animal Ecology. W. B. Saunders Co., Philadelphia and London.
- Alvarez-Bulla, E. R. and M. Slatkin. 1991. Finding confidence limits on population growth rates. *Trends in Ecology and Evolution*: 6:221-224.
- Alvarez-Bulla, E. R. and M. Slatkin. 1993. Finding confidence limits on population growth rates: Monte Carlo test of a simple analytic method. *Oikos* 68:273-282.
- Anderson, D. R. and R. S. Posphala. 1970. Correction of bias in belt transect studies of immotile objects. *Journal of Wildlife Management* 34:141-146,
- Anderson, D. R., Burnham, K. P. and B. R. Crain. 1978. A log-linear model approach to estimation of population size using the line-transect sampling method. *Ecology* 59:190-193.
- Anderson, D.R., A.P. Wywialowski, and K.P. Burnham. 1981. Tests of the assumptions underlying life table methods for estimating parameters from cohort data. *Ecology* 62:1122-1124.
- Anderson-Sprecher, R. 1994. Model comparisons and R^2 . *The American Statistician* 48:113-117.
- Andrewartha, H. G. and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, IL.
- Bailey, N. T. J. 1951. On estimating the size of mobile populations from recapture data. *Biometrika* 38:293-306
- Bailey, N. T. J. 1952. Improvements in the interpretation of recapture data. *Journal of Animal Ecology* 21:120-127.
- Ballard, W. B., J. S. Whitman, and C. L. Gardner. (1987). Ecology of an exploited wolf population in South-central Alaska. *Wildlife Monographs* No. 98.
- Bart, J. and D.S. Robson. 1982. Estimating survivorship when the subjects are visited periodically. *Ecology* 63:1078-1090.
- Berger, J. 1986. Wild horses of the Great Basin. Univ. Chicago Press, IL 326 pp.
- Berkson, J. 1957. Tables for the maximum likelihood estimate of the logistic function. *Biometrics* 13:28-34.
- Beverton, R. J. H. and S. J. Holt. On the dynamics of exploited fish populations. Fisheries Investigations, Ministry of Agriculture, Fisheries and Food (G.B.), Series II Salmon and Freshwater fish. 19:533pp.

- Birch, L. C. 1948. The intrinsic rate of natural increase of an insect population. *Journal of Animal Ecology* 17:15-26.
- Bonner, W. N. 1975. Population increase of gray seals at the Farne Islands. *Rapports P.-v. Conseil int. Explor. Mer.* 169:366-370.
- Boyd, J.M. 1974. Introduction. pp.1-7 *in* *Island Survivors: the Ecology of the Soay Sheep of St. Kilda*. P. A. Jewell, C. Milner, and J. M. Boyd, *Editors*. Athlone Press, University of London, London, U, K.
- Boyce, M. S. 1992. Population viability analysis. *Annual Reviews Ecology and Systematics* 23:481-506.
- Boyd, J. M. (1974) Introduction. *in* *Island Survivors: The Ecology of the Soay Sheep of St. Kilda*. (Eds. P. A. Jewell, C. Milner, and J. M. Boyd) pp. 1-7. The Athlone Press, University of London, London, England
- Breiwick, J. M. 1994. Rate of increase of the California gray whale. Document SC/46/AS18 submitted to Scientific Committee of the International Whaling Commission at the 46th meeting, Puerto Vallarta, Mexico, 6 pp.
- Breiwick, J.M., L.L.Eberhardt, and H.W.Braham. 1984. Population dynamics of Western Arctic bowhead whales. *Canadian Journal of Fisheries and Aquatic Sciences*. 41:484-496.
- Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson. 1978. Statistical inference from band recovery data: a handbook. U.S. Department of the Interior, Fish and Wildlife Service Resource Publication 131. 212pp.
- Brownie, C., J.E. Hines, and J.D.Nichols. 1986. Constant-parameter capture-recapture models. *Biometrics* 42:561-574.
- Buckland, S. T., J. M. Breiwick, K. L. Cattanach, and J. L. Laake. 1993. Estimated population size of the California gray whale. *Marine Mammal Science* 9:235-249.
- Buckland, S. T., K. P. Burnham, and N. H. Augustin. 1997. Model selection: an integral part of inference. *Biometrics* 53:603-618.
- Bulmer, M G. 1975. The statistical analysis of density dependence. *Biometrics* 31:901-911.
- Burnham, K. P. and D. R. Anderson. 1976. Mathematical models for nonparametric inferences from line transect data. *Biometrics* 32:325-326.
- Burnham, K.P. and D.R. Anderson. 1979. The composite dynamic method as evidence for age-specific waterfowl mortality. *J. Wildl. Manage.* 43:356-366.
- Burnham, K.P., D.R. Anderson, and J.L. Laake. 1980. Estimating density from line transect sampling of biological populations. *Wildlife Monographs* 72:1-202.

- Burnham, K.P. and D.R. Anderson. 1996. Data-based selection of an appropriate biological model: the key to modern data analysis. pp. 16-30 in *Wildlife 2001: Populations*. D. R. McCullough and R. H. Barrett, Eds. Elsevier Applied Science, London and New York.
- Burnham, K. P. and D. R. Anderson. (1998) *Model selection and inference*. Springer-Verlag, New York.
- Carbyn, L. N., S. M. Oosenburg, and D. W. Anions. 1993. Wolves, bison, and the dynamics related to the Peace-Athabaska delta in Canada's Wood Buffalo National Park. *Circumpolar Research Series No. 4*, Canadian Circumpolar Institute, University of Alberta, Edmonton, Alberta
- Caswell, H. 1989. *Matrix population models*. Sinauer Assoc., Sunderland, Mass.
- Caswell, H. 2001. *Matrix population models*. Second Edition. Sinauer Assoc., Sunderland, Mass
- Carbyn, L. N., S. M. Oosenburg, and D. W. Anions. (1993) *Wolves, bison, and the dynamics related to the Peace-Athabaska Delta in Canada's Wood Buffalo National Park*. *Circumpolar Research Series No. 4*, Canadian Circumpolar Institute, University of Alberta, Edmonton, Alberta, Canada.
- Carney, S. M. and G. A. Petrides. 1957. Analysis of variation among pheasant cock-crowing participants. *Journal of Wildlife Management* 21:393
- Caughley, G. 1974. Bias in aerial survey. *Journal of Wildlife Management* 38:921-923.
- Caughley, G. 1976. Wildlife management and the dynamics of ungulate populations. Pp. 183-246 in *Applied Biology, Vol I*, T.H. Croaker, Ed. Academic Press, London
- Caughley, G. 1977. *Analysis of vertebrate populations*. J. Wiley, London.
- Caughley, G. 1994. Directions in conservation biology. *Journal of Animal Ecology* 63:215-244.
- Caughley, G. and J.H. Lawton. 1981. Plant-herbivore systems. pp. 132-166 in *Theoretical Ecology*, R.M. May, Ed. (Second Edition) Sinauer, Sunderland Mass.
- Caughley, G., R. Sinclair, and D. Scott-Kemmis. 1976. Experiments in aerial survey. *Journal of Wildlife Management* 40:290-300.
- Chapman, D. G. 1948 . A mathematical study of confidence limits of salmon populations calculated from sample tag ratios. *International Pacific Salmon Fisheries Commission. Bulletin* 2:69-85.
- Chapman, D. G. 1952. Inverse multiple and sequential sample censuses. *Biometrics* 8:286-306.
- Chapman, D. G. 1954. The estimation of biological populations. *Annals Mathematical Statistics* 25:1-15.

- Chapman, D.G. and D.S. Robson. 1960. The analysis of a catch curve. *Biometrics* 16:354-368.
- Chapman, D. G. and A. M. Johnson. 1968. Estimation of fur seal pup populations by randomized sampling. *Transactions American Fisheries Society* 97:264-270.
- Christian, J. J. and D. E. Davis. 1964. Endocrines, behavior, and population. *Science* 146:1550-1560.
- Cleveland, W. S. 1979. Robust locally-weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.* 74:829-836
- Clutton-Brock, T. H., F. E. Guinness, and S. D. Albon. 1982. *Bred deer: behavior and ecology of two sexes.* University of Chicago Press, Chicago, Illinois.
- Clutton-Brock, T. H. , O. F. Price, S. D. Albon, and P. A. Jewell. (1991). Persistent instability and population regulation in Soay sheep. *Journal of Animal Ecology* 60:593-608.
- Cochran, W.G. 1977. *Sampling techniques.* (Third Edition) J.Wiley and Sons, New York.
- Cochran, W. G. 1983. *Planning and Analysis of Observational Studies.* John Wiley and Sons.
- Cole, L.C. 1954. The population consequences of life history phenomena. *Quarterly Review of Biology* 29:103-137.
- Cole, G. C. 1983. A naturally regulated elk population. Pages 62-81, in F. L. Bunnell, D. S. Eastman, and J. M. Peek, eds., *Symp. on natural regulation of wildlife populations.* Univ. Idaho, Moscow, Idaho.
- Cormack, R. M. 1966. A test for equal catchability. *Biometrics* 22:330-342.
- Cormack, R. M. 1968. The statistics of capture-recapture methods. *Oceanographic Marine Biology Annual Review* 6:455-506.
- Cox, D. R. 1969. Some sampling problems in technology. In: N. L. Johnson and H. Smith, Jr. (Eds.) *New developments in Survey Sampling.* 506-507. Wiley-Interscience, Wiley, New York.
- Cox, D. R. 1972. Regression model and life tables. *Jour. Royal Stat. Soc.* B34,187-220.
- Craighead, J. J., F. C. Craighead, Jr., R. L. Ruff, and B. W. O'Gara. 1973. Home ranges and activity patterns of nonmigratory elk of the Madison drainage herd as determined by biotelemetry. *Wildl. Mono. No. 33.*
- Craighead, J.J., J.R. Varney, and F. C. Craighead, Jr. 1974. A population analysis of the Yellowstone grizzly bears. *Bull. 40, Montana Forest and Conservation Experiment Station, Univ. of Montana, Missoula MT*

- Crain, B. R., K. P. Burnham, D. R. Anderson, and J. L. Laake. 1978. A Fourier Series Estimator of Population Density for Line Transect Sampling. Utah State University Press.
- Dale, B.W., L.G. Adams, and R.T. Bowyer. 1994. Functional response of wolves preying on barren-ground caribou. *J. Anim. Ecol.* 63:644-652.
- Darroch, J. N. 1958. The multiple-recapture census, I. Estimation of a closed population. *Biometrika* 45:343-359.
- Darroch, J. N. 1959. The multiple-recapture census, II. Estimation when there is immigration or death. *Biometrika* 46:336-351.
- David, J. F. 1995. Note on the use of the basic equation of demography. *Oikos* 73:285-288
- Deevey, E.S., Jr. 1947. Life tables for natural populations of animals. *Quarterly Review of Biology* 22:283-314.
- DelGiudice, G. D., R. O. Peterson, and W. M. Samuel. 1997. Trends of winter nutritional restriction, ticks, and numbers of moose on Isle Royale. *J. Wildl. Manage.* 61:895-903.
- DeLury, D. B. 1947. On the estimation of biological populations. *Biometrics* 3:145-167.
- DeLury, D. B. 1951. On the planning of experiments for the estimation of fish populations. *Journal Fisheries Research Board of Canada.* 8:281-307.
- DeMaster, D.B., M.C.S. Kinglsey, and I. Stirling. 1980. A multiple mark and recapture estimate applied to polar bears. *Canadian Journal of Zoology* 58:633-638.
- DeMaster, D.B., D. J. Miller, D. Goodman, R. L. DeLong, and B. S. Stewart. 1982. Assessment of California sea lion fishery interactions. *Transactions 47th North American Wildlife Conference*, pp. 253-264
- Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. *Ecological Monographs* 61:115-143.
- Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. *Ecological Monographs* 64:205-224.
- Despain, D. G. 1990. *Yellowstone vegetation*. Roberts Rhinehart Publishers, Boulder, Colorado.
- Despain, D. G., A. Rodman, P. Schullery and H. Shovic. 1989. Burned area survey of Yellowstone National Park: the fires of 1988. Unpublished report, Division of Research and Geographic Information Systems Laboratory, Yellowstone National Park, Wyoming.

- Devries, P. G. 1974. Multistage line intersect sampling. *Forest Science* 20:129-133,
- Dice, L. R. 1941. Methods for estimating populations of mammals. *Journal of Wildlife Management* 5:398-407.
- Doak, D. F. and L. S. Mills. 1994. A useful role for theory in conservation. *Ecology* 75:615-626.
- Doi, T. Further development of whale sighting theory. In W. E. Schevill (Ed.) *The Whale Problem: A Status Report*:359-368. Harvard University Press, Cambridge Mass.
- Draper, N. R. and H. Smith. 1998. *Applied regression analysis*. Third Edition. J. Wiley and Sons, New York.
- Edwards, W. R. and L. L. Eberhardt. 1967. Estimating cottontail abundance from live-trapping data. *Journal of Wildlife Management* 31:87-96.
- Eberhardt, L. E., L. L. Eberhardt, B. L. Tiller, and L. L. Cadwell. 1996. Growth of an isolated elk population. *Journal of Wildlife Management* 60:369-373.
- Eberhardt, L. L. 1960. Estimation of Vital Characteristics of Michigan Deer Herds. Ph.D. Thesis. Michigan State University. Game Division Report. No. 2282, Michigan Dept. of Conservation, Lansing, MI. 192 pp.
- Eberhardt, L. L. 1968. A preliminary appraisal of line transects. *Journal of Wildlife Management* 32:82-88.
- Eberhardt, L. L. 1969. Population estimates from recapture frequencies. *Journal of Wildlife Management* 33:28-39.
- Eberhardt, L.L. 1970. Correlation, regression and density dependence. *Ecology* 51(2):306-310.
- Eberhardt, L.L. 1972. Some problems in estimating survival from banding data. pp.153-171 in: *Population Ecology of Migratory Birds: A Symposium*. U.S. Department of Interior, Wildlife Research Report 2. 278 pp.
- Eberhardt, L.L 1977a. "Optimal" management policies for marine mammals. *Wildlife Society Bulletin* 5:162-169.
- Eberhardt, L.L. 1977b. Optimum policies for conservation of large mammals with special reference to marine ecosystems. *Environmental Conservation* 4:205-212.
- Eberhardt, L.L. 1977c. Relationship between two stock-recruitment curves. *Journal of the Fisheries Research Board of Canada*. 34:425-428.
- Eberhardt, L.L. 1978a. Appraising variability in population studies. *Journal of Wildlife Management* 42:207-238.

- Eberhardt, L. L. 1978b. Transect methods for population studies. *Journal of Wildlife Management* 42:1-31.
- Eberhardt, L. L. 1979. Line-transects based on right-angle distances. *Journal of Wildlife Management* 43:768-774.
- Eberhardt, L. L. 1981. Population dynamics of the Pribilof fur seals. Pages 197-220 in C. W. Fowler and T. D. Smith, (Eds.) *Dynamics of large mammal populations*. J. Wiley and Sons, N.Y.
- Eberhardt, L.L. 1985. Assessing the dynamics of wild populations. *Journal of Wildlife Management* 49:997-1012.
- Eberhardt, L.L. 1987. Population projections from simple models. *Journal of Applied Ecology* 24:103-118.
- Eberhardt, L.L. 1988. Using age structure data from changing populations. *J. Appl. Ecol.* 25:373-378.
- Eberhardt, L.L. 1988. Testing hypotheses about populations. *J. Wildl. Manage.* 52:50-56.
- Eberhardt, L. L. 1992. An analysis of procedures for implementing the dynamic response method. *Marine Mammal Science* 8:201-212.
- Eberhardt, L. L. 1995. Using the Lotka-Leslie model for sea otters. *Journal of Wildlife Management* 59:222-227.
- Eberhardt, L. L. (1995b) Population trend estimates from reproductive and survival data. *Yellowstone grizzly bear investigations for 1994*. (Eds. R. R.
- Eberhardt, L. L. 1997. Is wolf predation ratio-dependent? *Canadian Journal of Zoology* 75:1940-1944.
- Eberhardt, L. L. 1998. Applying difference equations to wolf predation. *Canadian Journal of Zoology* 76:380-386.
- Eberhardt, L.L. 2000. Reply: Predator=prey ratio dependence and regulation of moose populations. *Canadian Journal of Zoology* 78
- Eberhardt, L. L. and R. C. Van Etten. 1956. Evaluation of the pellet group count as a deer census method. *J. Wildl. Manage.* 20(1):70-74.
- Eberhardt, L. L. and W.H. Rickard, 1963. Pacific Northwest Laboratory Internal Report.
- Eberhardt, L.L. and D.B. Siniff. 1977. Population dynamics and marine mammal management policies. *Journal Fisheries Research Board of Canada* 34:183-190.
- Eberhardt, L.L., A.K. Majorowicz, and J.A. Wilcox. (1982) Apparent rates of increase for two feral horse herds. *Journal of Wildlife Management* 46:367-374.
- Eberhardt, L.L. and M.A. Simmons. 1987. Calibrating population indices by double sampling. *J. Wildl. Manage.* 51:665-675.

- Eberhardt, L.L. and K.W. Pitcher. 1992. A further analysis of the Nelchina caribou and wolf data. *Wildl. Soc. Bull.* 20:385-395.
- Eberhardt, L.L. and M.A. Simmons. 1992. Assessing rates of increase from trend data. *J. Wildl. Manage.* 56:603-610.
- Eberhardt, L. L. and T. J. O'Shea. (1995) Integrating manatee life history data and population model development. *Population biology of the Florida manatee*. (eds T.J. O'Shea, B.B. Ackerman and H. F. Percival), Information and Technology Report 1, National Biological Service Tech. Report Series, Ft. Collins, CO
- Eberhardt, L. L. and R. R. Knight. 1996. How many grizzlies in Yellowstone? *Journal of Wildlife Management* 60:416-421.
- Eberhardt, L. L. and R. O. Peterson. (1999) Predicting the wolf-prey equilibrium point. *Canadian Journal of Zoology* 77:494-498.
- Eberhardt, L. L., T. J. Peterle, and R. Schofield. 1963. Problems in a rabbit population study. *Wildlife Monographs* 10:1-51.
- Eberhardt, L.L., D.G. Chapman, and J.R.Gilbert. 1979. A review of marine mammal census methods. *Wildlife Monographs* 63:1-46.
- Eberhardt, L.L., A.K. Majorowicz, and J.A. Wilcox. 1982. Apparent rates of increase for two feral horse herds. *Journal of Wildlife Management* 46:367-374.
- Eberhardt, L. L., B.M. Blanchard, and R.R. Knight. 1994. Population trend of the Yellowstone grizzly bear as estimated from reproductive and survival rates. *Canadian Jour. of Zoology* 72:360-363.
- Eberhardt, L. L., R. A. Garrott, P. J. White, and P. J. Gogan. (1998). Alternative approaches to aerial censusing of elk. *Journal of Wildlife Management* 62;1046-1055.
- Eberhardt, L. L., R. A. Garrott, and B. L. Becker. (1999) Using trend indices for endangered species. *Marine Mammal Science* 15:766-785.
- Emlen, J. T. 1971. Population densities of birds derived from transect counts. *The Auk* 88:323-342.
- Efron, B. and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- Ehrlich, P. R. and L. C. Birch. 1967. The "balance of nature" and "population control". *American Naturalist* 101:97-107.
- Fay, F. H., L. L. Eberhardt, B. P. Kelly, J. J. Burns, and L. T. Quackenbush. Status of the Pacific walrus populaion, 1950-1989. *Marine Mammal Science* 13:537-565.
- Finney, D. J. 1978. *Statistical Method in Biological Assay*. 3rd Edition. Charles Griffin and Co. Ltd., London.

- Fisher, R. A. and E. B. Ford. 1947. The spread of a gene in natural conditions in a colony of the moth, *Panaxia dominula* L. *Heridity*, London 1:143-174.
- Fowler, C.W. 1981. Density dependence as related to life history strategies. *Ecology* 62:602-610.
- Fowler, C.W. 1994. Further consideration of nonlinearity in density dependence in large mammals. *Reports of the International Whaling Commission* 44:385-391.
- Fowler, C. W. and D. B. Siniff. 1992. Determining population status and the use of biological indices in the management of marine mammals. pp 1025-1037 in *Wildlife 2001: Populations*. D. R. McCullough and R. H. Barrett, Eds. Elsevier Applied Science, London and New York.
- Fredin, R. A. 1984. Levels of maximum net productivity in populations of large terrestrial mammals. *Reports International Whaling Commission, Special Issue* 6.
- Fryxell, J. M., W. E. Mercer, and R. B. Gellately. 1988. Population dynamics of Newfoundland moose using cohort analysis. *Journal of Wildlife Management* 52:14-21.
- Fuller, T.K. 1989. Population dynamics of wolves in north-central Minnesota. *Wildl. Mono. No. 105*, 41pp.
- Gaillard, J-M, M Festa-Blanchet and N. G. Yoccoz. 1998. Population dynamics of large herbivores: variable recruitment with constant adult survival. *Trends in Ecology and Evolution*: 13:58-63.
- Gaillard, J-M, M Festa-Blanchet, N. G. Yoccoz, A. Loison, and C. Toigo. 2000. Temporal variation in fitness components and population dynamics of large herbivores. *Annual Reviews of Ecology and Systematics*. 31:3617-393.
- Garrott, R.A. and L. Taylor. 1990. Dynamics of a feral horse population in Montana. *J. Wildl. Manage.* 54:603-612.
- Garrott, R. A., B.B. Ackerman, J. R. Cary, D. M. Heisey, J. R. Reynolds, III, and J. R. Wilcox. 1994. Trends in counts of Florida manatees at winter aggregation sites. *J. Wildl. Manage.* 58:642-654.
- Garrott, R. A., B.B. Ackerman, J. R. Cary, D. M. Heisey, J. R. Reynolds, III, and J. R. Wilcox 1995. Assessment of trends in sites of manatee populations at several Florida aggregation sites. pp34-55 in *Population Biology of the Florida Manatee*, T. J. O'Shea, B. B. Ackerman, and H. F. Percival (Eds.) *Information and Technology Report 1*, National Biological Service, USDI, Washington, D. C.
- Gasaway, W. C., R. D. Boertje, D. V. Grandgaard, D. G. Kelleyhouse, R. O. Stephenson, and D. G. Larsen. 1992. The role of predation in limiting moose at low densities in Alaska and Yukon and implications for conservation. *Wildlife Monographs No. 120*.

- Gates, C. E. 1969. Simulation study of estimators for the line transect sampling method. *Biometrics* 25:317-328.
- Gates, C. E., Marshall, W. H. and D. P. Olson. 1968. Line transect method of estimating grouse population densities. *Biometrics* 24:135-145.
- Gerrodette, T. 1987. A power analysis for detecting trends. *Ecology* 68:1364-1372
- Gerrodette, T. 1991. Models for power of detecting trends -- a reply to Link and Hatfield. *Ecology* 72:1889-1892.
- Gilbert, J. R. and M. S. Udevitz. 1997. Adaptation of a stage-projection model for species with multiple year reproductive cycles. *Ecological Modelling* 97:47-57.
- Gilmartin, W. G. and L. L. Eberhardt. 1995. Assessing status of the Hawaiian monk seal population. *Canadian Journal of Zoology* 73:1185-1190.
- Gilmartin, W.G., T.C. Johanos, and L. L. Eberhardt. 1993. Survival rates for the Hawaiian monk seal (*Monachus schauinslandi*) *Marine Mammal Science* 9:407-420.
- Greig-Smith, P. 1964. *Quantitative plant ecology*. (2nd 2d.) Butterworths: London.
- Grosenbaugh, L. R. 1965. Generalization and reparameterization of some sigmoid and other nonlinear functions. *Biometrics* 21:708-714.
- Grubb, P. 1974. Population dynamics of the Soay sheep. Pp. 242-272 in *Island Survivors, the ecology of the Soay sheep of St. Kilda*, P. A. Jewell, C. Milner, and J. M. Boyd, Athlone Press, Univ. of London, U.K.
- Hairston, N. G., F. E. Smith, and L. B. Slobodkin. 1960. Community structure, population control, and competition. *American Naturalist* 94:421-425.
- Haldane, J. B. S. 1955. The calculation of mortality rates from ringing data. *Proceedings Xith International Ornithological Congress Basel* 454-458.
- Hanson, W. C. and L. L. Eberhardt. 1971. A Columbia River goose population, 1950-1970. *Wildlife Monographs* No. 28, 61 pp.
- Harris, R. B. 1986. Reliability of trend lines obtained from variable counts. *Journal of Wildlife Management* 50:165-171.
- Hayne, D. W. 1949. An examination of the strip census method for estimating animal abundance. *Journal of Wildlife Management* 13:145-157.
- Heincke, F. 1913. Investigations on the plaice. General Report. 1. The plaice fishery and protective measures. Preliminary brief summary of the most important points of the report. *Conseil International Exploration de Mer. Rapports et Proces-verbaux*. 16:1-67.
- Heisy, D.M. and T.K. Fuller. 1985. Evaluation of survival and cause-specific mortality rates using telemetry data. *Journal of Wildlife Management* 49:668-674.

- Hemingway, P. 1971. Field trials of the line transect method of sampling large populations of herbivores. In: E. Duffy and A. S. Watts. (Eds.) *The Scientific Management of Animal and Plant Communities for Conservation*. 405-411. Blackwell, Oxford.
- Hesselton, W. T., C. W. Severinghaus, and J. E. Tanck. 1965. *New York Fish and Game Journal* 12:17-30s
- Hickey, F. 1960. Death and reproductive rates of sheep in relation to flock culling and selection. *N.S. J. Agric. Res.* 3:332-344.
- Houston, D.B. 1982. *The northern Yellowstone elk herd*. Macmillan Publishing Co., New York.
- Hovey, F. W. and B. N. McLellan. 1996. Estimating population growth of grizzly bears from the Flathead River drainage using computer simulations of reproduction and survival rates. *Canadian Journal of Zoology* 74:1409-1416.
- Ito, Y. 1972. On the methods for determining density-dependence by means of regression. *Oecologia* 10:347-372.
- Jackson, C. H. N. 1933. On the true density of tsetse flies. *Journal of Animal Ecology* 2:204-209.
- Jackson, C. H. N. 1937. Some new methods in the study of *Glossina morsitans*. *Proceedings Zoological Society of London* 1936:811-896.
- Jackson, C. H. N. 1939. The analysis of an animal population. *Journal of Animal Ecology* 8:238-246.
- Jackson, C. H. N. 1940. The analysis of a tsetse-fly population I. *Annals Eugenics London*. 10:322-369.
- Jackson, C. H. N. 1948. The analysis of a tsetse-fly population III. *Annals Eugenics London*. 14:91-108
- Johanos, T.C., B.L. Becker, and T.J. Ragen. 1994. Annual reproductive cycle of the female Hawaiian monk seal (*Monachus schauinslandi*) *Marine Mammal Science* 10:13-30.
- Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration – stochastic model. *Biometrika* 52:225-247.
- Junge, C. O. A quantitative evaluation of the bias in population estimates based on selective samples. In: *North Atlantic Fish Marking Symposium*, International Council North Atlantic Fisheries Special Publication NO. 4, pp. 26-28.
- Kaplan, E. L. and P. Meier. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 53:457-481.

- Keith, L. B. 1983. Population dynamics of wolves. Pages 66-77 in L. N. Carbyn, ed. Wolves in Canada and Alaska. Can. Wildl. Serv. Rep. Ser. 45.
- Kenyon, K. W., V. B. Scheffer, and D. G. Chapman. 1954. A population study of the Alaska fur seal herd. Special Scientific Report N. 12, U. S. Department of the Interior, U. S. Fish and Wildlife Service, Washington D. C.
- Keyfitz, N. 1968. Introduction to the mathematics of population. Addison-Wesley Publishing Co., Reading, Mass.
- Knight, R.R. and L.L. Eberhardt. 1985. Population dynamics of the Yellowstone grizzly bears. Ecology 66:323-334.
- Knight, R. R., B. M. Blanchard, and L. L. Eberhardt. 1995. Appraising status of the Yellowstone grizzly bear population by counting females with cubs-of-the-year. Wildlife Society Bulletin 23:245-248.
- Krebs, C. J. 1964. The lemming cycle at Baker Lake, Northwest Territories, during 1959-62. Arctic Institute of North America. Technical Paper No. 15.
- LeBouef, B. J. and M. L. Bonnell. 1980. Pinnipeds of the California Islands: abundance and distribution. pp. 475-493 in The California Islands (D. M. Power, ED.). Santa Barbara Museum of Natural History, Santa Barbara, California
- LeBouef, B. J. 1981. The elephant seal. pp. 291-301 in Problems in management of locally abundant wild mammals. P. A. Jewell and S. Holt, Editors, Academic Press, New York.
- Lebreton, J., K.P. Burnham, J. Clobert, and D.R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecol. Mono. 62:67-118.
- Leopold, A. 1933. Game Management. C. Scribner
- Leslie, P.H. 1945. On the use of matrices in certain population mathematics. Biometrika 33:183-212.
- Leslie, P.H. 1948. Some further notes on the use of matrices in certain population mathematics. Biometrika 33:183-212.
- Leslie, P. H. 1952. The estimation of population parameters from data obtained by means of the capture-recapture method. II: the estimation of total numbers. Biometrika 39:363-388.
- Leslie, P. H. 1958. Statistical Appendix. Journal of Animal Ecology. 27:84-86.
- Leslie, P. H. 1966. The intrinsic rate of increase and the overlap of successive generations in a population of guillemots (*Uria aalge* Pont.). Journal of Animal Ecology 35:291-301.
- Leslie, P. H. and D. H. S. Davis. 1939. An attempt to determine the absolute number of rats on a given area. Journal of Animal Ecology 8:94-113.

Leslie, P.H. and R.M. Ranson 1940. The mortality, fertility and rate of natural increase of the vole (*Microtus agrestis*) as observed in the laboratory. *The Journal of Animal Ecology*, 9:27-52.

Leslie, P. H. and D. Chitty. 1951. The estimation of population parameters from data obtained by means of the capture-recapture method I. The maximum likelihood equations for estimating the death rate. *Biometrika* 38:269-292.

Leslie, P. H., D. Chitty, and H. Chitty. 1953. The estimation of population parameters from data obtained by means of the capture-recapture method. III. An example of the practical applications of the method. *Biometrika* 40:137-169.

Link, W. A. and J. S. Hatfield. 1990. Power calculations and model selection for trend analysis: a comment. *Ecology* 71:1217-1220.

Link, W. A. and B. A. Hoover. 1991. The detection of density-dependence from a series of annual censuses-- a comment. *Ecology* 72:1514-1515.

Link, W. A. and J. R. Sauer. 1997. Estimation of population trajectories from count data. *Biometrics* 53:488-497.

Lorenzen, G. 1990. A unified approach to the calculation of growth rates. *The American Statistician* 44:148-150.

Lotka, A.J. 1907. Relation between birth rates and death rates. *Science, N.S.*, 26: 21-22.

Lotka, A.J. 1925. *Elements of Physical Biology*, Williams and Wilkins, Baltimore MD

Lotka, A.J. 1939. *Theorie analytique des associations biologiques. Deuxieme partie: Analyse demographique avec application particuliere a l-espece humaine. Actualites Scientifiques et industrielles.*, 780. Hemann et Cie. Paris., pp1-149.

Lowe, V. P. W. 1969. Population dynamics of the red deer (*Cervus elaphus* L.) on Rhum. *Journal of Animal Ecology* 38:425-457.

Lucas, H. A. and G. A. F. Seber, 1977. Estimating coverage and particle density using the line-intersect method. *Biometrika* 64:618-622.

Mangel, M. and C. Tier. 1994. Four facts every conservation biologist should know about persistence. *Ecology* 75:607-614.

Manly, B. J. F. and M. J. Parr. 1968. A new method of estimating population size, survivorship and birth rate from capture-recapture data. *Transactions Society British Entomology* 18:81-89.

Matson, P. and A. Berryman (Editors. 1992. Ratio-dependent predator-prey theory. *Ecology* 73:1529-1566.

May, R. M. (Editor) 1981. *Theoretical Ecology. (Second Edition)* Sinauer, Sunderland Mass.

- May, R.M., J.R. Beddington, C.W. Clark, S.J. Holt, and R.M. Laws. 1979. Management of multispecies fisheries. *Science* 205:267-277.
- McCaffrey, K. R. 1976. Deer trail counts as an index to populations and habitat use. *Journal of Wildlife Management* 40:308-328.
- McCullough, D. C., 1981. Population dynamics of the Yellowstone grizzly bear. PP.173-196 in C. W. Fowler and T. D. Smith (Eds) *Dynamics of Large Mammal Populations*, John Wiley and Sons., New York
- McCullough, D. C. 1983. Rate of increase of white-tailed deer on the George Reserve: a response. *Journal of Wildlife Management* 47:1248-1250.
- McLaren, L. a. 1961. Methods of determining the numbers and availability of ringed seals. *Arctic* 14:162-175.
- McRoberts, R. e., L. D. Mech, and R. O. Peterson. 1995. The cumulative effect of consecutive winters' snow depth on moose and deer populations: a defence. *Journal of Animal Ecology* 64:131-135.
- Mech, L. D., R. E. McRoberts, R. O. Peterson, and R. E. Page. 1987. Relationship of deer and moose populations to previous winters' snow. *Journal of Animal Ecology* 56:615-627.
- Mech, L. D., L. G. Adams, T. J. Meier, J. W. Burch, and B. W. Dale. (1998) *The wolves of Denali*. University of Minnesota Press, Minneapolis, MN
- Messier, F. 1991. The significance of limiting and regulating factors on the demography of moose and white-tailed deer. *Journal of Animal Ecology* 60:377-393.
- Messier, F. 1994. Ungulate population models with predation: a case study with the North American moose. *Ecology* 75:478-488.
- Messier, F. and M. Crete. 1985. Moose-wolf dynamics and the natural regulation of moose populations. *Oecologia* 65:503-512.
- Meyer, J. S., C. G. Ingersoll, L. L. McDonald, and M. S. Boyce. 1986. Estimating uncertainty in population growth rates: jackknife vs bootstrap techniques. *Ecology* 67:1156-1166.
- Moran, P. A. P. 1951. A mathematical theory of animal trapping. *Biometrika* 38:307-311.
- Mizroch, S. A. 1981. Analyses of some biological parameters of the Antarctic fin whale (*Balenopter physalus*) Rep. Int. Whaling Comm. 31:425-434.
- Murdoch, W. W. 1966. "Community structure, population control, and competition"-- a critique. *American Naturalist* 100:219-226.
- Murdoch, W. W. 1994. Population regulation in theory and practice. 1994. *Ecology* 75:271-287.

- Murray, B.G. and L. Garding. 1984. On the meaning of parameter x in Lotka's discrete equations. *Oikos* 42:323-326.
- National Academy of Science. 1974. Report of Committee on the Yellowstone grizzlies, Washington D.C.
- Nixon, C. M., W. R. Edwards, and L. L. Eberhardt. 1967. Estimating squirrel abundance from live-trapping data. *Journal of Wildlife Management* 31:96-101.
- Norton-Griffiths, M 1975. Counting Animals. African Wildlife Leadership Foundation, Nairobi.
- Overton, W. s. and D. E. Davis. 1969. Estimating the numbers of animals in wildlife populations. In: R. H. Giles, Jr. (Ed.) *Wildlife Management Techniques*, 3rd Ed. Pp. 403-455. The Wildlife Society, Washington D. C.
- Ozaga, J. J. and L. J. Verme. 1982. Physical and reproductive characteristics of a supplementally-fed white-tailed deer herd. *Journal of Wildlife Management* 46:281-301
- Parzen, S. 1972. *Stochastic Processes*. Holden-Day, Inc. San Francisco.
- Patil, G. P. and C. R. Rao. 1978. Weighted distributions and size biased sampling with application to wildlife populations and human families. *Biometrics* 34:179-189.
- Paulik, G.J. 1962. Use of the Chapman-Robson survival estimate for single- and multiple-release tagging experiments. *Transactions American Fisheries Society* 91:95-98.
- Paulik, G.J. 1963, Estimates of mortality rates from tag recoveries. *Biometrics* 19:28-57.
- Pearl, R. 1926. *The biology of population growth*. A. A, Knopf, New York, N.Y.
- Pearson, O. P. 1966. The prey of carnivores during once cycle of mouse abundance. *Journal of Animal Ecology* 35:217-233
- Peterson, R. O. 1995. *The wolves of Isle Royale: a broken balance*. Willow Creek Press, Minocqua, WI.
- Peterson, R. O., J. D. Woolington, and T. N. Bailey. (1984) *Wolves of the Kenai Peninsula, Alaska. Wildlife Monographs* No. 88.
- Pollard, E., K. H. Lakahani, and P. Rothery. 1987. The detection of density-dependence from a series of annual censuses. *Ecology* 68:2046-2055.
- Pollock, K. H. 1978. A family of density estimators for line transect sampling. *Biometrics* 34:475-478.
- Pollock, K.H. and D.G. Raveling. 1982. Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. *Journal of Wildlife Management* 46:88-98.

Pollock, K.H.; Winterstein, S.R.; Bunck, C.M.; Curtis, P.D. (1989): Survival analysis in telemetry studies: the staggered entry design. *Journal of Wildlife Management* 53, 7-15.

Pollock, K.H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. *Wildl. Mono. No. 107*, 97pp.

Post, E. and N. C. Stenseth. 1998. Large-scale climatic fluctuations and population dynamics of moose and white-tailed deer. *Journal of Animal Ecology* 67:537-543.

Quenouille, M. H. 1956. Notes on bias reduction. *Biometrika* 43:353-360

Reddingiusm J. and P. J. den Boer. 1989. On the stabilization of animal numbers. Problems of testing. 1 Power estimates and estimation errors. *Oecologia* 78:1-8.

Reddingiusm J. and P. J. den Boer. 1989. On the stabilization of animal numbers. Problems of testing. 2 Confrontation with data from the field. *Oecologia* 79:143-149.

Reid, V. H., R. M. Hansen, and A. L. Ward. 1966. Counting mounds and earth plugs to census mountain pocket gophers. *Journal of Wildlife Management* 30:327-334.

Ricker, W. E. 1958. Handbook of computations for biological statistics of fish populations. *Bulletin fisheries Research Board of Canada* 119:1-300.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. *Bulletin Fisheries Research Board of Canada*. 191:382pp.

Robinette, W. L., Loveless, C. M., and D. A. Jones. 1974. Field tests on strip census methods. *Journal of Wildlife Management* 38:81-96.

Robson, D. S. and D. G. Chapman. 1961, Catch curves and mortality rates. *Transactions American Fisheries Society* 90:181-189.

Ryel L. A. 1971. Evaluation of Pellet Group Survey for Estimating Deer Populations in Michigan. PhD Thesis, Michigan State University, East Lansing, Mich.

Saether, B-E. 1997. Environmental stochasticity and population dynamics of large herbivores: a search for mechanisms. *Trends in Ecology and Evolution*: 12:143-149..

Samuel, M. D., E. O. Garton, M. W. Schlegel, and R. G. Carson. 1987. Visibility bias during aerial surveys of elk in northcentral Idaho. *Journal of Wildlife Management* 51:622-639.

Schnabel, Z. E. 1938. The estimation of the total fish population of a lake. *American Mathematical Monthly* 45:348-352,

Scheffe, H. 1959. *The Analysis of Variance*. John Wiley and Sons.

Schnute, J. T. and L. J. Richards. 1990. A unified approach to the analysis of fish growth, maturity, and survivorship data. *Canadian Journal of Fisheries and aquatic science* 47:24-40.

Schumacher, F. X. and R. W. Eschmeyer 1943. The estimation of fish populations in lakes and ponds. *Journal of Tennessee Academy of Science* 18:228-249.

Seber, G.A.F. 1962. The multiple-sample single recapture census. *Biometrika* 49:33-349.

Seber, G.A.F. 1965. A note on the multiple recapture census. *Biometrika* 52:330-349.

Seber, G.A.F. 1982. *The estimation of animal abundance*.(Second edition) Macmillan Publishing Co. New York.

Shenk, T. M., G. C. White, and K. P. Burnham. 1998. Sampling variance effects on detecting density dependence from temporal trends in natural populations. *Ecology* 68:445-463.

Sibly, R. M. and R. H. Smith. 1998. Identifying key factors using λ contribution analysis. *Journal of Animal Ecology* 67:17-24.

Siler, W. 1979. A competing-risk model for animal mortality. *Ecology* 60:750-757.

Sinclair, A.R.E. 1977. *The African buffalo, a study of resource limitation of populations*. Univ. of Chicago Press, Chicago, IL 355 pp.

Siniff, D. B. and R. O. Skoog. 1964. Aerial censusing of caribou using stratified random sampling. *Journal of Wildlife Management* 88:391-401.

Siniff, D.B. and K. Ralls. 1991. Reproduction, survival and tag loss in California sea otters. *Marine Mammal Science* 7:211-229.

Skellam, J. G. 1955. The mathematical approach to population dynamics. Pp. 31-45 in: *The Numbers of Man and Animals*. J. B. Cragg and N. W. Pirie, Eds. Oliver and Boyd, London.

Skellam, J. G. 1958. The mathematical foundations underlying the use of line transects in animal ecology *Biometrics* 14:385-400.

Skogland, T. 1990. Density dependence in a fluctuating wild reindeer herd; maternal vs. offspring effects. *Oecologia* 84:442-450.

Slobodkin, L. B., F. E. Smith, and N. G. Hairston. 1967. Regulation in terrestrial ecosystems, and the implied balance of nature. *American Naturalist* 101:109-124.

Smith, R. H. and S. Gallizioli. 1965. Predicting hunter success by means of a spring call count. *Journal of Wildlife Management* 29:806-813,

Snedecor, G. W. AND W. G. Cochran. 1967. *Statistical Methods* (6th Ed.) Iowa State University Press, Ames, Iowa.

- Solomon, M. E. 1949. The natural control of animal populations. *Journal of Animal Ecology* 18:1-35.
- Solow, A. R. 1990. Testing for density dependence: a cautionary note. *Oecologia* 83:47-49.
- Soule, M. E. (Editor). 1987. *Viable Populations for Conservation*. Cambridge Univ. Press, Cambridge MA.
- Spencer, D. L. and C. J. Lensink. 1970. The muskox of Nunivak Island, Alaska. *Journal of Wildlife Management* 34:1-15
- Stewart, B. S., P.K. Yochem, H. R. Huber, R. L. DeLong, R. J. Jameson, W. J. Sydeman, S. G. Allen, and B. J. LeBouef. 1994. History and present status of the northern elephant seal population. pp 29-49 *in* *Elephant seals: population ecology, behavior and physiology*. B. J. Lebouef and R. M. Laws, Eds. University of California Press, Berkeley, California
- Tait, D.E.N. and F. L. Bunnell. 1980. Estimating rate of increase from ages at death. *J. Wildl. Manage.* 44:296-299.
- Thompson, S. K. 1992. *Sampling*. J. Wiley and Sons, Inc.
- Thompson, S. K. and G. A. F. Seber. 1996. *Adaptive Sampling*. John Wiley and Sons, Inc.
- Tukey, J. W. 1958. Bias and confidence in not quite large samples. *Annals Mathematical Statistics*. 29:614,
- Van Sickle, J., C.A.M. Atwell, and G.C. Craig. 1987. Estimating population growth rate from an age distribution of natural deaths. *J. Wildl. Manage.* 51:941-948.
- Vickery, W. L. and T. D. Nudds. 1984. Detection of density-dependent effects in annual duck censuses. *Ecology* 65:96-104.
- Volterra, V. 1931. *Lecons sur la Theorie Mathematique de la Lutte pour la vie*. Gauthier-Villiers, Paris.
- Watt, K. E. F. 1964. Density dependence in population fluctuations. *Canadian Entomologist* 96:1147-1148.
- Watt, K. E. F. 1968. *Ecology and resource management*. McGraw-Hill, New York.
- White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. *Journal of Wildlife Management* 47:716-728.
- Whitlock, S. C. and L. Eberhardt. 1956. Large-scale dead deer surveys: methods, results and management implications. *Trans. 21st North American wildlife Conference*: 555-566.
- Wynne-Edwards, V. C. 1962. *Animal dispersion in relation to social behavior*. Oliver and Boyd, Edinburgh, Scotland.

- Yapp, W. B. 1956. The theory of line transects. *Bird Study* 3:93-104,
- Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. *Biometrics* 12:163-169.
- Zippin, C. 1958. The removal method of population estimation. *Journal of Wildlife Management* 22:82-90.
- Zhou, Z. and W. Pan. 1997. Analysis of the viability of a giant panda population. *Journal of Applied Ecology* 34:363-374.

APPENDIX A BOOTSTRAPPING

This Appendix provides some programs to do bootstrapping and data-fitting. The bootstrap programs depend on EXCEL macros, but users do not need to know details of the programming. All one needs to do is insert the data in a spreadsheet and run the macro. This is accomplished by opening the “TOOLS” menu, and selecting “MACRO” from the submenu. A window labeled “Macro” comes up, and hitting the “run” button starts the program. A keystroke alternative is available on computers running Windows (in some machines it is ALT F8). This brings up the same “Macro” window, and is a little faster than going thru the TOOLS menu. When you open one of these programs, a window may come up indicating that the program contains macros, and warning about viruses. Hit “Enable Macros” to use the program. Some virus protection software will also detect the code used and bring up a warning window.

GENERAL PURPOSE PROGRAMS

There are 4 programs available to produce the results of Chapters 2 and 3. The first, BOOT1, bootstraps single columns of data. It computes the bootstrapping example of Section 2.1. Contents of the EXCEL window are as follows:

	A	B	C	D	E
1	NOBS	NBOOT	BOOTSTRAP	69.1	PARAM
2	10	30	SAMPLE	139.4	INPUT
3	DATA			101	66
4	13	12	11	121.2	
5	106	13	301	86.6	
6	203	9	8	142.4	
7	131	5	106	88.1	
8	160	9	8	96.7	
9	8	4	13	107.9	
10	67	12	11	103.8	
11	61	10	67	112.2	
12	11	7	131	105.3	
13	301	9	8	64.2	
14		AVE	66.4	96.1	
15				90	
16				112.5	
17				136.9	
18				122.3	
19				125.6	
20				137	
21				82.5	
22				91.3	
23				109.7	
24				83	
25				110.7	
26				118.2	
27				156.6	
28				99.9	
29				95.1	
30				66.4	

The 5 left-hand columns of the worksheet should be reserved for the program. Various calculations can be carried out in the remainder of the spreadsheet, and will ordinarily be used to perform intermediate calculations in bootstrapping, as described below. To modify the program, one inserts the data in the lefthand column, the number of observations under NOBS and the number of bootstraps to be run under NBOOT. The actual bootstrap results appear in the 4th column which usually will have 1,000 to 2,000 entries (so nothing should be entered under this column in the worksheet, and one needs to remember to print out only the first page as simply hitting “print” may get you 40 pages or so!). The program presently is set up to provide bootstraps of the average as in Section 2.1.

Ordinarily there is little reason to bootstrap averages, and this worksheet is used only to show how to use the program and to correspond to the example of Section 2.1. The first column contains the actual observations so NOBS has to conform to the number of observations entered (10 here). Under NBOOT, enter the number of bootstraps you want to run. Ordinarily this will be 1,000 or so, but only 30 are used for illustration. The second column will show the serial number of the observation randomly selected by the program, and the third column shows the bootstrap sample (or “bootstrap replication”). Run the program and you will see these numbers change as each sample is drawn. Any function you want to bootstrap needs to be inserted somewhere. Presently it is in the third column below the bootstrap sample. In the present example, the average is calculated there. For convenience in keeping track of things, a label “AVE” has been inserted to the left of the average calculated from the bootstrap sample. Note, however, that if you use a larger number of observations, the program will write over the word “AVE”, and over the calculated value of the average, so you have to be careful to remove both “AVE” and the calculation to the right of that word before you change the number of observations.

In general, be careful what you do in the left-hand 5 columns or you may get some strange results. It is wise to save copies of the program as you experiment with changes, so that you can go back if something seems to be wrong with the current version.

The actual input to the program is in cell E3 (shaded in the worksheet, and just below “PARAM INPUT”) and if you highlight that cell you will see that it is directed to cell C14, which contains the function =AVERAGE (C4:C13) which averages the bootstrap replication. You can remove this cell and substitute another function. Try it with the variance function, =VAR(C4:C13) and run the program. It should produce 30 bootstrap values of the variance in column 4. The basic idea here is that you can compute anything that can be done directly on the EXCEL spreadsheet and enter it in cell E4 (“PARAM INPUT”) and 30 bootstrap values will be produced in column 4. Change NBOOT to 1000, and run the program. Then order the entries in column using the SORT function in the DATA window, and locate entries 25 and 975. These are the percentile confidence limits as described in Section 2.5.

Suppose you want the bootstrap standard error of eq.(2.1) for variances computed as above. The equation is:

$$\hat{se}_{boot} = \left\{ \frac{\sum [s(x^*b) - s(\cdot)]^2}{B-1} \right\}^{1/2} \quad (2.1)$$

Find the bootstrap mean (average of the bootstrap values in column D; this is $s(\cdot)$ in eq. (2.1)), and the calculation is shown in the following table:

NOBS	NBOOT	BOOTSTRAP	12861.2	PARAM
10	30	SAMPLE	14468	INPUT
DATA			12173.7	6754.84444
13	13	301	7102.1	
106	11	61	9249.78	
203	8	160	4737.57	
131	10	67	8640.27	
160	11	61	10479.2	
8	5	106	8991.57	
67	9	8	6107.61	
61	10	67	5704.9	
11	7	131	14642.4	
301	8	160	6204.9	
VAR		6754.8444	7834.23	
			13931.7	
			2359.88	
			4403.33	
			16701.6	
			8079.73	
			8469.73	
			4544.71	
			6070.67	
			4557.6	
			5182.44	
			5182	
			8124.77	
			13266.1	
			9682.27	
			5299.43	
			6754.84	
	MEAN		8393.61	

You can then conduct the calculation in column F by subtracting the overall mean from the individual values of variances [$s(x^{*b})$ in the equation above] and squaring (remember to use the \$ notation to “freeze” the value of the bootstrap mean), sum up the column of squared values, divide by B-1 and find the square root. This is illustrated for the variance below [all we do is add a column to the table above, sum, divide by B-1 (30 - 1 = 29 here) and take the square root]. Note that the additional calculations are simply added to the worksheet after the actual bootstrapping is done. However, if you want to see how the bootstrap standard error behaves in repeated runs, just run the macro repeatedly and observe how the result varies.

Bootstrap standard error calculation:

NOBS	NBOOT	BOOTSTRAP	12861.2	PARAM	19959076
10	30	SAMPLE	14468.0	INPUT	36898367
DATA			12173.7	6754.84444	14289344
13	13	301	7102.1		1667994
106	11	61	9249.8		733026
203	8	160	4737.6		13366642
131	10	67	8640.3		60840
160	11	61	10479.2		4349738
8	5	106	8991.6		357554
67	9	8	6107.6		5225784
61	10	67	5704.9		7229153
11	7	131	14642.4		39047395
301	8	160	6204.9		4790445
VAR		6754.84444	7834.2		312901
			13931.7		30670826
			2359.9		36405907
			4403.3		15922296
			16701.6		69022169
			8079.7		98518
			8469.7		5795
			4544.7		14814011
			6070.7		5396059
			4557.6		14714961
			5182.4		10311575
			5182.0		10314429
			8124.8		72276
			13266.1		23741173
			9682.3		1660640
			5299.4		9573920
			6754.8		2685548
	MEAN		8393.6		393698361.5
					SUM
					3684.536
					BOOTSTRAP.
					STANDARD
					ERROR

In the above case we produced the bootstrap values and then operated on these to find the bootstrap standard error. In other situations one may want to insert several stages of calculations before doing the bootstrapping. Consider the “parametric regression” calculations of Section 2.6. One needs first to find the residuals about regression. This can be done separately or on the same worksheet. For convenience the regression calculation is listed separately here, as follows.

EXCEL calculation for a regression line:

ORIGINAL DATA		<i>Regression Statistics</i>			
X	Y	Multiple R	0.9172596		
10	12.672	R Square	0.8413651		
12	8.9391	Adjusted R Square	0.8215358		
14	13.934	Standard Error	1.9880968		
15	16.377	Observations	10		
17	13.252				
21	19.121	<i>ANOVA</i>			
23	17.821		<i>df</i>	<i>SS</i>	<i>MS</i>
28	18.879	Regression	1	167.70690	167.70690
30	21.047	Residual	8	31.62023	3.95253
35	25.213	Total	9	199.32714	
			<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>
		Intercept	6.15279	1.74062	3.53483
		X Variable 1	0.51574	0.07918	6.51385
					<i>P-value</i>
					0.0076
					0.0001

RESIDUAL OUTPUT

<i>Observation</i>	<i>Predicted Y</i>	<i>Residuals</i>
1	11.31021	1.36179
2	12.34170	-3.40260
3	13.37318	0.56082
4	13.88893	2.48807
5	14.92041	-1.66841
6	16.98338	2.13762
7	18.01487	-0.19387
8	20.59358	-1.71458
9	21.62506	-0.57806
10	24.20378	1.00922

Copy the residuals into the first column of BOOT1, and copy the “Predicted” values in a column to the right of the “active” part of the sheet (first 5 columns). Construct new Y values by adding the bootstrapped residual to the predicted values. Calculate a slope (using the SLOPE function) from the new Y values and the X-values and enter this in the PARAM INPUT box (one could calculate the slope directly in this box, but it may be best to calculate it elsewhere on the spreadsheet and set PARAM INPUT equal to this value. If you want to also bootstrap the intercepts, obtain means of the new Y and the X

values and compute an intercept using the slope value. For simplicity, the present version of BOOT1 only handles one parameter, but one can run the program twice, tabulating slopes and then intercepts. The modified program follows. One would usually also order the slope values and find confidence limits, and make a histogram to plot the frequency distribution. With a little effort all of this can appear on the same spreadsheet so you can have all of your results together, and then pick off what's needed for a report. Run the program as set up above to try it out and better understand the descriptions given here.

Regression bootstrap:

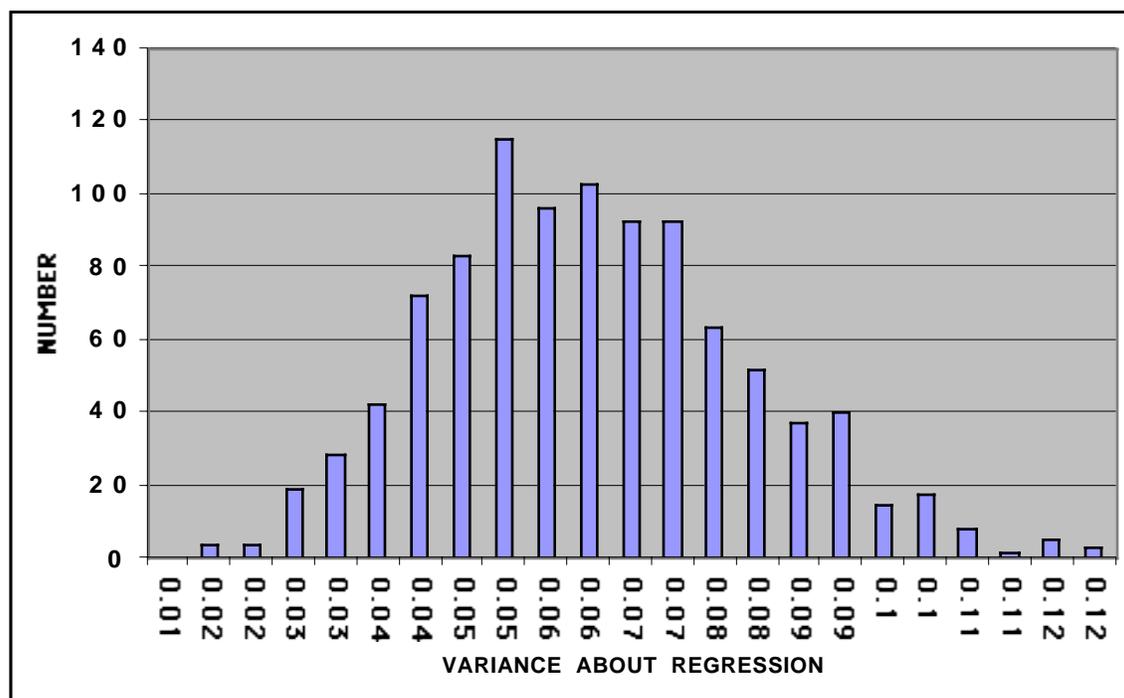
NOBS	NBOOT	BOOTSTRAP	0.5906 PARAM				
	10	20 SAMPLE	0.5250 INPUT				
DATA			0.5118	0.57185669	PREDICTED	NEW Y	x
1.3618	5	-3.40260	0.5497		11.31021	7.90761	10
-3.4026	7	2.48807	0.4476		12.34170	14.82977	12
0.5608	7	2.48807	0.5756		13.37318	15.86126	14
2.4881	13	1.00922	0.4399		13.88893	14.89815	15
-1.6684	12	-0.57806	0.5550		14.92041	14.34235	17
2.1376	6	0.56082	0.5187		16.98338	17.54420	21
-0.1939	4	1.36179	0.4442		18.01487	19.37665	23
-1.7146	10	-0.19387	0.4936		20.59358	20.39971	28
-0.5781	7	2.48807	0.5645		21.62506	24.11314	30
1.0092	13	1.00922	0.5426		24.20378	25.21300	35
			0.5342				
			0.5004		SLOPE	0.57186	
			0.4737				
			0.4817				
			0.4568				
			0.4901				
			0.5719				

BOOT2

To use nonparametric bootstrapping for regression data one needs a program that handles 2 columns of data. This is done in BOOT2 where the first 2 columns are input into a program. The version of BOOT2 furnished is set up to bootstrap the same data set as used above. Parametric regression bootstrapping should be used for smaller data sets (less than, say, 10 pairs of observations). One can see why this is so by running the program and plotting a frequency distribution of outcomes for smaller samples. This is an arbitrary sort of rule, based on experience running the 16 sets of data in Eberhardt (1987) with both BOOT1 and BOOT2. Thus I believe BOOT2 should be used for larger samples (say $n > 10$) inasmuch as it depends on less restrictive assumptions. The parametric approach requires that the assumed model holds for the data used. In the example shown here parametric and nonparametric regression gave about the same frequency distributions. For a more realistic example we use the larger data set of Example 3.4 and bootstrap the variance about regression. The first table below shows BOOT2 as it is currently set up, while the second table gives it with the example used here ($n = 19$).

To bootstrap the variance about regression, one adds calculations for means, slope and intercept as above and uses these to calculate a sum of squares [eq.(1.10)] and variance estimate which is then loaded into PARAM INPUT. The bootstrap results (B=1000 here) can then be summarized by HISTOGRAM (not shown here) and a frequency diagram constructed with the CHART WIZARD. The frequency plot indicates that the estimate of the variance about regression is itself quite variable. The value for the original data set is 0.0645 while the average of 1,000 bootstraps is 0.0583, suggesting a bias [eq.(3.2)] of 0.0062 which is likely not of much concern, considering the variability of the estimate as displayed below.

FREQUENCY DISTRIBUTION OF VARIANCES ABOUT REGRESSION FOR BEAR DATA



The next two programs likely do not need much explanation. They serve to jackknife one (JACK1) and two-column (JACK2) data. Applying JACK1 to the data of Section 2.2 gives the result shown below. All one needs to do is load the data in the left hand column and enter the number of observations. Running the macro then produces the basic jackknife data set, and one can proceed to add whatever calculations are desired. JACK2 is used in the same way, and is demonstrated here with the bear data (only part of the 19 sets of output is shown).

JACK1.xls

	A	B	C	D	E	F	G	H	I	J	K	L
1	NOBS											
2	10											
3	DATA											
4	13		106	13	13	13	13	13	13	13	13	13
5	106		203	203	106	106	106	106	106	106	106	106
6	203		131	131	131	203	203	203	203	203	203	203
7	131		160	160	160	160	131	131	131	131	131	131
8	160		8	8	8	8	8	160	160	160	160	160
9	8		67	67	67	67	67	67	8	8	8	8
10	67		61	61	61	61	61	61	61	67	67	67
11	61		11	11	11	11	11	11	11	11	61	61
12	11		301	301	301	301	301	301	301	301	301	11
13	301											

JACK2.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	NOBS												
2	19												
3	DATA												
4	1	2.8332		2	2.5649	1	2.8332	1	2.8332	1	2.8332	1	2.8332
5	2	2.5649		3	2.1972	3	2.1972	2	2.5649	2	2.5649	2	2.5649
6	3	2.1972		4	2.5649	4	2.5649	4	2.5649	3	2.1972	3	2.1972
7	4	2.5649		5	2.4849	5	2.4849	5	2.4849	5	2.4849	4	2.5649
8	5	2.4849		6	2.6391	6	2.6391	6	2.6391	6	2.6391	6	2.6391
9	6	2.6391		7	2.3979	7	2.3979	7	2.3979	7	2.3979	7	2.3979
10	7	2.3979		8	2.5649	8	2.5649	8	2.5649	8	2.5649	8	2.5649
11	8	2.5649		9	2.8332	9	2.8332	9	2.8332	9	2.8332	9	2.8332
12	9	2.8332		10	2.1972	10	2.1972	10	2.1972	10	2.1972	10	2.1972
13	10	2.1972		11	3.2189	11	3.2189	11	3.2189	11	3.2189	11	3.2189
14	11	3.2189		12	2.5649	12	2.5649	12	2.5649	12	2.5649	12	2.5649
15	12	2.5649		13	2.9444	13	2.9444	13	2.9444	13	2.9444	13	2.9444
16	13	2.9444		14	2.7726	14	2.7726	14	2.7726	14	2.7726	14	2.7726
17	14	2.7726		15	3.2189	15	3.2189	15	3.2189	15	3.2189	15	3.2189
18	15	3.2189		16	3.1781	16	3.1781	16	3.1781	16	3.1781	16	3.1781
19	16	3.1781		17	3.1355	17	3.1355	17	3.1355	17	3.1355	17	3.1355
20	17	3.1355		18	2.9957	18	2.9957	18	2.9957	18	2.9957	18	2.9957
21	18	2.9957		19	2.9957	19	2.9957	19	2.9957	19	2.9957	19	2.9957
22	19	2.9957											

The two columns of input data are on the left and the first 6 jackknife results appear to the right, with the first pair left out, then the second, third and so on. One can add various calculations (e.g., slopes) directly on the spreadsheet. It is wise to clear out all but the input columns before running the programs, i.e., JACK1 should appear as shown below before you run it. Inasmuch as it only takes a few minutes to run most of the programs, it is advisable to make a second run starting from the original program if anything looks questionable in your results. When using 2000 bootstraps with some mildly complicated secondary calculations, such as BOOT2 with the variance about regression calculations as shown above, it does take some time for the program to run. On a MAC PowerBook

G3 it took about 15 minutes, while a HP Pavilion portable (100mhz CPU) took about 10 min.

BASIC ARRANGEMENT FOR JACK 1

NOBS

19

DATA

1	2.8332
2	2.5649
3	2.1972
4	2.5649
5	2.4849
6	2.6391
7	2.3979
8	2.5649
9	2.8332
10	2.1972
11	3.2189
12	2.5649
13	2.9444
14	2.7726
15	3.2189
16	3.1781
17	3.1355
18	2.9957
19	2.9957

SPECIAL PURPOSE PROGRAMS USING BOOTSTRAPPING

The programs given above are general purpose in that they are written to use one or two columns of data, and the results can be modified on the spreadsheets to provide a variety of results. Some further programs follow, designed for more specialized purposes. However, the same underlying theme is pursued. That is, the MACROS bootstrap a data set, and the user makes changes on the EXCEL spreadsheet to accomplish a variety of different purposes.

Approximating the Lotka-Leslie model

The first special purpose program to be considered here is APPLMB, which has been prepared to bootstrap eq.(11.9), the approximation suggested by Eberhardt (1985) and further evaluated in Eberhardt (2002). The underlying equation is:

$$\lambda^a - s\lambda^{a-1} - \mathbf{l}_a \mathbf{m} \left[\mathbf{1} - \left(\frac{s}{\lambda} \right)^{\mathbf{w} - \mathbf{a} + \mathbf{1}} \right] = 0 \quad (11.9)$$

An example of the program using data on manatees follows, with the equation associated with the box labeled RES shown.

PROGRAM APPLMB:

APPLMB.xls											
	A	B	C	D	E	F	G	H	I	J	K
1											
2		NOBS	RATE								
3		S	1.01								
4		S	SJUV	m	a	w	lbd	RES	Estimate of lambda	Residuals	No. of iterations
5		0.9632	0.8444	0.1665	4	50	1.0660	-0.00009			
6		0.9548	0.8889	0.1327					1.0437	-0.0001	105
7		0.9648	0.7556	0.1629					1.0571	-0.0001	112
8		0.9655	0.8557	0.1432					1.0584	-0.0001	107
9		0.9690	0.8222	0.1567					1.0549	-0.0001	118
10		0.9828	0.8567	0.1515					1.0804	-0.0001	125
11		0.9833	0.7778	0.1583					1.0555	-0.0001	113
12		0.9751	0.8000	0.1783					1.0784	-0.0001	123
13		0.9680	0.8000	0.1493					1.0580	-0.0001	117
14		0.9632	0.8444	0.1665					1.0660	-0.0001	109
15											
16											

The main program does not do any bootstrapping. All of the bootstrapping is done in advance by using BOOT1 and BOOT2, and is done on the three input variables, S, SJUV, and m before using APPLMB. The only purpose of APPLMB is to produce estimates of lambda from the input data. One thus has to bootstrap adult survival data (S) and enter, say, 1,000 bootstrap values in that column. Values of m are similarly bootstrapped and entered in the column headed by m. The age of first reproduction (a) and the maximum age (w) are fixed and need only be entered once. SJUV corresponds to l_a in eq.(11.9). If survival is assumed to be at the adult rate after the first year of life then the first-year survival can be entered in the column headed SJUV (this was done for manatees as there were no estimates of early survival beyond age 1). In some examples, however, there are at least 2 early survival values, so SJUV has to be calculated as the product of two or more survival rates. Each of these rates has to be bootstrapped separately and the results combined (multiplied) to get SJUV, and the equation governing RES modified accordingly. Consequently, one will have several different bootstrapping sheets containing the data to be entered into APPLMB. In the example shown above there are only 9 entries just to show how the program is set up; in the usual case there will be at least 1,000 entries in the three left-hand columns.

The actual calculations are done by an equation in the spreadsheet associated with the box headed RES. The program runs thru all 1,000 rows, calculates lambda in the box headed lbd, and stores the estimates of lambda in the column headed "Estimate of lambda". The value of the residuals is also listed to the right, so that one can check to see that the program converged properly. A maximum of 2000 iterations is used to solve eq.(11.9), so if the last column contains the number 2000, the program likely did not converge adequately. Users need to understand that eq.(11.9) is nonlinear in nature, and the program may thus fail to converge. Hence a very important item appears in the box headed RATE. This is the initial "guesstimate" of lambda. I suggest setting NOBS at 30 (which limits the number of bootstraps used to the first 30), and running the program. If there are problems with convergence (i.e., 2000 shows up repeatedly in the box headed "Number of iterations"), then experiment with the value used for RATE until the number of iterations is less

than 2000. EXCEL contains the very useful program, SOLVER, in the TOOLS menu. When you activate SOLVER, a box comes up asking for the “Target” cell, and this should be set as the box under RES. It also asks for the cell to vary and this is the box under “lbd”. When you run the program APPLMB, you will note that it sequentially loads all of the data into the first line and solves eq. (11.9) with this data. Hence if you want to use SOLVER to check a result you have to copy data corresponding to the case in question into the first 3 columns of the first row (columns headed S, SJUV, and m), and then activate SOLVER. Be sure to read the “help” items associated with SOLVER, as there are various things that you can do to facilitate getting a solution, and to check that the solution is appropriate. I have used the routine built into APPLMB on many data sets, and have not had any important problems, so long as I used good “guesstimates” to start with. You can demonstrate failure of the routine by putting in an unreasonable values of RATE – -it will yield all sorts of strange outcomes. Incidentally, λ should never be set equal to s. I usually start with $\lambda = 1.01$. If you get error messages, usually the box that comes up will have a button titled “END”, and pressing this gets you back to the spreadsheet.

For an example, we use the data on Yellowstone grizzly bears that produced the frequency distributions of Fig. 11.8. The data shown there were produced by a program written especially for the purpose. Using the programs given here reduces the effort required considerably and does not require a knowledge of a programming language. Applications of BOOT1 and BOOT2 produce the basic data. Survival can be estimated in various ways, but in this case radiotelemetry was used, so the individual records are those of “radio-days” of observation per individual bear. Some of the individuals died during the course of the study, and a simple estimator was used:

$$S = (1 - d/\text{days})^{365}$$

Where d = number of deaths observed and days = number of “radio-days” tallied. The bootstrapping program contains a series of records of individual bears with deaths tallied as 1’s and survival tallied as 0’s in the y-column while the corresponding days of observation for the individual appear in the x-column. The data are as shown in the table below, and were inserted in BOOT2 or BOOT1 (for reproductive rate) and 1,000 bootstrap samples drawn. Bootstrap values (means of 1,000 bootstraps) are shown in the table along with the direct estimates of the several parameters.

When running APPLMB, be sure to pick the correct Macro out of the list that comes up when you select Macro in the TOOLS menu. When you copy in data from the various individual bootstrap sheets (for adult survival, etc.), the Macros for each such subset become available, so if you select the wrong one, you will get an error message.

INPUT DATA FOR PROGRAM APPLMB APPLIED TO DATA ON YELLOWSTONE GRIZZLY BEARS

AD SURV	DAYS	SUBADULT SURV		CUB SURV	DEATHS	REPRO
DEATHS		DEATHS	DAYS	SURVIVE		RATE
1	521	1	488	2	0	0.5
0	730	1	363	2	0	0.6
0	863	0	130	2	0	0.888889
1	143	0	259	3	0	1
1	126	0	342	2	0	1
0	732	0	187	2	1	0.571429
1	447	0	930	1	1	0.714286
0	1684	0	190	3	0	0.571429
0	2512	1	248	2	0	0.833333
0	1100	1	209	3	0	0.8
0	540	0	238	2	0	0.666667
0	977	0	967	1	2	1
0	1056	0	49	3	0	0.833333
0	2182	1	80	2	0	0.571429
0	1462	0	465	2	0	0.6
0	1097	0	229	3	0	0.777778
0	17	0	259	2	0	1
0	1053	0	952	2	0	0.5
0	961	0	951	1	1	1
0	713	0	441	0	2	0.666667
0	715	0	196	2	0	0.666667
0	537	0	89	3	0	1.2
0	775	0	70	1	2	0.5
0	924	5	8332	1	0	0.4
0	951			1	0	0.5
0	699	surv	0.803	1	1	0.4
0	831	BOOT		3	0	1.25
0	840	MEAN	0.798	2	0	0.666667
0	29			3	0	0.75
0	868			1	1	1
0	647			58	11	1
0	650					0.333333
0	655			SURVIVAL	0.810	0.333333
0	935			BOOT		1
0	491			MEAN	0.843	0.75
1	58					0.666667
0	786					0.333333
0	500					0
0	415					1.5
5	31222					0
						0.333333
S=	0.943					0
BOOT						0
MEAN	0.942					0.5
						1
						0.666667
						1.5
						1
				AVE		0.695
				BOOT	MEAN	0.695

As noted previously, cub survival was multiplied by subadult survival to get the data to enter for JSURV in the final program, using the equation:

$$JSURV = (\text{cub surv})(\text{subadult surv})^3$$

The final program was adapted from APPLMB by changing the formula attached to RES, so that the reproductive rate was divided by 2, to use female births per female. Also, there is a single entry for SJUV due to the calculation shown above being done on a separate spreadsheet. One has to be sure to get the proper equation in place for RES, but this makes it possible to use one program here instead of several variants. The first part of the final program follows, showing the input data and estimates of lambda (the full output contains 1,000 values).

RES = =G5^E5-B5*G5^(E5-1)-(D5/2)*C5*(1-(B5/G5)^(F5-E5+1))

APPLMBbear											
	A	B	C	D	E	F	G	H	I	J	K
1											
2		NOBS									
3		1000									
4		S	SJUV	m	a	w	lbd	RES	Estimate	Residuals	No. of
5		0.966442	0.29994	0.689666	4	20	1.0186	9.52E-05	of lambda		iterations
6	1	0.94974	0.353568	0.761789					1.044361	-9.21E-05	73
7	2	0.91558	0.36263	0.63991					1.006698	5.96E-05	110
8	3	0.949687	0.36766	0.61701					1.026748	9.6E-05	103
9	4	0.972916	0.404718	0.753158					1.072843	-9.71E-05	131
10	5	0.956332	0.584835	0.661657					1.089553	-9.25E-05	110
11	6	0.954833	0.403478	0.679315					1.049542	-9.49E-05	123
12	7	0.920598	0.555587	0.654663					1.059422	-8.87E-05	85
13	8	0.92984	0.529066	0.642824					1.057422	-9.09E-05	92
14	9	0.969526	0.373506	0.750248					1.061492	-9.52E-05	139
15	10	0.946021	0.378807	0.68292					1.037859	9.57E-05	49
16	11	0.967661	0.361823	0.636359					1.04	4.53E-05	1
17	12	0.978505	0.553779	0.679117					1.100384	-8.95E-05	131
18	13	0.942578	0.362558	0.743667					1.03999	9.01E-05	2
19	14	0.974791	0.470335	0.799074					1.097914	-9.37E-05	127
20	15	0.966787	0.479183	0.693304					1.078493	-9.13E-05	120

Observant readers will note that this example varies a little from that shown earlier in that the box for RATE was inserted after this example was run (the starting value was originally in the program, and has now been moved to the spreadsheet).

The bootstrap value of lambda was 1.053, as is the value calculated from the original data, using SOLVER. The 95% confidence interval for lambda is 0.97 to 1.12. These wide limits are a consequence of the limited data on subadult survival. Using the delta method shows that 77% of the variability of the estimate of lambda is due to the small sample for subadult survival (Eberhardt 2002:Table 2). Radiotelemetry has been continued on the population since the above data were collected, and hopefully a better estimate will eventually be available (a trend estimate provided by Eberhardt et al. (1999) gives a somewhat better notion of the likely status of the population). It is assumed here that the different sets of data used for inputs are independent. If they are not independent, the confidence limits will be narrower than they should be.

Hence, it is a good idea to check the correlations between the bootstraps. In this example, they were all quite small (less than 0.10).

Survival estimation with the Jolly-Seber method

Estimating survival is an essential feature of any study of population dynamics. The program used here (BOOTSURV) follows the approach of Seber (1982), and provides survival estimates by the Jolly-Seber method. There is a necessary dichotomy in applications of capture-recapture methods. The original theory assumes that all previously unmarked animals are marked and released as they are captured. This is usually not feasible with large mammals, where studies usually begin with a substantial marked sample, with new "batches" marked when it is practicable to do so. When the actual capture-recapture study begins, the first capture of a marked individual is then effectively regarded as the initial capture and marking.

The essential equation for the Jolly-Seber method is (Seber 1982:200):

$$\hat{M}_i = \frac{R_i z_i}{r_i} + m_i \quad (i = 2, 3, \dots, s-1)$$

Here R_i denotes the number of marked individuals released at time i , z_i is the number captured before and after i , but not at time i , r_i denotes those animals released at i , and caught again later, and m_i are previously marked animals caught at time i . The z_i and r_i are usually obtained by constructing two intermediate tables as illustrated by Seber (1982:206). A different approach is used in the present program, in that z is estimated from two other parameters, i.e., $z = C - k$ in the table produced by BOOTSURV (see Section 8.5 for an explanation).

The basic data are represented by a rectangular matrix in which the rows represent the capture history of individuals with 1 denoting capture on the i th occasion and zero, no capture, and the columns contain the data for the capture occasions.

The Jolly-Seber method estimates survival as (Seber 1982, eq.(5.9), p. 200):

$$\hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + R_i} \quad (i = 2, 3, \dots, s-2)$$

BOOTSURV uses 3 different worksheets. The first one (DATA) contains the original data, which will be, as noted above, a matrix of zeros and ones. The data must start in the upper left-hand corner of the sheet, and no other information appears on the sheet – it is called by the programs associated with the other two sheets. The second sheet is titled "Original Data" and calculates survival by the equations given above. The box titled "NROWS" contains the number of rows of data (number of individuals for which data is available), and the second box, titled "NCOLS" contains the number of columns of data, which correspond to the number of capture occasions. When you run the program associated with this sheet, it prints out the data on the right,

permitting one to check to be sure that the data are correct. The program is illustrated with data on the Florida manatee collected at the Blue Springs site. There were 15 capture occasions, and the equation for survival above produces survival estimates for occasions 2,...,s-2 (where s is the number of capture occasions), and these 13 values appear in the right-hand column of output, and are averaged below. You start the program either by using the Macro call in the Tools menu, or by hitting Alt & F8. When you do this a box comes up that says "Bootstrap1" and "Original data". Be sure to highlight "Original data" in order to carry out the analysis described above. "Bootstrap1" is automatically highlighted, so if you just hit "return" you will get an error message (just press "End" in the box that comes up to get back to the worksheet). You must highlight the appropriate name to run the corresponding program. The picture on the next page shows the Original Data worksheet for the Blue Springs manatee data. Note that the columns headed M-hat to Jssurv contain calculations, so do not erase these in preparation for a new run. The program calculates C, k, R and m from the DATA sheet.

Once you have run the analysis on the sheet titled "Original data", switch to the sheet titled "Bootstraps" (a picture of that sheet follows the Original Data sheet below). The box labeled "NBOOT" in the upper left-hand corner determines the number of bootstraps. I suggest starting out with 30 until you are sure you have the program running properly. The program obtains the number of rows and the number of columns from the previous sheet ("Original data") and operates only on that part of the data. However, if it has been used for a different data set at some previous time, there may be more rows of data than called for by "Original data", which can be confusing. I suggest clearing the columns below C,k,R,m and r before you start. If you clear more columns you will destroy the outputs you want (just close down and start over to restore the computations). When you run the sheet the above columns are filled in and the survival rate is computed under "surv". This program is set up to tally only an average survival rate, so you need to be sure that an average of the presently used data is calculated below the "surv" column and that the box under "PARAM INPUT" is set equal to this average. Another program is available that prints out ALL of the bootstrap survival estimates and does not average them. It can be used to study the pattern of survival over time, but is not considered here for the present. The data shown here are in the copy of the program included with this Appendix, so readers can experiment with running the program with this data.

	A	B	C	D	E	F	G	H	I	J	K	L
1	NROWS	C	k	R	m	r	M-hat	p-hat	JSsurv			
2	68	13	13	16	13	16	13.00	1				
3	NCOLS	16	16	19	16	19	16.00	1	1			
4	15	19	18	21	18	21	19.00	0.94737	1			
5		21	20	23	21	21	22.10	0.95043	1.0043			
6		21	21	26	22	25	22.00	1	0.913			
7		25	18	20	18	20	25.00	0.72	0.9615			
8		25	23	32	25	29	27.21	0.91888	1.0077			
9		31	30	36	30	35	31.03	0.96685	0.9071			
10		34	34	40	36	36	36.00	1	0.9722			
11		35	34	39	35	38	36.03	0.97151	0.9007			
12		38	37	40	38	39	39.03	0.97372	0.975			
13		38	35	42	37	40	40.15	0.92154	0.9787			
14		40	39	50	42	45	43.11	0.97423	0.9548			
15												
16												
17								AVERAGE	0.9646			
18												
19												
20												

	A	B	C	D	E	F	G	H	I	J	K
1	NBOOT	66	C	k	R	m	r	surv	M-hat	0.96465	PARAM
2	1000	51	10	10	13	10	13		10	0.95478	INPUT
3		36	13	13	18	13	18	1	13	0.96482	0.9651
4		12	18	17	19	17	19	1	18	0.96556	
5		55	19	18	20	19	19	1.003	20.05	0.96899	
6		36	19	19	21	20	20	0.95	20	0.98281	
7		48	20	14	15	14	15	0.952	20	0.96331	
8		22	20	20	33	21	30	1	21	0.97514	
9		62	30	30	37	30	34	0.909	30	0.96799	
10		51	34	34	39	34	36	0.919	34	0.96325	
11		6	35	28	33	29	32	0.929	36.22	0.97164	
12		49	37	37	40	39	38	0.97	39	0.97102	
13		46	38	36	40	36	40	0.95	38	0.96454	
14		53	32	32	52	42	38	1	42	0.95811	
15		60								0.97328	
16		19					AVE	0.965		0.98007	
17		7								0.98309	

Another example of the use of the approximation to the Lotka-Leslie equation

The Blue Springs manatee data provide another example of the use of APPLMB, using the survival data obtained above with the Jolly-Seber method. The bootstrap results were used to introduce use of program APPLMB and the survival data were used to illustrate the use of BOOTSURV above, with the original survival data incorporated in the appended copy of BOOTSURV. The bootstrap means were very close to the estimates from original data, with the adult survival estimate from original data being 0.965, the same as the bootstrap mean. The same applies to reproductive data with both averages being 0.151. The original data for first year survival gave 0.822, while the bootstrap mean was 0.821 (a second run gave 0.824). The adult survival data appear in the picture of the Original Data sheet above, while first year survival was computed from 37 survivals from 45 individuals tagged. Thus the first year survival was bootstrapped by using an input table with 37 values of 1, and 8 zeros in BOOT1. The reproductive rate data (births per female per year) are shown below for use by anyone wanting to duplicate the bootstrapping. Just copy the data below into the clipboard – remember to divide by 2 to get female births per female.

0.3333
 0.3571
 0.25
 0.5
 0.3077
 0.3333
 0.2
 0.2222
 0.1667
 0.2857
 0.2857
 0.3333
 0.3333
 0.5
 0.3333
 0
 0.25
 0.3333
 0.2
 0.5

Estimating a parameter using values of lambda from trend data.

When data are not available to estimate a parameter other than lambda (early survival, for instance), it may be possible to substitute estimates of lambda from trend data. This is readily done by rearranging eq.(11.9).

$$\lambda^a - s\lambda^{a-1} - l_a m [1 - (\frac{s}{\lambda})^{w-a+1}] = 0 \quad (11.9)$$

For the Blue Springs manatee data, l_a was estimated by assuming survival beyond the first year to the age of first reproduction (a) was at the adult rate, so that

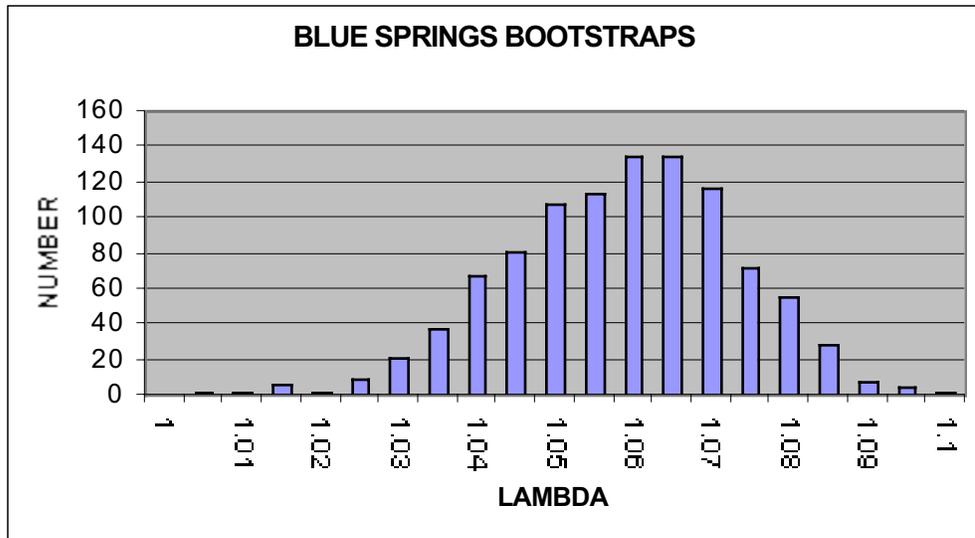
$$l_a = s_{juv}s^{a-1}$$

Thus we can estimate first year survival from [rearranging eq.(11.9):

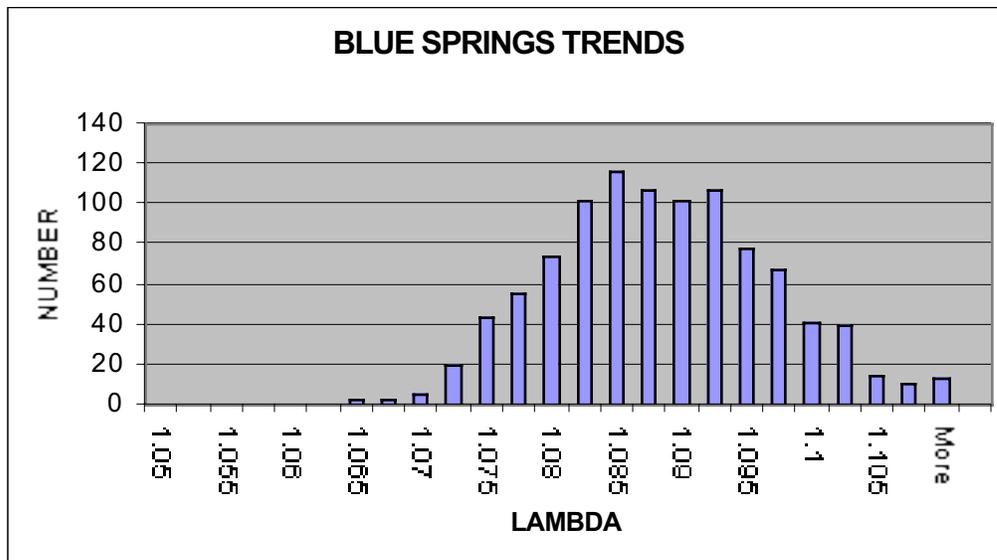
$$s_{juv} = \frac{\lambda^a - s\lambda^{a-1}}{s^{a-1}m[1 - (\frac{s}{\lambda})^{w-a+1}]}$$

and m can be estimated by a similar rearrangement. Estimating adult survival (s) would require an iterative solution, but, while APPLMB can be revised to do that, I do not recommend it because adult female survival is such an important parameter that it should be estimated directly. Bootstrapping can be arranged by using the above equation. Nineteen years of trend data were available for Blue Springs manatees and were bootstrapped with BOOT2 to get 1,000 estimates of λ from trend data. These can then be used with the available bootstraps of adult female survival and reproductive rate, arranged on a worksheet, and first year survival estimated from the above equation.

Most of the estimates of SJUV are unreasonably high. Why this should be the case can be appreciated by comparing a frequency plot of the bootstrap estimates of lambda estimated from reproductive and survival data with one estimated by bootstrapping trend data from Blue Springs manatees. These plots are shown below. The trend data give much higher estimates of λ than were obtained by bootstrapping reproductive and survival data. Eberhardt and O'Shea (1995) reported indications of immigration to the two sites for which trend data were obtained (Crystal River and Blue Springs). Estimates of early survival from the Crystal River site were not available, so the main indication that there might be immigration likely has to come from the failure of the above program to produce satisfactory estimates for that site, too, along with the impressions recorded by Eberhardt and O'Shea (1995).



Frequency plot of bootstraps of estimates of λ based on reproductive and survival data.



Frequency plot of estimates of λ obtained by bootstrapping trend data.

Bootstrapping multiple regressions

Multiple regressions may be useful in assessing trend data, as indicated in Section 9.10. Hence it is worthwhile to have two programs to bootstrap multiple regressions. MREGNONP takes the usual bootstrap approach, in which the set of observations is randomly sampled n times with replacement, and a multiple regression is fitted to each of the B bootstrap samples. The calculations are done on the sheet titled "DATA", and the bootstrapped values of the regression calculations appear on a sheet titled "BOOT". It is convenient to attach another sheet containing the usual EXCEL regression calculations,

and that sheet is labeled “REGRDATA” in the attached programs, which are based on the grizzly bear data of Example 9.6. The regressions are calculated by matrix operations which appear of the right of the sheet titled “DATA”. Up to 5 independent variables can be accomodated, but note that the box under NVAR should show p+1 where p is the number of independent variables. The first column (labeled Xo) has to contain a column of ones for calculating the intercept by matrix methods. The sheet labeled “DATA” is as follows (without the matrix calculations which are off to the right of the sheet):

MREGNONP.xls															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	NOBS	NBOOT	NVAR												
2	20	20	4					RAND							
3	Y	X0	X1	X2	X3	X4	X5	NO.	Y	X0	X1	X2	X3	X4	X5
4	2.8332	1	1	1.6	7.60			15	2.944	1	13	2.12	-15.7		
5	2.5649	1	2	1.5	10.40			6	2.197	1	3	1.3	-18.4		
6	2.1972	1	3	1.3	-18.40			15	2.944	1	13	2.12	-15.7		
7	2.5649	1	4	1.1	3.90			4	2.833	1	1	1.6	7.6		
8	2.4849	1	5	1.4	1.60			20	2.996	1	18	1.67	-6.33		
9	2.6391	1	6	1.6	-4.40			21	2.996	1	19	1.47	3.27		
10	2.3979	1	7	1.6	-9.80			19	3.136	1	17	1.65	0.47		
11	2.5649	1	8	1.2	9.60			8	2.485	1	5	1.4	1.6		
12	2.8332	1	9	2.3	2.90			7	2.565	1	4	1.1	3.9		
13	3.2189	1	11	3.12	-0.60			17	3.219	1	15	1.95	8.97		
14	2.5649	1	12	1.64	3.57			9	2.639	1	6	1.6	-4.4		
15	2.9444	1	13	2.12	-15.73			10	2.398	1	7	1.6	-9.8		
16	2.7726	1	14	1.86	-3.13			9	2.639	1	6	1.6	-4.4		
17	3.2189	1	15	1.95	8.97			11	2.565	1	8	1.2	9.6		
18	3.1781	1	16	2.65	-4.73			4	2.833	1	1	1.6	7.6		
19	3.1355	1	17	1.65	0.47			11	2.565	1	8	1.2	9.6		
20	2.9957	1	18	1.67	-6.33			8	2.485	1	5	1.4	1.6		
21	2.9957	1	19	1.47	3.27			14	2.565	1	12	1.64	3.57		
22	2.8332	1	20	1.47	-5.33			12	2.833	1	9	2.3	2.9		
23	3.4965	1	21	1.96	12.67			13	3.219	1	11	3.12	-0.6		
24															

Note that the estimates of regression coefficients appear as the lowest set of values just to the right of the material shown in the picture above. These values (4 in the example) are loaded into the sheet labeled “BOOT” as the bootstrapping proceeds. They can be compared with the values obtained from the original data on the sheet labeled “REGR DATA”. Usually the focus is on the second parameter which estimates the rate of change of the population, so one would obtain bootstrap confidence limits on that estimate. The trend data in this example give a narrower confidence interval than did the data used to illustrate APPLMB, which used survival and reproductive data to estimate λ .

The second program, MREGPARA, uses the parametric approach to the same data set, with the deviations from regression and “predicted” values calculated by the usual EXCEL regression program (attached to the program as sheet “REGRDATA”) in the lefthand 2 columns. The deviations are bootstrapped and attached to the predicted values to generate a new y-value next to the values of the independent variables as follows:

MREGPARA.xls															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	NOBS	NBOOT	NVAR												
2	20	20	4					RAND							
3	DEV.	PREDICT	DEV SELECTED	ORIG Y				WO.	Y	X0	X1	X2	X3	X4	X5
4	0.243	2.590674	-0.014	2.8332				18	2.577	1	1	1.6	7.6		
5	-0.05	2.614539	-0.094	2.5649				22	2.521	1	2	1.5	10.4		
6	0.02	2.177489	0.091	2.1972				17	2.269	1	3	1.3	-18.4		
7	0.114	2.450921	-0.020	2.5649				21	2.431	1	4	1.1	3.9		
8	-0.066	2.551125	0.135	2.4849				23	2.687	1	5	1.4	1.6		
9	0.083	2.555726	-0.122	2.6391				11	2.433	1	6	1.6	-4.4		
10	-0.126	2.523973	-0.094	2.3979				22	2.43	1	7	1.6	-9.8		
11	-0.122	2.687326	0.243	2.5649				4	2.93	1	8	1.2	9.6		
12	-0.139	2.972321	0.163	2.8332				15	3.135	1	9	2.3	2.9		
13	-0.03	3.248444	0.135	3.2189				23	3.384	1	11	3.12	-0.6		
14	-0.3	2.86451	0.083	2.5649				9	2.948	1	12	1.64	3.568		
15	0.163	2.781467	0.079	2.9444				20	2.86	1	13	2.12	-15.73		
16	-0.13	2.902119	0.163	2.7726				15	3.065	1	14	1.86	-3.132		
17	0.091	3.127435	0.161	3.2189				19	3.288	1	15	1.95	8.968		
18	-0.014	3.191673	0.079	3.1781				20	3.27	1	16	2.65	-4.732		
19	0.161	2.924505	0.163	3.1355				15	3.137	1	17	1.65	0.468		
20	0.079	2.917087	-0.094	2.9957				22	2.823	1	18	1.67	-6.332		
21	-0.02	3.015525	0.020	2.9957				6	3.035	1	19	1.47	3.268		
22	-0.094	2.926931	-0.066	2.8332				8	2.861	1	20	1.47	-5.332		
23	0.135	3.361032	-0.139	3.4965				12	3.222	1	21	1.96	12.67		
24															
25															
26															

Multiple regression calculations are then done on this data set.

Calculations for the Jolly-Seber and Manly Parr methods

Calculations for these methods as given in Sec. 8.2 are accomplished with the program JSMP. The data used are those of Examples 8.5 and 8.6, so that readers can compare outputs to those given there. The program is arranged in the same manner as the other programs presented here, i.e., with the data in a sheet labeled DATA, the basic calculations in a sheet labeled "Original data", and bootstrapping is accomplished in the sheet labeled BOOTSTRAPS. Equations for variance estimates are given by Seber (1982) and Pollock et al. (1990), but are not included in the present programs. This is because these variance equations are "large sample" results, whereas bootstrapping may be more efficient with smaller samples. In order to accommodate both the Jolly-Seber and Manly-Parr methods in one program, some of the final calculations are left to the user, with the essential data calculated on the sheets. Equations (8.25) and (8.26) and eqs. (8.29) to (8.32) can all be calculated from the data generated on the "Original data" sheet, recalling that $z_i = C_i - k_i$. The desired calculations can be installed to the right of the parameter estimates C thru r (but do not insert or remove any columns). Note that "batch" tagging requires insertion of the total number (n) caught each day as that information is not available from the DATA sheet, which only contains capture histories of those animals marked and subsequently caught at least once. Note that several basic calculations are already present on the sheets to the right of the basic items C thru r. In repeated use of the programs one needs to be careful to be sure that outputs from earlier uses with a larger number of capture periods do not carry over to a new use (i.e., make sure you do not have any extra rows in the "Original data" sheet when you make a new run).

The BOOTSTRAPS sheet produces the desired number of bootstraps (inserted below NBOOT in the left-hand corner of the sheet), with "M-hat" already calculated for the Jolly-Seber method. Spaces labeled "p" and "N" currently contain the remaining calculations for the Jolly-Seber method ("p-

hat" and "N-hat'). To change to the Manly-Parr method, change the formulas in these two spaces ("p" and "N") by eqs. 8.30 ("p-tilde") and eq. (8.31) for "N-tilde". The column with heavy borders just to the right of the column headed "N" should contain the value to actually be bootstrapped. That is, if you want to bootstrap whatever is in the column headed "N", simply set the values in the column with heavy borders equal to those in the column headed "N". When a data set with more than the three rows of the present program is to be bootstrapped, make a run with a small number in NBOOT and copy down those entries to the right as needed before running the program with NBOOT set equal to 1,000 or more (those entries under columns headed C thru r are set by the entries under NROWS and NCOLS on the "Original data" sheet).

Two additional sheets are attached to the program ("Laysan data" and "Lisianski data") which contain the original data for two sets of censuses of Hawaiian monk seals. These data can be transferred to the DATA sheet and the program run to see how it behaves with a large data set. The Laysan data were used in example 8.6 which gives the calculations from the "Original data" sheet and bootstrap confidence limits.

APPENDIX B

SOME DATA-FITTING PROGRAMS

The text includes various references to data-fitting programs, such as non-linear least-squares, maximum likelihood, and a data-smoothing program (“lowess”). Many of these procedures are available in commercial “packages”, such as SYSTAT, SAS, MINITAB, etc. Those used in various chapters were written in QUICKBASIC, but that language is no longer supported by Microsoft, being largely replaced by VISUAL BASIC and VISUAL BASIC FOR APPLICATIONS (which underlies the bootstrapping programs given in Appendix A). Readers who have access to one of the commercial packages may want to use them to analyze data. An alternative, used here, is to take advantage of the “R-language” which is now available free over the internet. It is much the same as the S-language which is, however, a proprietary commercial package. Both “S” and “R” require some effort to utilize efficiently. All that will be attempted here is to provide the essential statements that are needed to do some data-fitting.

Users need to download the language (look up CRAN on the web with a browser). It may take an hour or more to download. Once the program is downloaded and installed, you should see a bold “**R**” on your desktop. Double-click it and the “workspace” or “console” opens up, with a symbol `>` as a prompt. In Windows there should be drop-down menus at the top. Under “help” you can find a number of items, one of which is “Manuals” which contains an online manual. Various other manuals are available online. Another drop-down menu is “packages” which provides access to some “packages” that are not continuously available from the “console”, but are stored on your computer and can thus be brought up as needed. Below the drop-down menus there are a number of symbols, including “load image” and “save image” which let you save and recall programs that you work on.

There are a few basic commands one needs to know before doing any data-fitting (a summary of some useful commands is included at the end of this Appendix). One feature is that the “workspace” is automatically saved when you quit, so one needs to check to see what is there when you start again. Use the command `ls()` to see what’s in the workspace. If you don’t want to use what is there, type `rm(list=ls())`. This clears out the workspace [type `ls()` and you should get “character(0)” which means nothing is there].

You can also use `rm(y)` to remove an individual data set or whatever is labeled “y”. There is a menu at the top of the workspace that provides some useful features. You can also get some help by typing `?rm` or `?lowess`, but it takes awhile to figure out what some of the statements in the help files mean. When you are working in R it is very convenient to use the “arrow keys” to change or correct statements. Use the up arrow key to locate a statement, and the left-arrow key to find a position in the statement. Make a correction or change by using the delete key, and type in the change. Then use the right arrow key to return to the end of the statement, followed by a return. **IT IS IMPORTANT TO NOTE THAT THE R-LANGUAGE IS “CASE SENSITIVE”, I.E., WHEN LOWER CASE IS CALLED FOR AND YOU MISTAKENLY TYPE CAPS, IT DOESN’T WORK.**

LOWESS

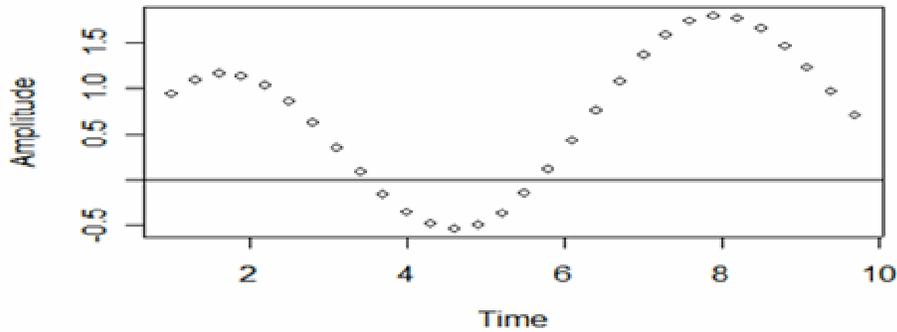
It is assumed initially here that you will have data in an EXCEL file to work with, or will copy data from the text. You can enter data directly, too (see the attached list of handy statements for instructions). The first program to be considered here is “lowess”, discussed in Section 1.14. Note that section says that “lowess” and “loess” are used interchangeably in the literature (and they were). However, there are now two different R-programs under those names. The one given here, and labeled “lowess” is described in Section 1.14, but a new program is present in R, called “loess”. It is more complicated, including optionally different weighting and possible use of a quadratic for projections in addition to the linear regression used in the earlier version. We will use “lowess” here. The example is artificial, being constructed from a sine curve with a small added component, i.e., $\sin(x)+0.1*x$. The basic data are on the following sheet, with headings of “Time” (which is x) and “Amplitude” which is the value of the sine of x. Copy the data into the clipboard, including the headings. The data given here was copied from EXCEL and you can copy it directly from the table below, or reconstruct it in EXCEL by using the formula given above. You may need to look up the Adobe instructions for copying from a PDF file, if you are using the PDF version of this Appendix.

Time	Amplitude
1.00	0.941471
1.30	1.093558
1.60	1.159574
1.90	1.13630
2.20	1.028496
2.50	0.848472
2.80	0.614988
3.10	0.351581
3.40	0.084459
3.70	-0.15984
4.00	-0.3568
4.30	-0.48617
4.60	-0.53369
4.90	-0.49245
5.20	-0.36345
5.50	-0.15554
5.80	0.115398
6.10	0.427837
6.40	0.756549
6.70	1.07485
7.00	1.356987
7.30	1.580437
7.60	1.72792
7.90	1.788941
8.20	1.760731
8.50	1.648487
8.80	1.464917
9.10	1.229098
9.40	0.964775
9.70	0.698239

Start R, and type: `z<-read.table("clipboard",header=T)`

The arrow comprised of `<` and `-` means “assign to”, and you should see the above data as soon as you type `z` at the prompt sign (`>`). (You can replace this “assign” statement with an equals sign (`=`) in most work in R). The statement “`header=T`” means that there is a header in the data (T stands for TRUE). If you don’t have or want a header, use “`header=F`” (actually you don’t need to do this, as the “default” condition is “`header=F`”). The program inserts headings of V1 and V2 over the data if you don’t use “`header=T`”. Now type “`plot(z)`”. You should get the following graph, with the exception that I have added a statement “`abline(h=0)`” in order to draw a horizontal line

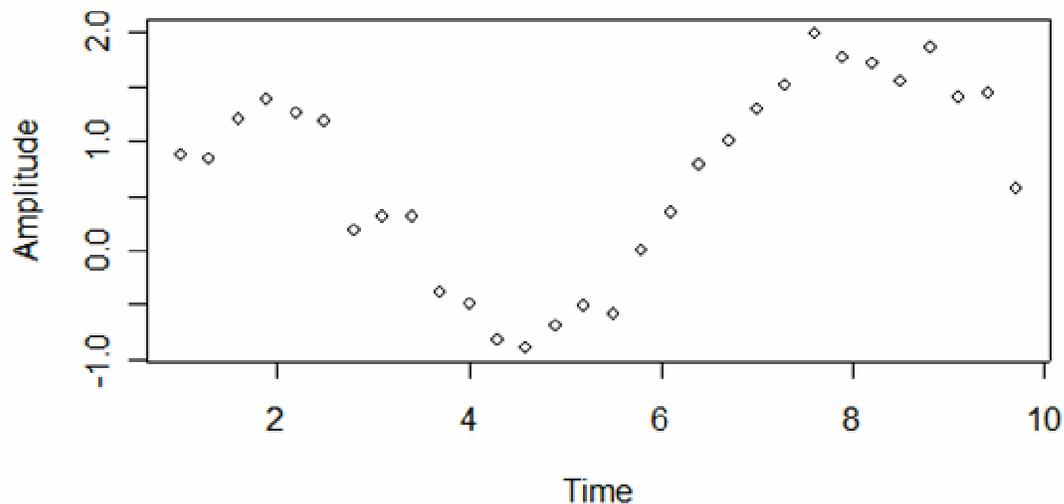
at zero (use “`abline(h=x)`” to draw a vertical line at point x (insert a number for x). We now have a plot of the basic sine curve.



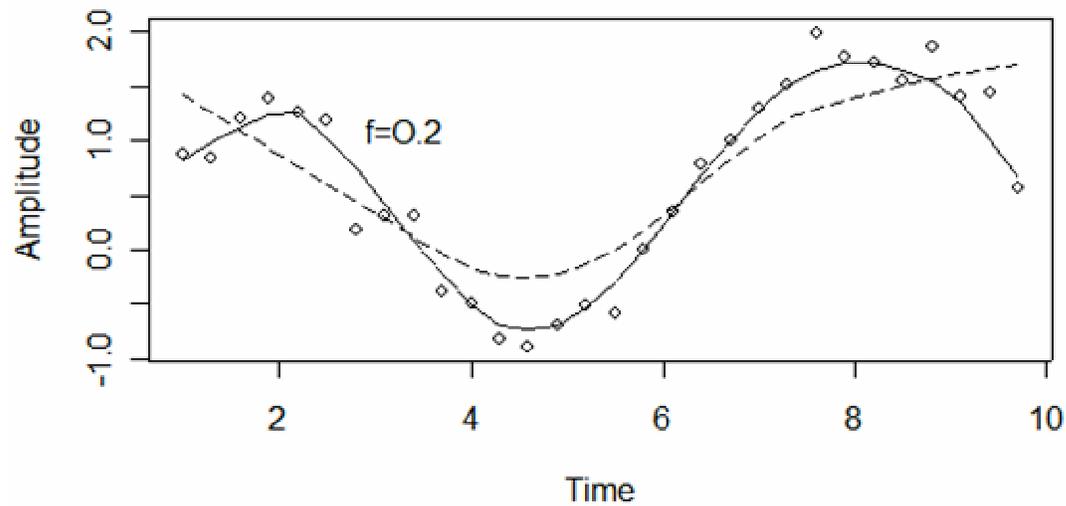
In order to introduce some variability into data to use with LOWESS, I added uniform random variables (but reduced by multiplying the tabulated (RND in EXCEL) value by 0.2). These new values are as follows:

Time	Amplitude
1.00	0.881425
1.30	0.838022
1.60	1.208425
1.90	1.391595
2.20	1.268166
2.50	1.195099
2.80	0.178271
3.10	0.304744
3.40	0.303463
3.70	-0.37718
4.00	-0.49484
4.30	-0.82425
4.60	-0.90307
4.90	-0.68798
5.20	-0.51816
5.50	-0.57913
5.80	0.001813
6.10	0.347028
6.40	0.78352
6.70	1.001751
7.00	1.291588
7.30	1.506389
7.60	1.996448
7.90	1.771884
8.20	1.723499
8.50	1.545846
8.80	1.85936
9.10	1.402233
9.40	1.439906
9.70	0.567258

Copy the data to the clipboard ,and use the command `“y<-read.table(“clipboard”,header=T)”`, and then the command `“plot(y)”`, and you should get the following plot.

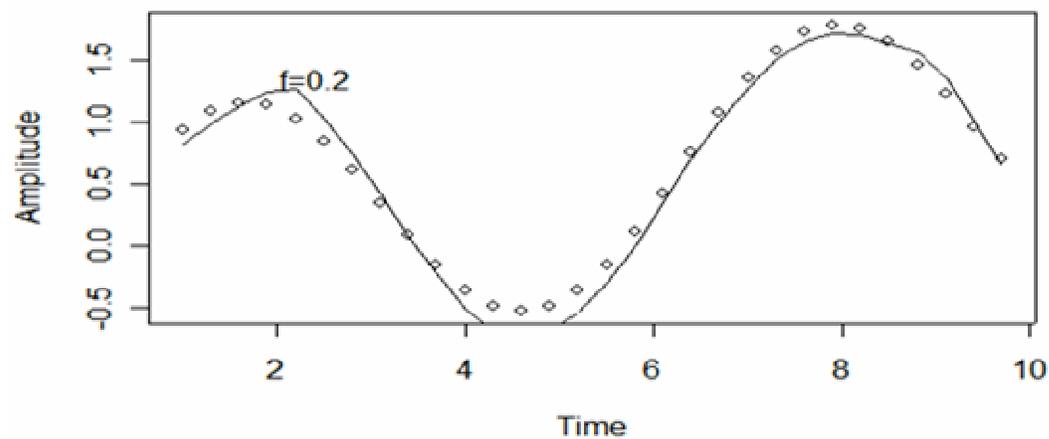


This is the sine curve with some added variability, and we now can try “lowess”. It is most conveniently introduced by way of a “lines” command, which draws the line generated by lowess on your plot. Type `“lines(lowess(y,f=.2),lty=1)”` and you should get the fitted lowess line. Put on another line `“lines(lowess(y,f=.6),lty=5)”` to get a second fitted line. The added `“lty=5”` makes the line dashed whereas the first line was solid (you don’t need to use `“lty=1”` as the program automatically draws a solid line if you leave off the `lty` command). There are a number of `lty` choices, which you can see by trying different numbers in the `lty` command. In order to identify lines on the plot, use the handy “locator” command. It places text on the plot. Type `“text(locator(1),”f=0.2”,adj=0)”` and hit return. Nothing happens until you move the cursor next to the solid line and again hit return. Then the text is printed on your plot at that location. The parentheses determine the number of locations – use (1) for a single introduction. Using `“adj=0”` makes the text start at the cursor location. The plot is as follows:



The broken line shows the results for $f=0.6$. Using $f=0.2$ seems to give a better fit, as it uses only points in the immediate neighborhood. Comparing the fitted line ($f=0.2$) with the original sine values gives the following plot:

LOWESS CURVE vs. ACTUAL



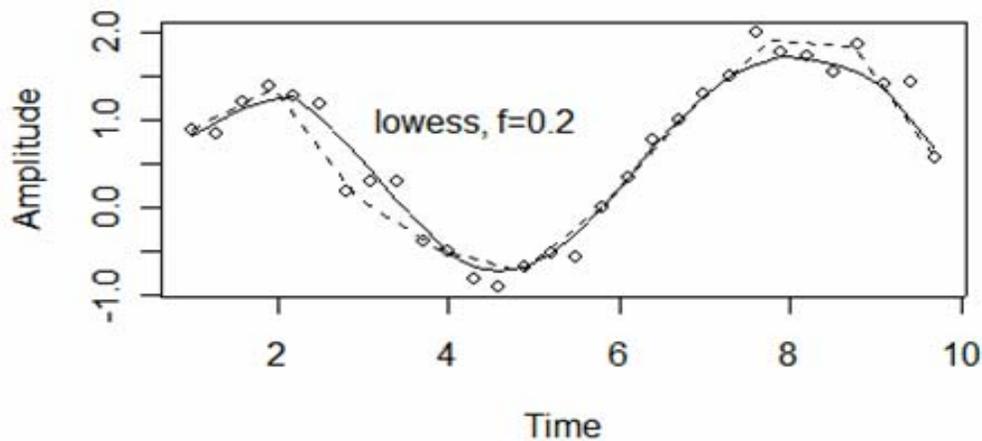
The title on this graph was produced by the statement
`“title(main=“LOWESS CURVE vs. ACTUAL”)` This is a somewhat better approximation than one might expect to get in many practical applications.

SPLINES

We can consider another “smoother” on the same data as used to illustrate the lowess method. As with lowess, it is conveniently executed with the “lines” statement. There are 3 versions, but we will illustrate with one and add it to the plot of lowess:

```
> lines(spline(z,n=10,method="natural"),lty=2)
```

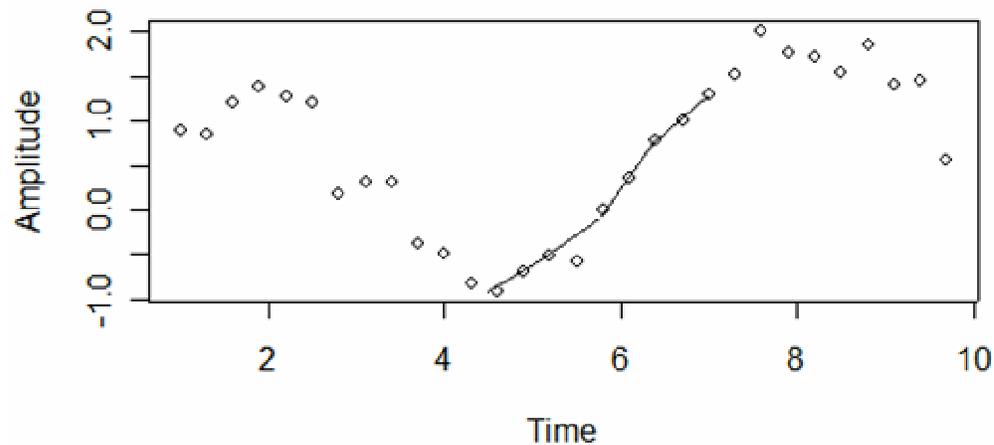
This gives the following plot:



where the broken line shows the “spline” fit from the above statement, and the solid line is the earlier fit of lowess to the same data. The value of n determines the number of points used in interpolation for the fit, i.e., 10 points are uniformly spread over the range of the data. Using more points results in line that follows the data points more closely. One can restrict the range of points fitted by a statement like the following:

```
> lines(spline(y,n=5,xmin=4.5,xmax=7,method="natural"))
```

This gives the following plot (after using “plot(y)” to get just the points on a new plot):



There are two additional “methods”. Use “periodic” or “fmm” to experiment with them. The latter fits a cubic to the first and last 4 points to produce the end conditions of the fit. Note that leaving off the “lines()” part of the statements used above results in the program printing out the points used for interpolation and the corresponding y-values. One can thus get interpolated values accurately, and can use this approach to get a better understanding of the methods. Using this on the plot above, i.e.

```
>spline(y,n=5,xmin=4.5,xmax=7,method="natural")
```

gives the 5 points used to approximate the line and the corresponding y-values:

```
$x
```

```
[1] 4.500 5.125 5.750 6.375 7.000
```

```
$y
```

```
[1] -0.92033436 -0.52407423 -0.09690946 0.75362089 1.29158847
```

NONLINEAR LEAST-SQUARES

Exponential curve

Using nonlinear least-squares was briefly discussed in Section 1.13, and a method for simulating the exponential distribution appears in exercise 1.14.6. To demonstrate fitting by nonlinear least-squares we use the

exponential, and the data on muskoxen from Fig.9.3. (Use “rm(list=ls())” to clear the previous work from the console). The model is:

$$y = ae^{bx}$$

where a and b are the parameters to be estimated. While the model looks simple, one has to have fairly good initial estimates of the parameters to avoid non-convergence in the fitting process. These are easy to get as one can take logarithms of the values and use linear regression to get initial parameter estimates for the fitting (but remember to transform the intercept (a) from the linear fit by taking exp(a) to get the initial value to use for that parameter in the nonlinear program).

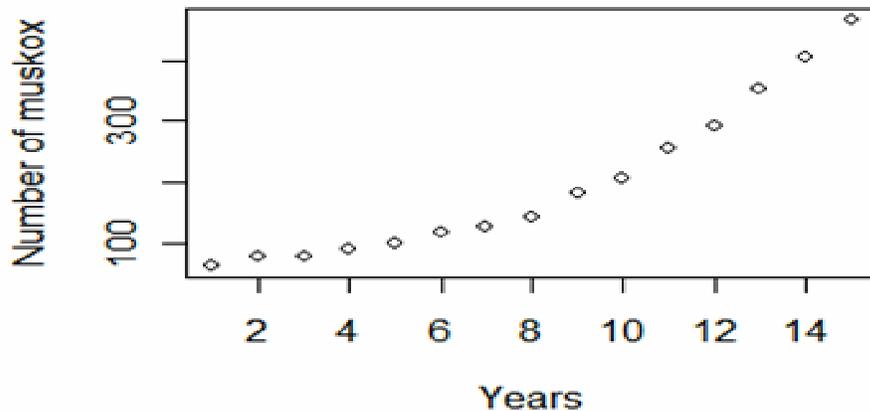
Thus far, we have used data in matrix form (i.e., we read two columns of data into a single “object”, z, and worked with z. Now we need to bring in single columns to correspond to x and y above. The data are as follows:

YEAR	NUMBER
1	61
2	76
3	77
4	90
5	100
6	116
7	126
8	143
9	181
10	206
11	256
12	293
13	353
14	406
15	467

Read in x and y separately, using the “read.table” statement. There is a problem here in that one has to use statements: x=as.matrix(x), and y=as.matrix(y) before the data will plot. Whenever, one reads in data it is a good idea to plot it before trying anything else. That way, it is possible to identify difficulties with the “read” statement early on. Now try:

```
> plot(x,y,ylab="Number of muskox",xlab="Year")
```

to get a plot of the data:



To get the fit one can use the following statement:

```
fn<-function(p)sum((y-(p[1]*exp(p[2]*x)))^2
```

There are several things to note here. One is that there is no punctuation between “function(p)” and “sum”. This is unusual in R, and may thus be confusing. Also the parameters are identified by numerals in SQUARE brackets. Be sure to get the parentheses right, or you may get some strange results! Most R commands won’t work unless things are “right”, but these are an exception.

```
Now type >out<-nlm(fn,p=c(50,.2))
```

“nlm” is the program that minimizes a function (fn), and the trial parameters are included in the vector “ c(50,.2)”. Type “>out” to get the results:

```
>out
$minimum
[1] 0.0002419648
$estimate
[1] 49.9999575 0.1469777
$gradient
[1] -1.66196345 0.09648096
$code
[1] 2
$iterations
[1] 12
```

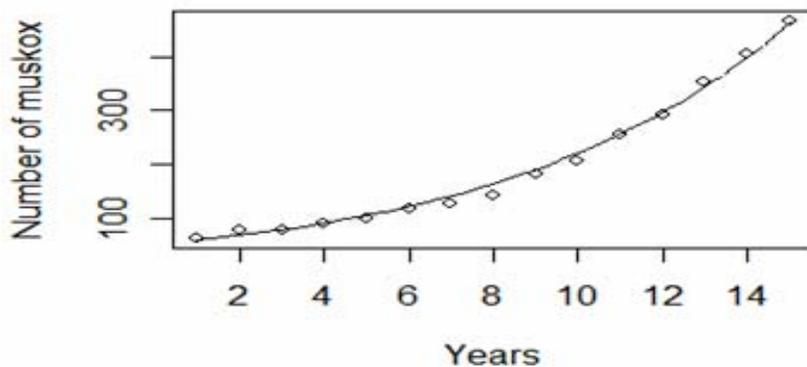
The first entry is the minimum sum of squares (note that `fn` is set up to calculate a sum of squared deviations), the second gives parameter estimates, and the third gives slopes at the parameter estimates, the fourth is the number of parameters, and the final entry (12) gives the number of iterations used. As noted above, you need reasonably good initial estimates, or you may get some strange results. You can check your results by typing:

```
>xfit<-seq(1:15)
>yfit<-50*exp(.14698*xfit)
```

The sequence in `xfit` is any convenient sequence of points that covers the range of the data. In this case it is exact, but in others it may not be; the point is to select a sequence of x-values for plotting, while `yfit` produces a plot using the parameter estimates over that range. Then use the “splines” function:

```
>lines(spline(xfit,yfit))
```

to get the following plot:



It is often desirable to use the `xfit` and `yfit` statements along with the `lines(spline(xfit,yfit))` to try out some trial parameters before beginning the fitting process, just to be sure that you have reasonable starting parameters.

Note that the nonlinear model (`nlm`) uses the Newton-Raphson approach and is, like all nonlinear fitting procedures, not always guaranteed to find the “best” result. It is used here to demonstrate the function statement:

```
fn<-function(p)sum(y-(p[1]*exp(p[2]*x)))^2
```

which is useful in programming in R. For nonlinear fitting purposes, it seems preferable to use “nls” as follows:

```
out=nls(y~a*exp(b*x),start=list(a=50,b=.2))
```

because this lets one write the equation directly in a simple form, and also because more detail is obtained. Use

```
summary(out)
```

```
Formula: y ~ a * exp(b * x)
```

```
Parameters:
```

```
Estimate Std. Error t value Pr(>|t|)
a 45.217910  1.804085  25.06 2.17e-12 ***
b 0.156102  0.003134  49.81 3.15e-16 ***
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 8.038 on 13 degrees of freedom
```

```
Correlation of Parameter Estimates:
```

```
      a
b -0.9748
```

There is a lot more information available: Type “fitted(out)” to get:

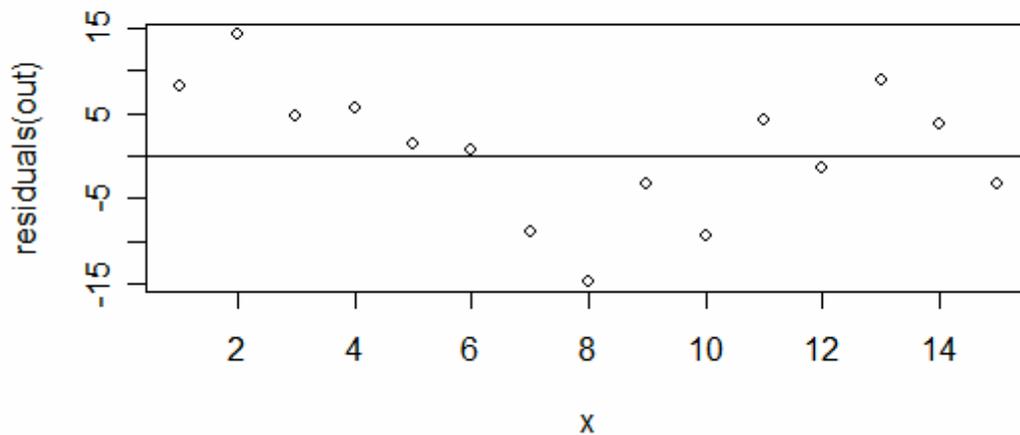
```
[1] 52.85725 61.78721 72.22584 84.42803 98.69171 115.36518 134.85553
 [8] 157.63869 184.27094 215.40258 251.79375 294.33301 344.05908 402.18610
[15] 470.13340
attr("label")
[1] "Fitted values"
```

These are values of the curve fitted at the x-values.

Also, “residuals(out)” gives:

```
residuals(out)
 [1]  8.1427546 14.2127924  4.7741587  5.5719711  1.3082861  0.6348242
 [7] -8.8555341 -14.6386890 -3.2709418 -9.4025779  4.2062548 -1.3330147
[13]  8.9409232  3.8138963 -3.1333954
attr("label")
[1] "Residuals"
```

Use `plot(x,residuals(out))` and `abline(h=0)` to plot the residuals as in the graph below. This is valuable in that it suggests that the observations may be “oscillating” around the trend line, rather than being “random errors”.



Note, too, that the parameter estimates from “nls” differ from those of “nlm”, so that the two fitting methods do not give exactly the same results.

Logistic function

Another example, the logistic function, has 3 parameters. The function is:

$$y = \frac{K}{1 + be^{-rx}}$$

where K is the asymptotic value, b is determined by the initial population size, and r is the rate of increase. Examining a plot of the data (Fig. 9.2) suggests the asymptote might be around 2200. However, the last 5 observations are rather erratic so are dropped here, and the data used are given on the next page. We can regress the logarithms of the first dozen observations on time and get a rough notion of r (the slope of a regression on the log scale) as about 0.3 (using just the first 12 counts).

```
> z1<-read.table("clipboard",header=T)
> lm.mod<-lm(log(y)~x,z1)
> lm.mod
Call:
lm(formula = log(y) ~ x, data = z1)
Coefficients:
(Intercept)      x
  2.6214      0.3039
```

Note that this brings in something new, a regression model (lm.mod), but using $\log(y)$ instead of y , and giving the data (z1) as part of the “command line”: `lm.mod<-lm(log(y)~x,z1)`

x	y
1	12
2	23
3	32
4	60
5	75
6	100
7	143
8	188
9	247
10	301
11	310
12	367
13	395
14	480
15	605
16	690
17	814
18	994
19	1173
20	1353
21	1225
22	1469
23	1566
24	1685
25	1701
26	1742

Copy y from the table above, and set `x=seq(1,26)`. Be sure to use `rm(list=ls())` to clear the R-console before starting, and remember to convert y by `y=as.matrix(y)`. Check by plotting x,y .

An estimate of b can be obtained from setting $K=2200$, $x=0,y=12$ (initial value) in the model:

$$y = \frac{K}{1 + be^{-rx}}$$

This gives:

```
out=nls(y~k/(1+b*exp(-r*x)),start=list(k=2200,b=180,r=.3))
```

and:

```
summary(out)
```

```
Formula: y ~ k/(1 + b * exp(-r * x))
```

```
Parameters:
```

```
Estimate Std. Error t value Pr(>|t|)
k 2.056e+03  8.739e+01  23.529 < 2e-16 ***
b 8.698e+01  1.493e+01   5.827 6.15e-06 ***
r 2.432e-01  1.384e-02  17.575 7.93e-15 ***
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 50.06 on 23 degrees of freedom
```

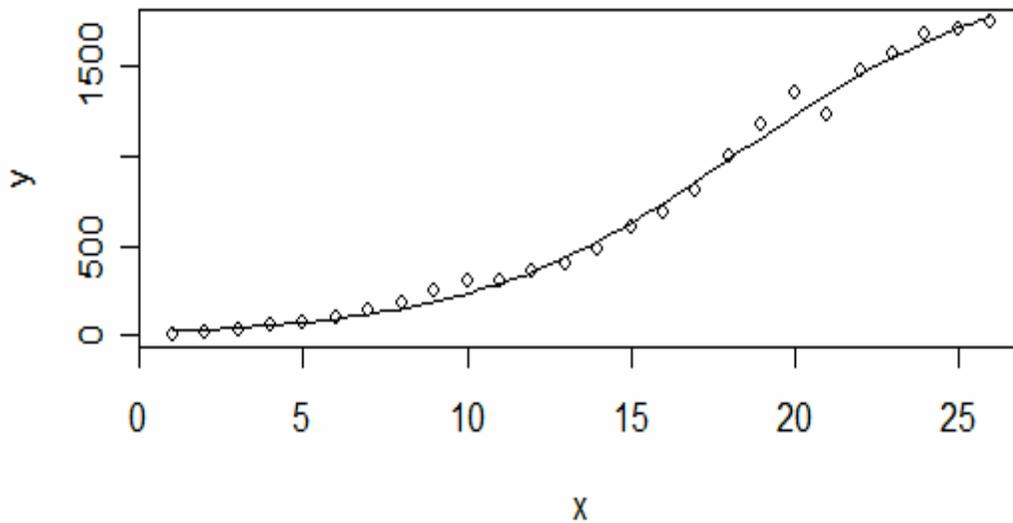
```
Correlation of Parameter Estimates:
```

```
      k      b
b -0.6385
r -0.8552 0.9345
```

```
Set yfit=2056/(1+87*exp(-.243*x))
```

```
And use lines(spline(x,yfit))
```

```
to get the fitted curve:
```



Michaelis-Menton curve

For another fit, consider the equation:

$$y = \frac{ax}{b + x}$$

and the data of Fig. 1.15, loaded with the vectors:

```
>x<-c(1,2,3,4,5,6,11,23)
>y<-c(.217,.462,.5,.6,.692,.4,.824,1)
```

Note that you can plot x,y directly without the conversions involved in getting data in from the clipboard.

Trying a=1 in the equation suggests b=4 might give an approximation, and we plot the data and try these parameters with:

```
> xfit=seq(1:25)
> yfit=xfit/(4+xfit)
> lines(spline(xfit,yfit))
```

which gives the solid line on the plot below. Then enter the function:

```
out=nlm(y~a*x/(b+x),start=list(a=1,b=4))
```

```
summary(out)
```

```
Formula: y ~ a * x/(b + x)
```

```
Parameters:
```

```
Estimate Std. Error t value Pr(>|t|)
a  1.1041    0.1758  6.280 0.000758 ***
b  3.9446    1.6461  2.396 0.053568 .
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.1229 on 6 degrees of freedom
```

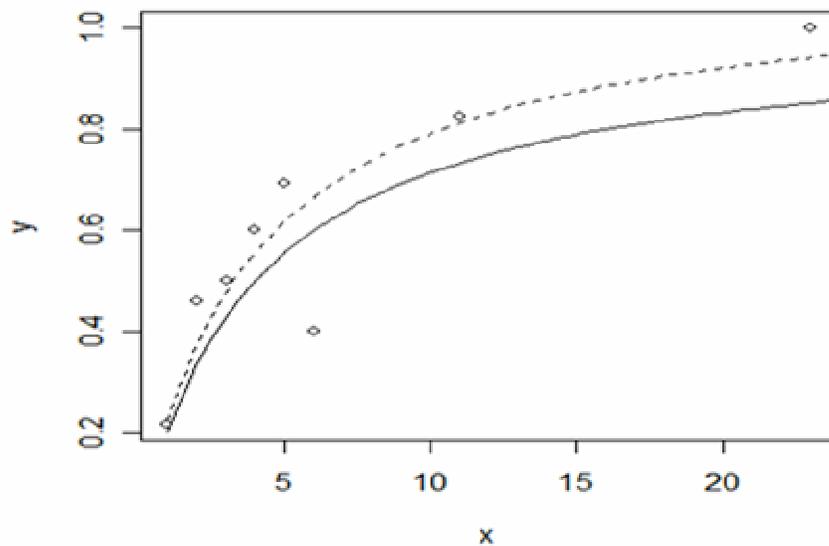
```
Correlation of Parameter Estimates:
```

```
      a
b 0.8988
```

Now plot the data with:

```
> yfit=1.104*xfit/(3.9446+xfit)
> lines(spline(xfit,yfit),lty=2)
```

and we get:



where the broken line now is the fit by nonlinear least-squares, and the solid line comes from the “guesstimates” used to run nls. This fit can be compared to Fig. 1.15.

Logistic regression

Another fit uses the model:

$$P = \frac{1}{1 + e^{-(a+bx)}}$$

discussed in Section 1.11, and originally used as a bioassay model. This is the logistic function used earlier here on the elephant seal data, but with the asymptotic parameter set equal to 1. The same data are used as above for the Michaelis-Menton model. An easy way to get trial values is to fit eq.(1.31) by linear regression to get estimates of a and b. One can go to EXCEL for the calculations, but instead we use the linear model in R. The data are given

in the last column of Table 1.9 and are shown again below (note that the last observation has to be dropped to avoid a calculation involving infinity).

Make new vectors using the data of the last column of Table 1.9:

```
>Y=c(-1.281,-.154,0,.405,.811,-.405,1.54)
```

```
>X=c(1,2,3,4,5,6,11)
```

Note that the fact that R-language is “case-sensitive” lets us use cap Y and X as new variables, without disturbing the full data set on the R-console as x and y from fitting the Michalis-Menton curve.

The linear regression model is:

```
>lm.mod<-lm(Y~X)
```

```
> lm.mod
```

Call:

```
lm(formula = Y ~ X)
```

Coefficients:

```
(Intercept)      X
   -0.8715      0.2193
```

and we use these values in:

```
out=nls(y~1/(1+exp(-(a+b*x))),start=list(a=-.87,b=.2))
```

to get:

```
summary(out)
```

```
Formula: y ~ 1/(1 + exp(-(a + b * x)))
```

Parameters:

```
Estimate Std. Error t value Pr(>|t|)
a -0.80842  0.39391 -2.052  0.0860 .
b  0.21003  0.08397  2.501  0.0464 *
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.1259 on 6 degrees of freedom
```

Correlation of Parameter Estimates:

```
      a
b -0.8527
```

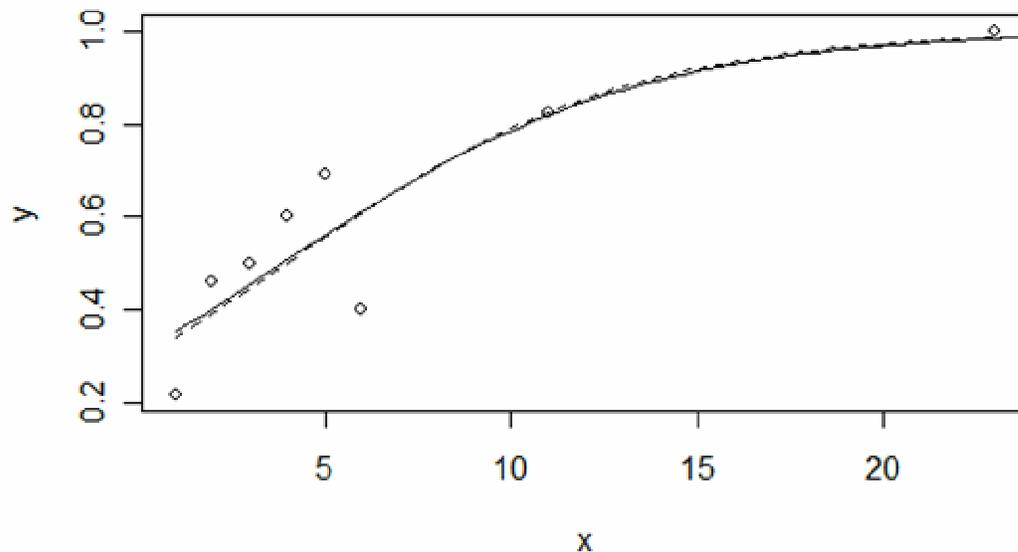
and plot the output with:

```
> plot(x,y)
```

```
> xfit<-seq(1:25)
```

```
> yfit<-1/(1+exp(-(-.8084+.21*xfit)))
```

```
> lines(spline(xfit,yfit))
```



The broken line shows the plot using the parameters from the regression fit, which are nearly identical to the least-squares fit. Note the large standard errors in the nls fit.

Monomolecular curve

This curve has the form (eq.(12.8)):

$$y = a(1 - e^{-bt})$$

Using the same data as for the Michaelis-Menton curve:

```
>x<-c(1,2,3,4,5,6,11,23)
```

```
> y<-c(.217,.462,.5,.6,.692,.4,.824,1)
```

Clearly the asymptotic value will be roughly 1.0, and expanding the exponential term as in Section 12.3 gives $y=abx$, so one can estimate b as about 0.1. We can then get a trial plot from:

```
> xfit<-seq(1:25)
```

```
> yfit<-(1-exp(-.1*xfit))
```

and plotting x and y , plot the trial curve with:

```
> lines(spline(xfit,yfit))
```

This gives the solid line on the plot below.

Using nls:

```
out=nls(y~a*(1-exp(-b*x)),start=list(a=1,b=.1))
```

```
summary(out)
```

```
Formula: y ~ a * (1 - exp(-b * x))
```

```
Parameters:
```

```
Estimate Std. Error t value Pr(>|t|)
a 0.92554  0.12398  7.465 0.000298 ***
b 0.22419  0.06967  3.218 0.018185 *
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.1382 on 6 degrees of freedom
```

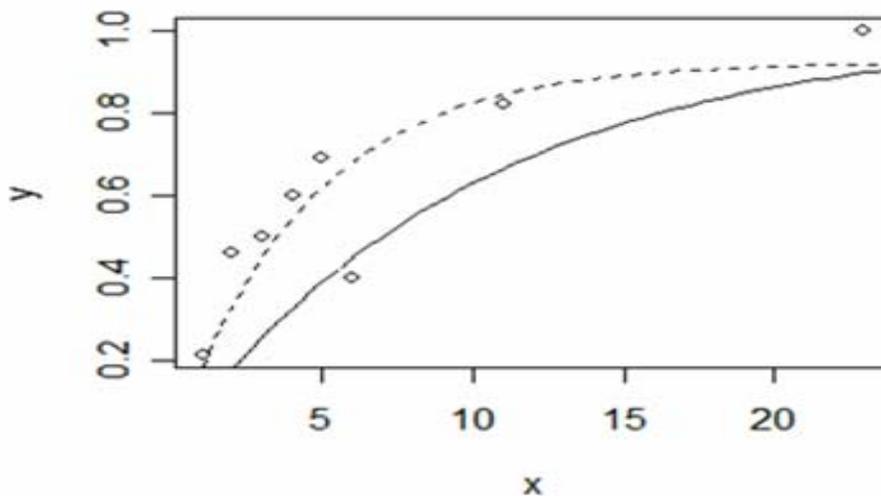
```
Correlation of Parameter Estimates:
```

```
      a
b -0.8086
```

and we can then plot the resulting curve with:

```
> yfit<-.9255*(1-exp(-.2242*xfit))
> lines(spline(xfit,yfit),lty=2)
```

with the results as follows:



where the solid line shows the plot with trial values and the broken line the fit from “nls”.

Another exponential-type curve

We can add yet a 4th curve:

$$y = ae^{-b/x}$$

taking logarithms, we can again construct new variables for initial parameter estimates:

```
Y<-log(y)
```

```
X<-1/x
```

```
> fit<-lm(Y~X)
```

```
> fit
```

Call:

```
lm(formula = Y ~ X)
```

Coefficients:

```
(Intercept)      X
   -0.1831     -1.3608
```

converting the intercept estimates a as $\exp(-.1831)=0.8327$ and we insert these parameter estimates:

```
out=nls(y~a*exp(-b/x),start=list(a=.8,b=1.4))
```

```
summary(out)
```

Formula: $y \sim a * \exp(-b/x)$

Parameters:

	Estimate	Std. Error	t value	Pr(> t)
a	0.9279	0.1183	7.844	0.000227 ***
b	1.7551	0.5731	3.063	0.022148 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1397 on 6 degrees of freedom

Correlation of Parameter Estimates:

	a
b	0.7802

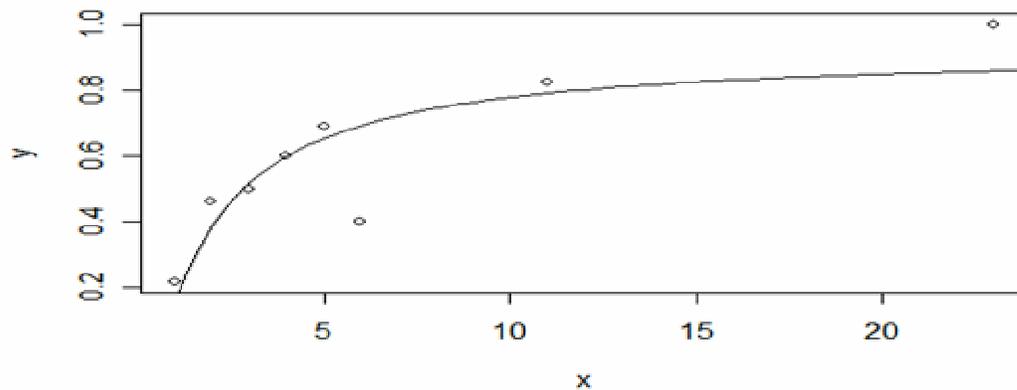
and plotting with:

```
>xfit<-seq(1:25)
```

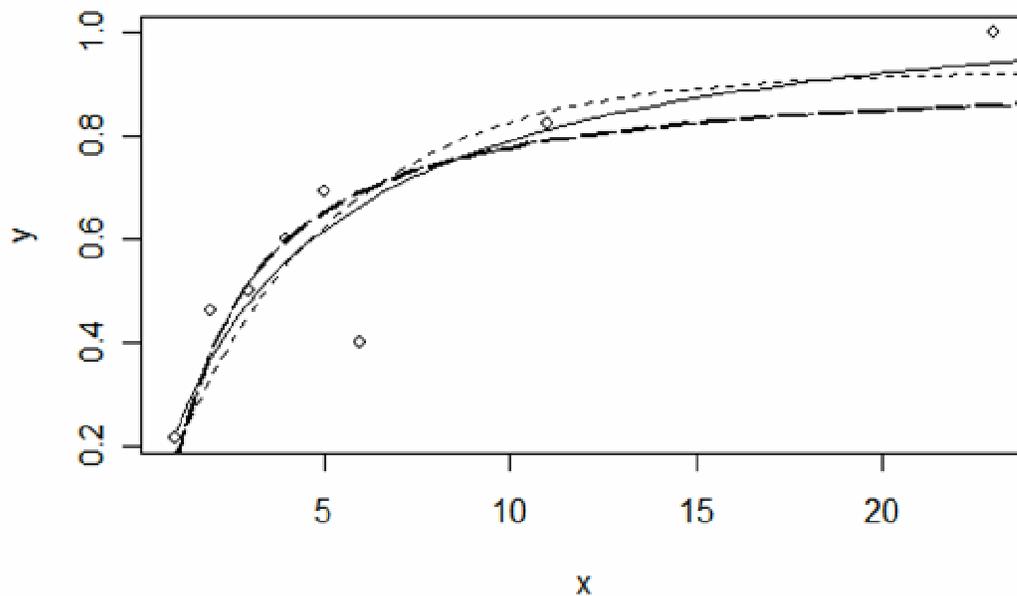
```
> yfit<-0.9279*exp(-1.7551/xfit)
```

```
> lines(spline(xfit,yfit))
```

gives:



Note that three of the lines evaluated above are very similar, the Michaelis-Menton (solid line below), monomolecular (broken line), and the last, exponential-type curve (heavy broken line). Hence, it is useful to have more data if one wants to try to find the “best” curve.



As was noted earlier, one can often use an equals sign (=), instead of the “assignment operator” (<-). This substitution seems to work in most cases, but some references insist on the assignment operator, as used in the S-language.

MAXIMUM LIKELIHOOD ESTIMATORS

Normal distribution

The normal distribution is one of the most commonly used frequency distributions, and estimates of its two parameters are usually derived in textbooks that discuss maximum likelihood estimators. It is assumed that the observed random variables, x_i , are independent, and the process starts out with the product of frequency distributions of n such random variables:

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} e^{-\frac{\sum (x_i - \mu)^2}{2\sigma^2}}$$

and logarithms are taken to obtain the log-likelihood.

$$L = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2$$

This equation is then differentiated with respect to σ^2 and μ , the resulting equations set equal to zero and solved for the parameters σ^2 and μ , which locates the maximum likelihood values. We can equivalently take the negative value of L , and find a minimum, which is what is done with the R-language program `nlm`. Some artificial data were generated to demonstrate the program, using “`rnorm`”, which generates normal random variables with zero means and unit standard deviations. The statements used and output follow (copy the list of values given below for `x` in order to get the same results as given below; generate a new set from `e<-rnorm(20)` to experiment further if you want to change the outcomes):

```
> e<-rnorm(20)
> x<-8+4*e
```

```

      x
9.2934
8.4137
 13.21
12.595
11.778
10.026
 1.6196
 7.3556
 5.7192
10.437
 1.4914
12.316
 3.1239
 5.4366
11.235
 7.0967
 9.0129
10.635
 8.2053
 6.5455

```

The negative log-likelihood is inserted in R (the first term in L does not contain either parameter so has no effect on the results and is thus not included, similarly the magnitude of n does not affect the location of the minimum). We now use nlm to find a minimum:

```
>fn<-function(theta){sum(.5*(x-theta[1])^2/theta[2]+.5*log(theta[2]))}
```

and we insert guesses at the parameters in:

```
>out<-nlm(fn,theta<-c(8,12))
```

which gives the output;

```

$minimum
[1] 34.50804
$estimate
[1] 8.277286 11.597658
$gradient
[1] -1.321974e-07 -1.286587e-08

```

```

$code
[1] 1
$iterations
[1] 7

```

The parameter estimates can be checked with:

```

mean(x)
  V1
8.27729
sum(((x-mean(x))^2)/20)
[1] 11.59766

```

This is because we already know that μ is estimated by the mean of the x_i , and σ^2 is estimated by: $\sum (x_i - \bar{x})^2 / n$ (note that the maximum likelihood estimator of the variance is biased and is usually corrected by using $n-1$ rather than n in the denominator).

Exponential distribution

The exponential distribution is a one-parameter function, and, like the normal distribution, parameter estimates can be obtained directly by differentiating the log-likelihood function. Estimates are obtained from the following:

$$\prod_{i=1}^n \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum x_i}$$

the log-likelihood is:

$$L = n \log \theta - \theta \sum x_i$$

differentiating and setting the result equal to zero gives:

$$\frac{\partial L}{\partial \theta} = \frac{n}{\theta} - \sum x_i = 0$$

the estimate is then:

$$\hat{\theta} = \frac{n}{\sum x_i}$$

For an example we use 30 observations generated using the RAND() function in EXCEL, transforming the uniform random variables by taking

the negative logarithm and dividing by 3 (the parameter of the simulated exponential distribution).

```
x
0.720755
0.035991
1.063578
0.444527
0.342697
0.868936
0.117598
0.143798
0.012229
0.27057
0.118674
0.297912
0.495389
0.789919
0.224876
0.230535
0.406611
0.634387
0.239877
0.337368
0.123611
0.447395
0.567364
0.171177
0.036534
0.120707
0.189452
0.322679
0.227388
0.021487
```

Write the function as:

```
>fn<-function(theta){theta*sum(x)-30*log(theta)}
```

and try nlm (inserting a trial value as 2):

```
> out<-nlm(fn,p=c(2))
```

```
> out
```

```
$minimum
```

```
[1] -2.886392
```

```
$estimate
```

```
[1] 2.992809
```

```
$gradient
```

```
[1] -2.409779e-07
```

```
$code
```

```
[1] 1
```

```
$iterations
```

```
[1] 6
```

The direct estimate is that given by the maximum likelihood estimator derived above with the sum of the x 's obtained from the 30 "observations":

$$\hat{\theta} = \frac{30}{10.02402} = 2.99$$

The above two examples illustrate the basis for maximum likelihood estimation. These examples are constructed from the actual probability models (normal and exponential distributions). Many of the procedures used in statistical analysis assume that the "error structure" is normally distributed, but this is not necessarily the case, and going further with maximum likelihood methods involves linking models with various types of error structure, e.g., Poisson, binomial, gamma, etc.

BOOTSTRAPPING

A ratio estimator

Bootstrapping is available in the R-language, and offers some advantages over the "macros" available in Appendix A. As with the R-language in general, there are various difficulties in learning the methods, so the presentation here has been kept simple. Once users become familiar with the approach, it should become feasible to handle more complex examples. The "boot" package may have to be unloaded from the internet (if that was not done in the original installation – check in the pull-down menu under "packages" for "boot"). If necessary, activate your internet access, go to CRAN and download the available packages. When you begin a new session in R, it is again necessary to pull down the "packages" menu and hit "boot", as the package does not stay loaded in the R "environment" between uses, as do the other features used thus far. If you forget, R will return a message saying "boot not located", and you can just pull down the "packages" menu and load it.

Type `?boot` to get a description of the package. As with all of the "help" features, this is likely to be confusing to the non-initiate. However, at the end there is a simple example which is used here.

Type “data(city)” and then “city” which prints out the trial data:

```
> city
  u x
1 138 143
2  93 104
3  61  69
4 179 260
5  48  75
6  37  63
7  29  50
8  23  48
9  30 111
10 2  50
```

Now type in:

```
> ratio<-function(d,w){sum(d$x*w)/sum(d$u*w)}
```

This is the function to be bootstrapped, which is the ratio of the sums of the two columns.

The actual bootstrapping is accomplished with:

```
>result<-boot(city,ratio,R=5000,stype="w")
> result
ORDINARY NONPARAMETRIC BOOTSTRAP
Call:
boot(data = city, statistic = ratio, R = 5000, stype = "w")
Bootstrap Statistics :
  original    bias  std. error
t1* 1.520313 0.04035454  0.2213885
```

Because bootstrapping is a sampling process, results will vary a little between trials. The first item (the “statistic” given in “?boot”) is just the ratio of the sums of the original data, as one can confirm by adding up the two columns and finding the ratio of the sums. The second item is the bias, defined in eq.(3.2), and is the mean of the bootstrap samples minus the value of the original “statistic”, i.e., the ratio of sums of the original data. The term labeled “std error” is the square root of the variance of the bootstrap sample (eq.(2.1)). The “w” is inserted to permit use of weights which was not done

here (weights are automatically assumed equal by the program if they are not specified in the “call” statement).

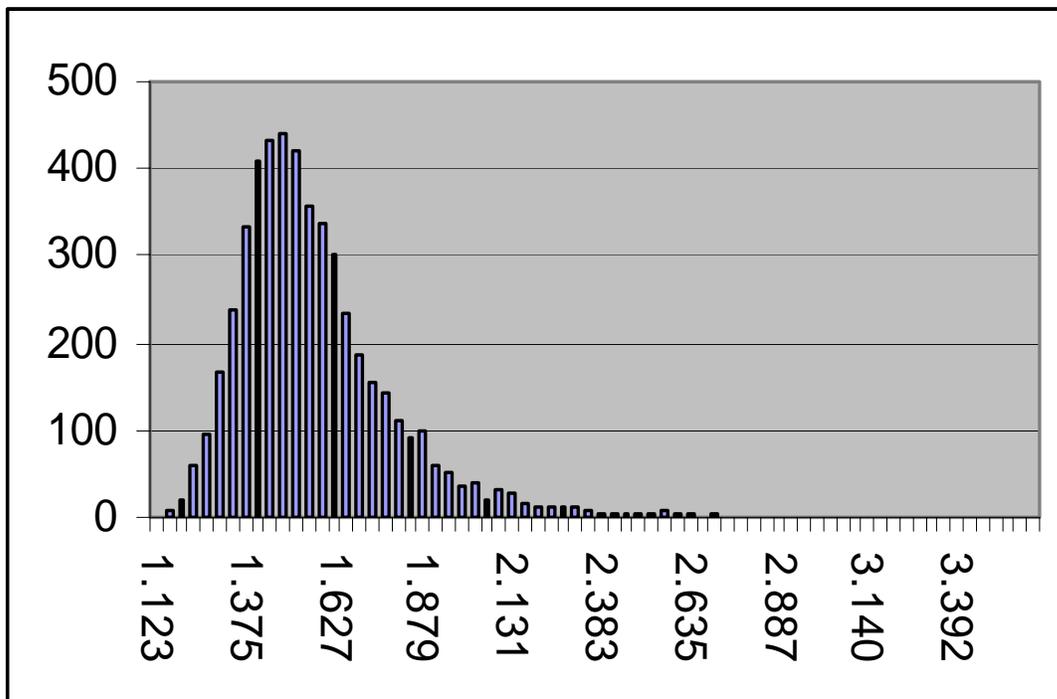
For comparison it is useful to run BOOT2 with the same data and 5,000 bootstraps. It gave:

```

1.520313 Estimate
1.549695 Bootstrap mean
0.029383 Bias
  0.04692 Bootstrap variance
0.216611 Standard error

```

The bias appears smaller, but another run gave a bias of 0.0436, in line with the values from the R-language program. A frequency diagram of the BOOT2 data is given below and shows how skewed the results are. The 95% percentile confidence interval from BOOT2 was 1.247 to 2.091.



Non-parametric regression

We can illustrate nonparametric regression with the grizzly bear data used to illustrate BOOT2 in Appendix A. The data are:

x	y
1	2.8332
2	2.5649
3	2.1972
4	2.5649
5	2.4849
6	2.6391
7	2.3979
8	2.5649
9	2.8332
10	2.1972
11	3.2189
12	2.5649
13	2.9444
14	2.7726
15	3.2189
16	3.1781
17	3.1355
18	2.9957
19	2.9957

Read in the data with

```
>beardat<-read.table("clipboard",header=T)
```

Then calculate the regression with:

```
> lm.mod<-lm(y~x,data=beardat)
```

and print out the results with:

```
> lm.mod
```

Call:

```
lm(formula = y ~ x, data = beardat)
```

Coefficients:

```
(Intercept)      x
  2.37756    0.03752
```

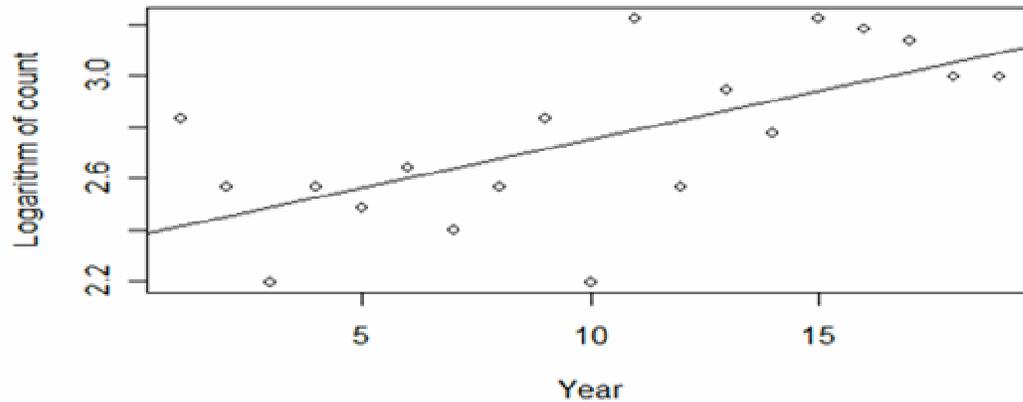
plot the data with:

```
> plot(beardat,xlab="Year",ylab="Logarithm of count")
```

and add the regression line with:

```
>abline(2.3776,.03752)
```

We now have a plot of the data as follows:



So far, this is just ordinary regression. To accomplish nonparametric regression by using bootstrapping, we need to use a simple program which follows:

```
linreg.1<-function(data,indices)
{tmp.dat=data[indices,]
z=lm(y~x,data=tmp.dat)
coefficients(z)}
```

Here we have another function, but it takes several lines to include the details, and these are enclosed in “curly brackets”, which are useful to delineate entries. This is the first case where it has been necessary to use several lines in a program, and one needs to copy the entire program and enter it as a unit.

Now type linreg.1 and you should see:

```
> linreg.1
function(data,indices)
{tmp.dat=data[indices,]
z=lm(y~x,data=tmp.dat)
coefficients(z)}
```

This only tells you that the program has been loaded properly. No computing is done yet.

The data are stored as “tmp.dat” for the bootstraps which are accomplished by:

```
>z=lm(y~x,data=tmp.dat), and the last line of the little program above stores the coefficients.
```

Be sure to load the “boot” package from the “packages” menu.

The actual bootstrapping is accomplished by entering the following statement. Note that there is a perceptible pause before the prompt symbol (>) comes up. This means that the program is running. In most of the examples so far, the time to run the programs has been short enough that one might not notice.

```
> out<-boot(linreg.1,data=beardat,R=1000)
```

print out the results with “out”:

```
> out
ORDINARY NONPARAMETRIC BOOTSTRAP
Call:
boot(data = beardat, statistic = linreg.1, R = 1000)
Bootstrap Statistics :
  original    bias  std. error
t1* 2.37755965 -0.0091140901 0.118493057
t2* 0.03751825  0.0006573894 0.009504197
```

t1* is the intercept, t2* is the slope, and these are the same as the regression calculation given earlier here. Notice that your estimates of bias and standard error differ somewhat from those given here. Try another run with:

```
> out<-boot(linreg.1,data=beardat,R=5000)
```

and notice how long it takes for the prompt (>) symbol to come up. A run with R=5000 gave the following results:

ORDINARY NONPARAMETRIC BOOTSTRAP

Call:

```
boot(data = beardat, statistic = linreg.1, R = 5000)
```

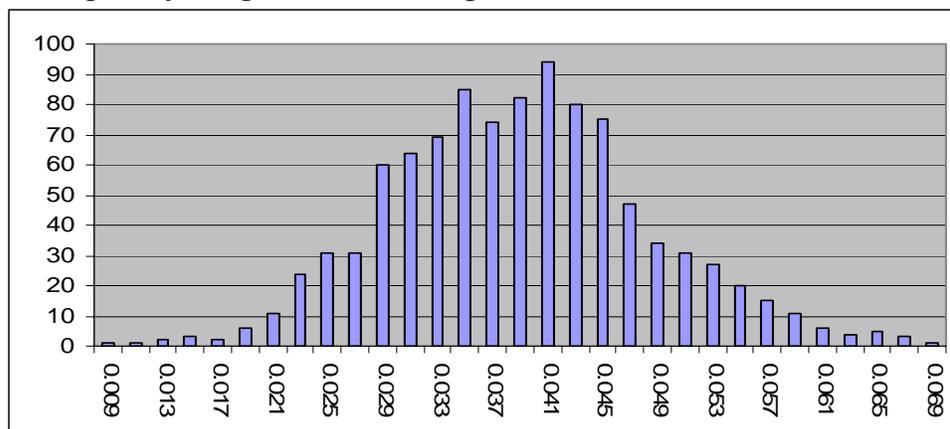
Bootstrap Statistics :

	original	bias	std. error
t1*	2.37755965	-0.0107787913	0.121005448
t2*	0.03751825	0.0006637228	0.009579154

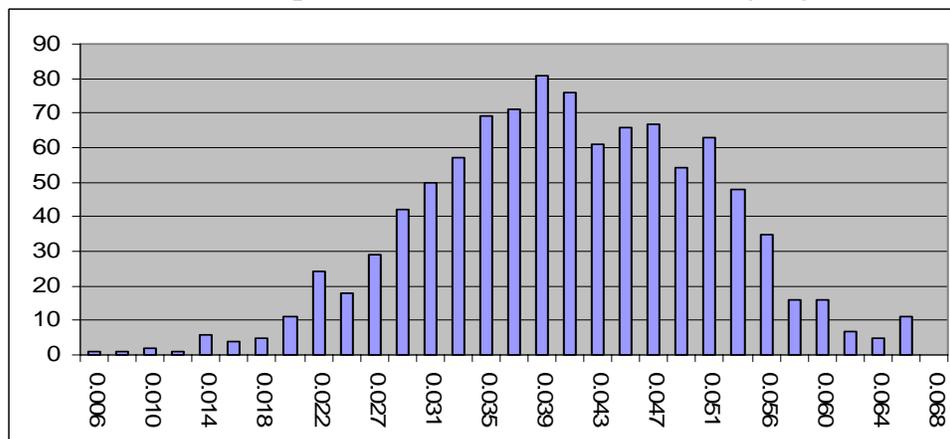
Running BOOT2 with the bear data and calculating slopes gave the following results:

BOOT MEAN	0.039878
VARIANCE	0.0001135
SE	0.0106525
BIAS	0.002359

A frequency diagram of the output of BOOT2 is:



and this can be compared with one from the R-language bootstrap program:



These plots are of the slopes of the regression lines (remember the log-transform) which estimate rate of change of the bear population. The intercepts are of very little interest. The two programs (BOOT2 and the R version) thus give similar results, as they should (but the data are quite variable).

If you type:

```
>names(out)
```

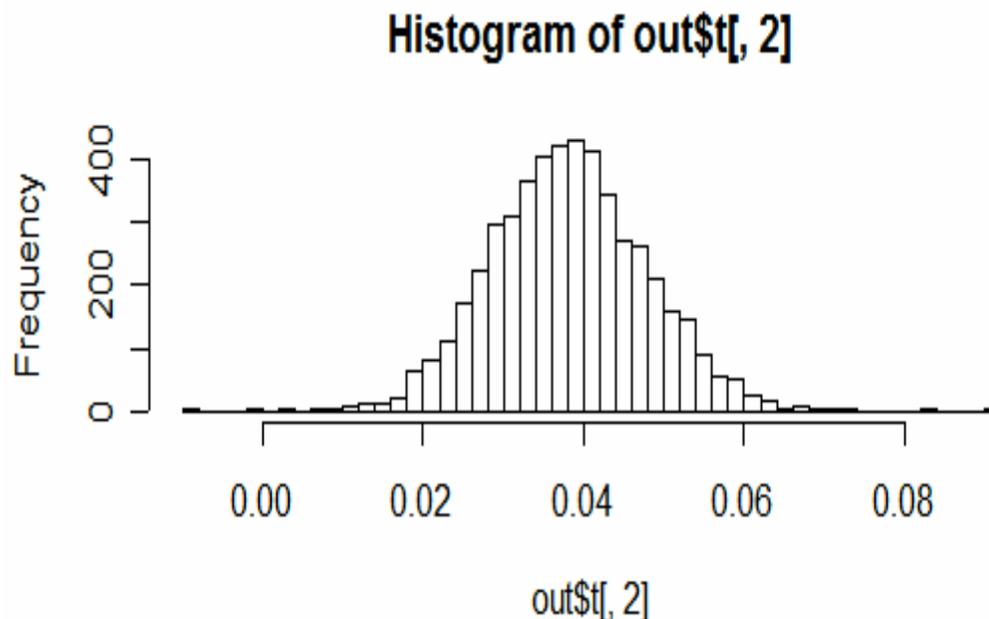
the following list comes up.

```
[1] "t0"    "t"     "R"     "data"  "seed"  "statistic"
[7] "sim"   "call"  "stype" "strata" "weights"
```

>out\$R produces the number of bootstraps used:

```
[1] 1000
```

>out\$t yields the bootstrap results, while `hist(out$t[,2],breaks=40)` produces a histogram of the slope values from the latest run (“breaks=40” determines the number of bars in the histogram).



This plot was produced from a later “run” of the BOOT program with $R = 5000$, so looks a little different from the one given earlier.

SOME USEFUL STATEMENTS IN THE R-LANGUAGE

It is advisable to store statements in Microsoft Word (or any other word-processing program), and copy them with the clipboard to load into the R-workspace. That way you can go back and change statements if they don't work properly. Similarly, if something "works" in R, you can copy it from the R-workspace and store it in the word-processing program for later reference. If there is a problem, you can go back to the Word file and try changes. Working only in the R-console soon gets one into trouble remembering details.

```
rm(list=ls())          ## clears the workspace of everything.
rm(z)                 ## takes out the single item, z.
```

The ## symbols can be used to add a comment to a line if you save a program. They can appear on the "command line" (after any commands) or on following or preceding lines as in most programming languages, e.g.

```
ls()                  ## prints a list of the contents of the current workspace.
```

```
z<-read.table("clipboard",header=T)    ## assigns the contents of the
                                         ## clipboard to z.
```

Using "header = T" lets one carry a header into R. Just using `read.table("clipboard")` gives the same result as including "header=F". Some references indicate that you should spell out TRUE or FALSE if there is a prospect that the symbols T or F will be used in the same program (we haven't done much actual programming here!). If you just transfer a vector of data, (y) by "y=read.table("clipboard"), and try to plot it, you may get a confusing message – use the statement "y=as.matrix(y)" to fix it. You can also use the formula "is.matrix(y)" to check; if the answer is "FALSE", there is a problem somewhere.

```
plot(z)              ## produces a plot of the contents of z
```

(assuming there are two columns in z, if not you will still get a plot with an x-axis supplied). It is wise to always plot data when you enter it just to be sure you have what you thought you entered or computed.

`plot(z,ylab="Amplitude",xlab="Time")` `##` Adds labels
if they were not in the material transferred via the clipboard.

`title(main=" text")` `##` adds a title to the plot
(with "text" being the title you want).

`text(locator(1),"Label",adj=0)`

followed by a line return lets you position the cursor on a plot whereupon another line return prints "Label" to the right of the cursor position.

`abline(v=x)`
adds a vertical line at point x on the graph (use `h=x` to get a horizontal line).

`abline(a,b)` `##` adds a line with intercept a and slope b to a graph.

You can use the regression program in EXCEL to fit a linear regression to data. The "abline" command is handy if you want to then do the plotting in R, after doing your regression analysis in EXCEL. However, it is at times desirable to get the regression coefficients directly while working in R. The simple statement `lm(y~x)` will produce estimates of the intercept and slope.

`lines(z)` `##`connects points on a graph of z.

`locator(n)` `##`lets you get coordinates of n points on a plot.

If you want x,y coordinates for a point on a graph, use `locator(1)`. Nothing happens until you hit return; then put the cursor on the first point that you want to locate and hit return. The location is printed out.

`x<-c(x1,x2)` `##`loads any number of values in an x-vector.
`y<-c(y1,y2)` `##`loads in the y-values.

If you cannot load data using the "read.table("clipboard")" command, you can type it in directly with the above command. A statement useful with vectors is the `data.frame` statement. You can combine vectors x and y by `>df<-data.frame(x=x,y=y)`. `df` is then equivalent to `z`, as used above. Use just `df<-data.frame(x,y)` unless you want to rename the vectors being transferred.

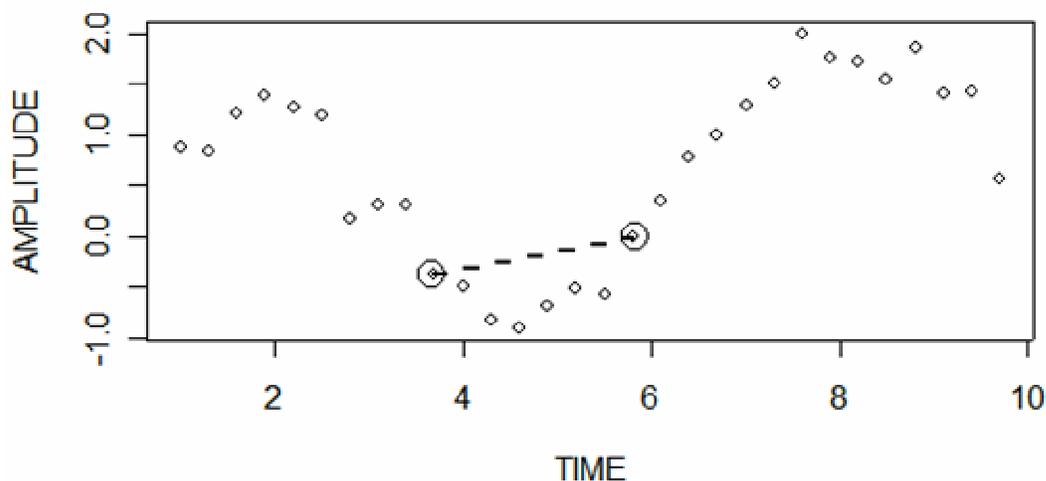
```
points(x,y,pch=1,cex=2)
```

Serves to put points on a graph. `pch=1` gets a circle, `cex=2` makes it fairly large. Try different values of `pch` and `cex`.

```
lines(x,y,lty=2,lwd=2)
```

Draws a line connecting points in the two vectors on your plot. “`lty`” determines line type and “`lwd`” line width.

An example using data shown previously (under “`lowess`”) follows, 2 points were located with `locator(2)`, plotted with `points()`, and connected with `lines()` as shown above.



If you have data stored in `z`, as with the sine curve above, you can get the `x`-values from: `x<-z[,1]` This statement will pick off column 1 from a matrix, `z`. The statement `>fix(z)` brings up a handy data-editing table. The same statement will also load a “function” statement for editing, assuming that it is in the workspace.

Figure margins can be adjusted by resetting a parameter before making the plot. One statement is `>par(mai=c(.75,.7,.05,.05))`. This sets margins (bottom, left, top, and right) in inches. Users should experiment as needed (i.e., if you use a title, the top margin should be wider). The default margins are wider, as evident in the graphs given above. There is a huge array of statements to set parameters [see `?par()`] for a list. Two useful such

commands are `las()` which sets orientation of the axis labels (1 for horizontal, 2 for vertical), and `lab()` which controls tick-marks. These can be combined in one statement, e.g.,

```
>par(mai=c(.7,.7,.05,.05),las=c(2),lab=c(10,6,1))
```

The first 2 numbers in `lab()` determine the number of tick marks on the x and y axes, and the last governs the length of the axis labels.

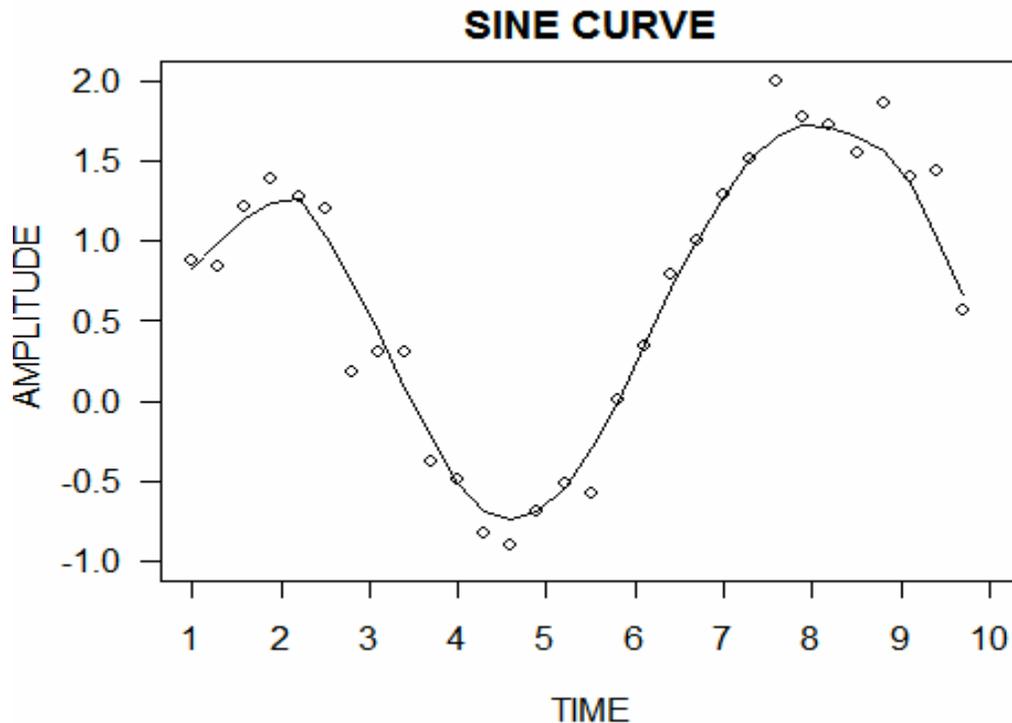
When one issues a plot command, the R-program sets up axes according to the data to be plotted. This can be unsatisfactory if there is more than one data set to be plotted, as any data in further commands that are outside the range of the set first plotted won't appear. The trick is to decide in advance what range of variables is to be used and set these by statements like:

```
z3<-data.frame(x2,y2)
where x2<-c(1989,2004)
and y2<-c(7.5,9.2), so that
> z3
  x2 y2
1 1989 7.5
2 2004 9.2
```

All of these commands can be loaded into a single program which will produce the desired plot. Using the data for a sine curve shown earlier, we use the following program (set off by curly brackets {}):

```
{y3=c(-1,2)           #range of y-variable
x3=c(1,10)           #range of x-variable
z3=data.frame(x3,y3) #frame for graph
par(mai=c(.8,.7,.4,1),las=c(1),lab=c(10,6,1)) #bottom,left,top,right margins
plot(z3,type="n",xlab="TIME",ylab="AMPLITUDE")
points(y)
lines(lowess(y,f=.2))
title(main="SINE CURVE")
}
```

This yields the following plot:



The program can be kept in a word processor document and loaded into R as needed. An important advantage is that the plot takes less space than those used thus far, which give much wider margins than one might want, and this approach controls the margins, labels, etc. to give a much better plot.

Under the “help” pull-down menu in R, one can find “Manuals”, and “An Introduction to R”, in which Chapter 12 gives more detail on graphs, including “Multiple figure environment” which shows how to include a number of graphs in one figure.