

Essentials of Biostatistics

Lisa M. Sullivan

Series Editor: Richard Riegelman

Essentials of Biostatistics

Lisa M. Sullivan, PhD

Chair, Department of Biostatistics
Boston University School of Public Health
Boston, Massachusetts



JONES AND BARTLETT PUBLISHERS

Sudbury, Massachusetts

BOSTON TORONTO LONDON SINGAPORE

World Headquarters

Jones and Bartlett Publishers
40 Tall Pine Drive
Sudbury, MA 01776
978-443-5000
info@jbpub.com
www.jbpub.com

Jones and Bartlett Publishers
Canada
6339 Ormindale Way
Mississauga, Ontario
L5V 1J2
Canada

Jones and Bartlett Publishers
International
Barb House, Barb Mews
London W6 7PA
United Kingdom

Jones and Bartlett's books and products are available through most bookstores and online booksellers. To contact Jones and Bartlett Publishers directly, call 800-832-0034, fax 978-443-8000, or visit our website www.jbpub.com.

Substantial discounts on bulk quantities of Jones and Bartlett's publications are available to corporations, professional associations, and other qualified organizations. For details and specific discount information, contact the special sales department at Jones and Bartlett via the above contact information or send an email to specialsales@jbpub.com.

Copyright © 2008 by Jones and Bartlett Publishers, Inc.

All rights reserved. No part of the material protected by this copyright may be reproduced or utilized in any form, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the copyright owner.

This publication is designed to provide accurate and authoritative information in regard to the Subject Matter covered. It is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional service. If legal advice or other expert assistance is required, the service of a competent professional person should be sought.

Library of Congress Cataloging-in-Publication Data

Sullivan, Lisa M. (Lisa Marie), 1961-

Essentials of biostatistics in public health / Lisa M. Sullivan.

p. ; cm.

Includes bibliographical references and index.

ISBN-13: 978-0-7637-3737-5 (pbk.)

ISBN-10: 0-7637-3737-2 (pbk.)

1. Public health--Statistical methods. 2. Medical statistics. I. Title.

[DNLM: 1. Biometry--methods. 2. Public Health Practice. WA 950 S949e 2008]

RA409.S73 2008

362.1--dc22

2007019656

6048

Production Credits

Publisher: Michael Brown

Associate Editor: Katey Birtcher

Production Director: Amy Rose

Production Editor: Tracey Chapman

Marketing Manager: Sophie Fleck

Manufacturing Buyer: Therese Connell

Composition: Publishers' Design and Production Services, Inc.

Cover Design: Kristin E. Ohlin

Cover Image: © Digital Vision/Getty Images

Printing and Binding: Malloy, Inc.

Cover Printing: John Pow Company

Printed in the United States of America

11 10 09 08 07 10 9 8 7 6 5 4 3 2 1

Table of Contents

Acknowledgments		ix
Preface		xi
Prologue		xiii
Chapter 1	Introduction	1
1.1	What Is Biostatistics?	1
1.2	What Are the Issues?	2
1.3	Summary	4
Chapter 2	Study Designs	5
2.1	Vocabulary	5
2.2	Observational Study Designs	6
2.3	Randomized Study Designs	11
2.4	The Framingham Heart Study	14
2.5	More on Clinical Trials	15
2.6	Sample Size Implications	17
2.7	Summary	17
2.8	Practice Problems	18
Chapter 3	Quantifying the Extent of Disease	19
3.1	Prevalence	19
3.2	Incidence	20
3.3	Relationships Between Prevalence and Incidence	23
3.4	Comparing the Extent of Disease Between Groups	23
3.5	Summary	25
3.6	Practice Problems	26

Chapter 4	Summarizing Data Collected in the Sample	29
4.1	Dichotomous Variables	30
4.2	Ordinal and Categorical Variables	34
4.3	Continuous Variables	43
4.4	Summary	54
4.5	Practice Problems	56
Chapter 5	The Role of Probability	59
5.1	Sampling	59
5.2	Basic Concepts	61
5.3	Conditional Probability	62
5.4	Independence	65
5.5	Bayes' Theorem	66
5.6	Probability Models	67
5.7	Summary	88
5.8	Practice Problems	88
Chapter 6	Confidence Interval Estimates	93
6.1	Introduction to Estimation	94
6.2	Confidence Intervals for One Sample, Continuous Outcome	95
6.3	Confidence Intervals for One Sample, Dichotomous Outcome	96
6.4	Confidence Intervals for Two Independent Samples, Continuous Outcome	98
6.5	Confidence Intervals for Matched Samples, Continuous Outcome	101
6.6	Confidence Intervals for Two Independent Samples, Dichotomous Outcome	104
6.7	Summary	110
6.8	Practice Problems	110
Chapter 7	Hypothesis Testing Procedures	113
7.1	Introduction to Hypothesis Testing	113
7.2	Tests with One Sample, Continuous Outcome	120
7.3	Tests with One Sample, Dichotomous Outcome	123
7.4	Tests with One Sample, Discrete Outcome	124
7.5	Tests with Two Independent Samples, Continuous Outcome	128
7.6	Tests with Matched Samples, Continuous Outcome	131
7.7	Tests with Two Independent Samples, Dichotomous Outcome	134
7.8	Tests with More Than Two Independent Samples, Continuous Outcome	137
7.9	Tests with Two or More Independent Samples, Discrete Outcome	144
7.10	Summary	148
7.11	Practice Problems	150

Chapter 8	Power and Sample Size Determination	155
8.1	Estimating Sample Size for Confidence Intervals Estimates	155
8.2	Estimating Sample Size for Hypothesis Testing	162
8.3	Summary	171
8.4	Practice Problems	171
Chapter 9	Multivariable Methods	175
9.1	Confounding and Effect Modification	176
9.2	The Cochran-Mantel-Haenszel Method	182
9.3	Introduction to Regression Analysis	184
9.4	Multiple Linear Regression Analysis	188
9.5	Multiple Logistic Regression Analysis	191
9.6	Summary	192
9.7	Practice Problems	193
Appendix		197
Index		209

Acknowledgments

I am very grateful to Professor Ralph D'Agostino Sr. for his constant guidance and support. I am also thankful to Dr. Richard Riegelman who was instrumental in providing invaluable insight from the development to the completion of the textbook. Special thanks to Lori Chibnik, Michelle Keyes, Michael Pencina, and Sharon Milewitz for their careful editing and reviewing. I thank my family for their constant support and most especially, I am thankful to Kimberly Dukes and Kevin Green for their unending love and support.

This textbook is dedicated to the memory of my cousin, Catherine Render. Catherine was a remarkable woman. She was a loving daughter, sister, wife, mother, cousin, and friend who excelled in every role. She lost a long battle with breast cancer in January 2006. She fought through many setbacks, always pushing forward optimistically, taking care of everyone around her, never asking why. She was and always will be an inspiration to her family and to so many others who were fortunate enough to know her.

Biostatisticians play an extremely important role in addressing important medical and public health problems. Unfortunately, there are many more problems than solutions. We must never lose sight of the fact that our work is important in improving health and well-being. We need qualified biostatisticians to work in research teams to address problems like breast cancer, cardiovascular disease, diabetes, and so many others.

Lisa M. Sullivan

Preface

Essentials of Biostatistics provides a fundamental and engaging background for students learning to apply and appropriately interpret biostatistical applications in the field of public health. The examples are real, important, and represent timely public health problems. The author aims to make the material relevant, practical and interesting for students. Throughout the textbook, the author uses data from the Framingham Heart Study and from clinical trials in a variety of major areas. The author presents example applications involving important risk factors such as blood pressure, cholesterol, smoking and diabetes and their relationships to incident cardiovascular and cerebrovascular disease throughout. Clinical trials investigating new drugs to lower cholesterol, to reduce pain and to promote healing following surgery are also considered. The author presents examples with relatively few subjects to illustrate computations while minimizing the actual computation time, as a particular focus is mastery of “by-hand” computations. All of the techniques are then applied to and illustrated on real data from the Framingham Study and large clinical trials. For each topic, the author discusses methodology—including assumptions, statistical computations and the appropriate interpretation of results. Key formulas are summarized at the end of each chapter.

Prologue

Learning to avoid being deceived by data is essential for all educated citizens. It is a key skill for future clinicians, public health practitioners, and health researchers. In *Essentials of Biostatistics*, Lisa Sullivan PhD has written a text that guides students through this maze. The text uses real and relevant examples drawn from her own experience working on the Framingham Heart Study and clinical trials.

Essentials of Biostatistics takes an intuitive, step-by-step, hands-on approach to walking students through statistical principles. The approach emphasizes understanding what questions to ask and what the answers mean.

Calculations are designed to enhance understanding. The accompanying Workbook utilizes Excel, making it familiar and engaging for students. The combination of text and workbook provides a solid foundation for future course work in statistics as well as the type of grounding needed to read and appreciate the health research literature.

Essentials of Biostatistics is a tried and true approach. Lisa Sullivan has spent 20 years teaching undergraduate students biostatistics and working with undergraduates as Assistant Dean for Undergraduate Education in Public Health at the Boston University School of Public Health. In these roles she has repeatedly won teaching awards for her skills and commitment to education in Biostatistics.

Dr. Sullivan possesses a unique combination of sophisticated biostatistics expertise with an engaging and articulate style—a rare combination indeed. Her recent appointment as Chair of the Department of Biostatistics confirms her expertise in biostatistics. Even a quick look at *Essentials of Biostatistics* will convince you of her skills in communication and education.

I am delighted that Lisa Sullivan has agreed to include her book and workbook in our Essential Public Health series. There is no better book to recommend for the anxious student first confronting the field of biostatistics. Students will find the book and workbook engaging and relevant. Just take a look and see for yourself.

Richard Riegelman MD, MPH, PhD
Series Editor—Essential Public Health series

