# SAMPLING AND STATISTICAL METHODS FOR BEHAVIORAL ECOLOGISTS

This book describes the sampling and statistical methods used most often by behavioral ecologists and field biologists. Written by a biologist and two statisticians, it provides a rigorous discussion, together with worked examples, of statistical concepts and methods that are generally not covered in introductory courses, and which are consequently poorly understood and applied by field biologists. The first section reviews important issues such as defining the statistical population when using nonrandom methods for sample selection, bias, interpretation of statistical tests, confidence intervals and multiple comparisons. After a detailed discussion of sampling methods and multiple regression, subsequent chapters discuss specialized problems such as pseudoreplication, and their solutions. It will quickly become *the* statistical handbook for all field biologists.

JONATHAN BART is a Research Wildlife Biologist, in the Biological Resources Division of the U.S. Geological Survey. He is currently based at the Snake River Field Station of the Forest and Rangeland Ecosystem Science Center in Boise, Idaho.

MICHAEL A. FLIGNER is Professor and Vice Chair of the Department of Statistics at the Ohio State University.

WILLIAM I. NOTZ is also a Professor in the Department of Statistics at the Ohio State University, and has served as Associate Editor for the *Journal of the American Statistical Association, Technometrics* and the *Journal of Statistical Education.*

To Susan
Alysha and Karen
Claudia

# SAMPLING AND STATISTICAL METHODS FOR BEHAVIORAL ECOLOGISTS

JONATHAN BART, MICHAEL A. FLIGNER, AND WILLIAM I. NOTZ

CAMBRIDGE UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011-4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia



First published 1998

Printed in the United Kingdom at the University Press, Cambridge

Typeset in Times NR 10/13 [SE]

*A catalog record for this book is available from the British Library*

*Library of Congress Cataloging in Publication data*

Bart, Jonathan.
Sampling and Statistical Methods for Behavioral Ecologists / Jonathan Bart, Michael A. Fligner, William I. Notz.
p. cm.
Includes index.
ISBN 0 521 45095 0 hardback.
ISBN 0 521 45705 X paperback.
1. Animal behavior – Statistical methods. 2. Sampling (Statistics)
I. Fligner, Michael A. II. Notz, William I. III. Title.
QL751.65.S73B37 1998
590′.1′5195 dc21 97–35086 CIP

ISBN 0 521 45095 0 hardback
ISBN 0 521 45705 X paperback

# Contents

# Preface

This book describes the sampling and statistical methods used most often by behavioral ecologists. We define behavioral ecology broadly to include behavior, ecology and such related disciplines as fisheries, wildlife, and environmental physiology. Most researchers in these areas have studied basic statistical methods, but frequently have trouble solving their design or analysis problems despite having taken these courses. The general reason for these problems is probably that introductory statistics courses are intended for workers in many fields, and each field presents a special, and to some extent unique, set of problems. A course tailored for behavioral ecologists would necessarily contain much material of little interest to students in other fields.

The statistical problems that seem to cause behavioral ecologists the most difficulty can be divided into several categories.

1. Some of the most difficult problems faced by behavioral ecologists attempting to design a study or analyze the resulting data fall between statistics – as it is usually taught – and biology. Examples include how to define the sampled and target populations, the nature and purpose of statistical analysis when samples are collected nonrandomly, and how to avoid pseudoreplication.
2. Some methods used frequently by behavioral ecologists are not covered in most introductory texts. Examples include survey sampling, capture–recapture, and distance sampling.
3. Certain concepts in statistics seem to need reinforcement even though they are well covered in many texts. Examples include the rationale of statistical tests, the meaning of confidence intervals, and the interpretation of regression coefficients.
4. Behavioral ecologists encounter special statistical problems in certain areas including index methods, detecting habitat 'preferences', and sampling behavior.

5. A few mathematical methods of use to behavioral ecologists are generally not covered in introductory methods courses. Examples include the statistical properties of ratios and other nonlinear combinations of random variables, rules of expectation, the principle of maximum likelihood estimation, and the Taylor series approximation.

This book is an attempt to address problems such as those above adopting the special perspective of behavioral ecology. Throughout the book, our general goals have been that behavioral ecologists would find the material relevant and that statisticians would find the treatment rigorous. We assume that readers will have taken one or more introductory statistics courses, and we view our book as a supplement, rather than a substitute, for these courses.

The book is based in part on our own research and consulting during the past 20 years. Before writing the text, however, we undertook a survey of the methods used by behavioral ecologists. We did this by examining every article published during 1990 in the journals *Behavioral Ecology and Sociobiology*, *Animal Behavior*, *Ecology*, and *The Journal of Wildlife Management* and all the articles on behavior or ecology published in *Science* and *Nature*. We tabulated the methods in these articles and used the results frequently in deciding what to include in the book and how to present the examples.

Chapter One describes statistical objectives of behavioral ecologists emphasizing how the statistical and nonstatistical aspects of data analysis reinforce each other. Chapter Two describes estimation techniques, introducing several statistical methods that are useful to behavioral ecologists. It is more mathematical than the rest of the book and can be skimmed by readers less interested in such methods. Chapter Three discusses tests and confidence intervals concentrating on the rationale of each method. Methods for ratios are discussed as are sample size and power calculations. The validity of $t$-tests when underlying data are non-normal is discussed in detail, as are the strengths and weaknesses of nonparametric tests. Chapter Four discusses survey sampling methods in considerable detail. Different sampling approaches are described graphically. Sample selection methods are then discussed followed by a description of multistage sampling and stratification. Problems caused by nonrandom sample selection are examined in detail. Chapter Five discusses regression methods emphasizing conceptual issues and how to use computer software to carry out general linear models' analysis.

The first five Chapters cover material included in the first few courses in statistical methods. In these Chapters, we concentrate on topics that

behavioral ecologists often have difficulty with, assuming that the reader has already been exposed to the basic methods and ideas. The subsequent Chapters discuss topics that are generally not covered in introductory statistics courses. We introduce each topic and provide suggestions for additional reading. Chapter Six discusses the difficult problem of pseudo-replication, introducing an approach which we believe might help to resolve the controversies in this area and focus the discussions on biological, rather than statistical, issues. Chapter Seven discusses special statistical problems that arise in sampling behavior. Chapter Eight discusses estimating and monitoring abundance, particularly by index methods. Chapter Nine discusses capture–recapture methods, while Chapter Ten emphasizes the estimation of survival. Chapter Eleven discusses resource selection and Chapter Twelve briefly mentions some other topics of interest to behavioral ecologists with suggestions for additional reading.

Appendix One gives a detailed explanation of frequently used statistical methods, whilst Appendix Two contains a set of tables for reference. They are included primarily so that readers can examine the formulas in more detail to understand how analyses are conducted. We have relegated this material to an appendix because most analyses are carried out using statistical packages and many readers will not be interested in the details of the analysis. Nonetheless, we encourage readers to study the material in the appendices as doing so will greatly increase one's understanding of the analyses. In addition, some methods (e.g., analysis of stratified samples) are not available in many statistical packages but can easily be carried out by readers able to write simple computer programs. Appendix Three contains detailed notes on derivation of the material in Appendix One.

This book is intended primarily for researchers who wish to use sampling techniques and statistical analysis as a tool but who do not have a deep interest in the underlying mathematical principles. We suspect, however, that many biologists will be interested in learning more about the statistical principles and techniques used to develop the methods we present. Knowledge of this material is of great practical use because problems arise frequently which can be solved readily by use of these methods, but which are intractable without them. Basic principles of expectation (by which many variance formulas may be derived) and use of the Taylor series approximation (by which nearly all the remaining variance formulas needed by behavioral ecologists may be derived) are examples of these methods. Maximum likelihood estimation is another statistical method that can be presented without recourse to complex math and is frequently of value to biologists. We introduce these methods in Chapter Two and

illustrate their use periodically in the rest of the book. These sections, however, can be skipped without compromising the reader's ability to understand later sections of the book.

Another approach of great utility in developing a deep understanding of the statistical methods we present is to prepare computer programs that carry out calculations and simulations. We encourage readers to learn some programming in an elementary language such as Basic or the languages included in many data bases or statistical packages and then to write short programs to investigate the material we present. Several opportunities for such projects are identified in the text, and all of the examples we mention are listed in the *Index* under the heading 'Computer programming, examples'. We have found that preparing programs in this manner not only ensures that one understands the fine structure of the analysis, but in addition frequently leads one to think much more deeply about how the statistical analysis helps us understand natural systems. Such efforts also increase one's intuition about whether studies can be carried out successfully given the resources available and about how to allocate resources among different segments of the study. Furthermore, data management, while not discussed in this book, frequently consumes far more time during analysis than carrying out the actual statistical tests, and in many studies is nearly impossible without recourse to computer programs. For all of these reasons, we encourage readers strongly to learn a programming language.

The authors thank the staff of Cambridge University Press for their assistance with manuscript preparation, especially our copy editor, Sarah Price. Much of the book was written while the senior author was a member of the Zoology Department at Ohio State University. He acknowledges the many stimulating discussions of biological statistics with colleagues there, especially Susan Earnst, Tom Grubb, and John Harder and their graduate students. JB also acknowledges his intellectual debt to Douglas S. Robson of Cornell University who introduced him to sampling techniques and other branches of statistics and from whom he first learned the value of integrating statistics and biology in the process of biological research.