Springer Texts in Statistics

Advisors:
George Casella Stephen Fienberg Ingram Olkin

Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Springer Texts in Statistics

Alfred: Elements of Statistics for the Life and Social Sciences
Berger: An Introduction to Probability and Stochastic Processes
Bilodeau and Brenner: Theory of Multivariate Statistics
Blom: Probability and Statistics: Theory and Applications
Brockwell and Davis: An Introduction to Times Series and Forecasting
Chow and Teicher: Probability Theory: Independence, Interchangeability, Martingales, Third Edition
Christensen: Log-Linear Models and Logistic Regression, Second Edition
Christensen: Plane Answers to Complex Questions: The Theory of Linear Models, Second Edition
Christensen: Advanced Linear Modeling: Multivariate, Time Series, and Spatial Data—Nonparametric Regression and Response Surface Maximization, Second Edition
Creighton: A First Course in Probability Models and Statistical Inference
Davis: Statistical Methods for the Analysis of Repeated Measurements
Dean and Voss: Design and Analysis of Experiments
du Toit, Steyn, and Stumpf: Graphical Exploratory Data Analysis
Durrett: Essentials of Stochastic Processes
Edwards: Introduction to Graphical Modelling, Second Edition
Finkelstein and Levin: Statistics for Lawyers
Flury: A First Course in Multivariate Statistics
Jobson: Applied Multivariate Data Analysis, Volume I: Regression and Experimental Design
Jobson: Applied Multivariate Data Analysis, Volume II: Categorical and Multivariate Methods
Kalbfleisch: Probability and Statistical Inference, Volume I: Probability, Second Edition
Kalbfleisch: Probability and Statistical Inference, Volume II: Statistical Inference, Second Edition
Karr: Probability
Keyfitz: Applied Mathematical Demography, Second Edition
Kiefer: Introduction to Statistical Inference
Kokoska and Nevison: Statistical Tables and Formulae
Kulkarni: Modeling, Analysis, Design, and Control of Stochastic Systems
Lehmann: Elements of Large-Sample Theory
Lehmann: Testing Statistical Hypotheses, Second Edition
Lehmann and Casella: Theory of Point Estimation, Second Edition
Lindman: Analysis of Variance in Experimental Design
Lindsey: Applying Generalized Linear Models
Madansky: Prescriptions for Working Statisticians

(continued after index)

Charles S. Davis

Statistical Methods for the Analysis of Repeated Measurements

With 20 Illustrations

Springer

Charles S. Davis
Senior Director, Clinical Operations and Biostatistics
Elan Pharmaceuticals
7475 Lusk Boulevard
San Diego, CA 92121
USA
Chuck.Davis@elancorp.com

Library of Congress Cataloging-in-Publication Data
Davis, Charles S. (Charles Shaw), 1952–
Statistical methods for the analysis of repeated measurements/Charles S. Davis.
p. cm. — (Springer texts in statistics)
Includes bibliographical references and index.
ISBN 0-387-95370-1 (alk. paper)
1. Multivariate analysis. 2. Experimental design. I. Title. III. Series.
QA278.D343 2002
519.5'35—dc21 2001054913

Printed on acid-free paper.

Production managed by Frank McGuckin; manufacturing supervised by Jerome Basma.
Camera-ready copy prepared from the author's LaTeX2e files using Springer's svsing2e.sty file.
Printed and bound by Edwards Brothers, Inc., Ann Arbor, MI.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-95370-1 SPIN 10853625

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

Preface

I have endeavored to provide a comprehensive introduction to a wide variety of statistical methods for the analysis of repeated measurements. I envision this book primarily as a textbook, because the notes on which it is based have been used in a semester-length graduate course I have taught since 1991. This course is primarily taken by graduate students in biostatistics and statistics, although students and faculty from other departments have audited the course. I also anticipate that the book will be a useful reference for practicing statisticians. This assessment is based on the positive responses I have received to numerous short courses I have taught on this topic to academic and industry groups.

Although my intent is to provide a reasonably comprehensive overview of methods for the analysis of repeated measurements, I do not view this book as a definitive "state of the art" compendium of research in this area. Some general approaches are extremely active areas of current research, and it is not feasible, given the goals of this book, to include a comprehensive summary and list of references. Instead, my focus is primarily on methods that are implemented in standard statistical software packages. As a result, the level of detail on some topics is less than in other books, and some more recent methods of analysis are not included. One particular example is the topic of nonlinear mixed models for the analysis of repeated measurements (Davidian and Giltinan, 1995; Vonesh and Chinchilli, 1996). With respect to some of the more recent methods of analysis, I do attempt to mention some of the areas of current research.

The prerequisites for a course based on this book include knowledge of mathematical statistics at the level of Hogg and Craig (1995) and a course

in linear regression and ANOVA at the level of Neter et al. (1985). Individuals without these prerequisites who have audited the graduate course or attended short courses have also been able to benefit from much of the material.

Because a wide variety of methods are covered, knowledge of topics such as multivariate normal distribution theory, categorical data analysis, and generalized linear models would also be useful. However, my philosophy is not to assume any particular knowledge of these areas and to present the necessary background material in the book.

When I began to develop my graduate course on the analysis of repeated measurements, no suitable text was available for the course as I envisioned it, and I made the decision to prepare my own notes. Since then, multiple books on the analysis of repeated measurements have been published. I regularly refer to the following books (listed chronologically): Hand and Taylor (1987), Crowder and Hand (1990) [updated as Hand and Crowder (1996)], Diggle (1990), Jones (1993), Diggle et al. (1994), Kshirsagar and Smith (1995), Vonesh and Chinchilli (1996), and Lindsey (1999), among others. Although some of the existing books are reasonably comprehensive in their coverage, others are more narrowly focused on specialized topics. This book is more comprehensive than many and is targeted at a lower mathematical level and focused more on applications than most. In summary, it is more oriented toward statistical practitioners than to statistical researchers.

Two obvious distinctions of this book are the extensive use of real data sets and the inclusion of numerous homework problems. Eighty real data sets are used in the examples and homework problems. These data sets are available from the website www.springer-ny.com (click on "author websites"). Because many of the data sets can be used to demonstrate multiple methods of analysis, instructors can easily develop additional homework problems and exam questions based on the data sets provided.

The inclusion of homework problems makes this book especially well-suited as a course text. Approximately 85% of the homework problems involve data analysis. The focus of these problems is not on providing a definitive analysis of the data but rather on providing the reader with experience in knowing when, and learning how, to select and apply appropriate methods of analysis. Although many of the examples and homework problems have a biomedical focus, the principles and methods apply to other subject areas as well.

My graduate course and short course notes include numerous examples of the use of, and output from, statistical software packages, primarily SAS (SAS Institute, 1999). I have purposely chosen not to include programming statements or computer output in the book. I do provide the raw data for nearly all examples as well as the key results of all analyses. In this way, readers will be able to carry out and verify the results of their own analyses using their choice of software.

The notes on which this book is based are in the form of overhead transparencies produced using TEX (Knuth, 1986). This format is well-suited for instructors. The course notes also include programming statements and computer output for the examples, prepared primarily using SAS. Course instructors interested in obtaining this supplemental material, as well as solutions to homework problems, should contact Springer-Verlag.

I would like to thank John Kimmel of Springer-Verlag for initially encouraging me to write this book and for his support and advice during its preparation. I am also grateful to the graduate students who have participated in my course since 1991 and to the attendees at external short courses; both groups have motivated me to develop and expand the notes on which this book is based. I also thank Michelle Larson for her assistance in the preparation of solutions to the homework problems and Kathy Clark for her careful review of the manuscript. Finally, I thank my wife, Ruth, and our children, Michael, Carrie, and Nathan, for their understanding and support during this endeavor.

San Diego, California Charles S. Davis
November 2001

Contents

List of Tables

List of Figures