ESSENTIAL PUBLIC HEALTH

Series Editor: Richard Riegelman

Essentials of Biostatistics in Public Health

Fourth Edition

Lisa M. Sullivan, PhD

Associate Dean for Education
Professor of Biostatistics
Boston University School of Public Health
Boston, Massachusetts

JONES & BARTLETT
LEARNING

World Headquarters
Jones & Bartlett Learning
25 Mall Road
Burlington, MA 01803
978-443-5000
info@jblearning.com
www.jblearning.com

23220-2

Production Credits
Vice President, Product Management: Marisa R. Urbano
Vice President, Content Strategy and Implementation: Christine Emerton
Director, Product Management: Matthew Kane
Product Manager: Sophie Fleck Teague
Director, Content Management: Donna Gridley
Manager, Content Strategy: Carolyn Pershouse
Content Strategist: Sara Bempkins
Director, Project Management and Content Services: Karen Scott
Manager, Program Management: Kristen Rogers
Project Manager: Angela Montoya
Senior Digital Project Specialist: Angela Dooley
Senior Marketing Manager: Susanne Walker
Content Services Manager: Colleen Lamy
Vice President, Manufacturing and Inventory Control: Therese Connell
Composition: Straive
Cover and Text Design: Michael O'Donnell
Media Development Editor: Faith Brosnan
Rights & Permissions Manager: John Rusk
Rights Specialist: Liz Kincaid
Cover Image (Title Page, Part Opener, Chapter Opener): © U-STUDIOGRAPHY DD59/Shutterstock
Printing and Binding: LSC Communications

Library of Congress Cataloging-in-Publication Data

Names: Sullivan, Lisa M., author.
Title: Essentials of biostatistics in public health / Lisa M. Sullivan.
Description: Fourth edition. | Burlington, Massachusetts : Jones & Bartlett Learning, [2023] | Series: Essential public health series | Includes index.
Identifiers: LCCN 2021047475 | ISBN 9781284231977 (paperback)
Subjects: LCSH: Public health—Statistical methods. | Medical statistics. | BISAC: MEDICAL / Biostatistics
Classification: LCC RA409 .S73 2023 | DDC 610.72—dc23/eng/20211115
LC record available at https://lccn.loc.gov/2021047475

6048

Printed in the United States of America
26 25 24 23 22 10 9 8 7 6 5 4 3 2 1

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 me 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.

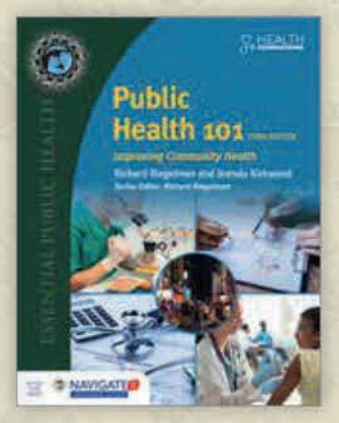
Brief Contents

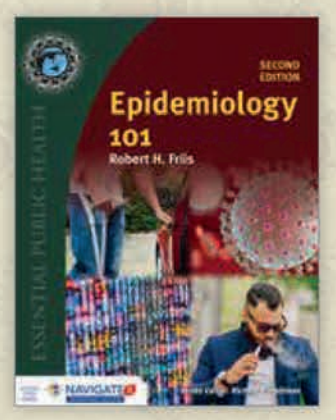
Contents

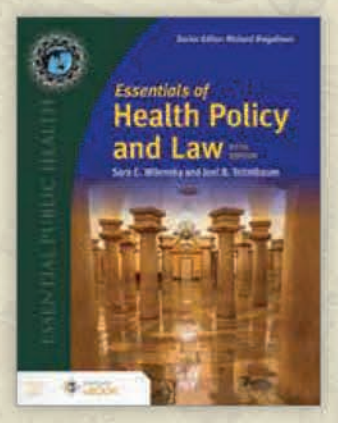
THE ESSENTIAL PUBLIC HEALTH SERIES

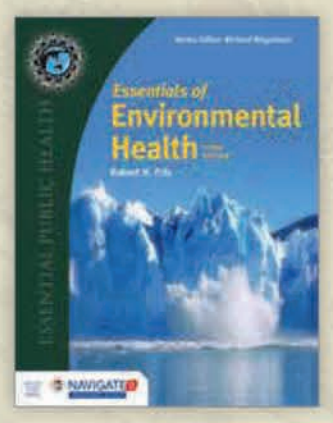
From the impact of COVID-19 to the cost of health care, this unique series will introduce you to the full range of issues that impact the public's health.

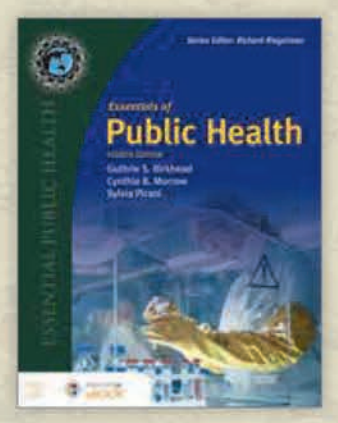
Current and Forthcoming Titles in the Essential Public Health Series:

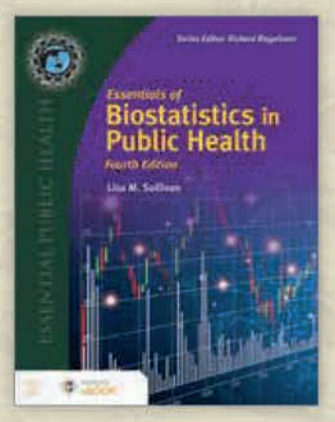


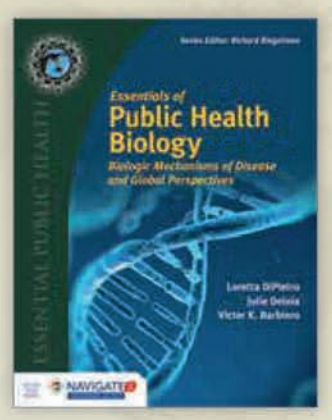


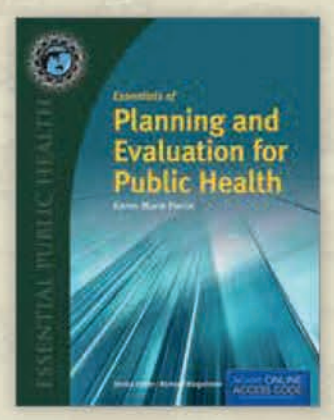


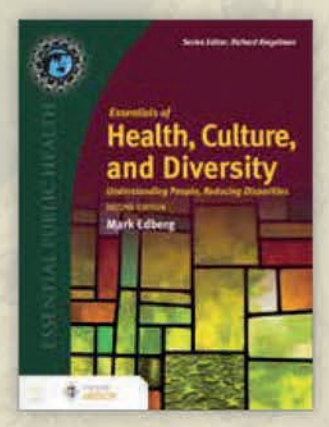


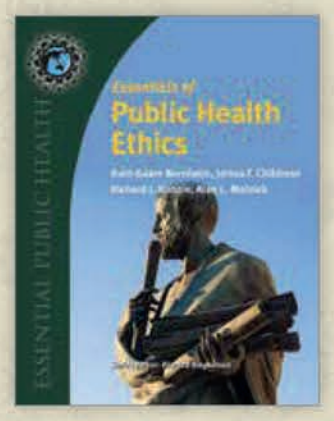


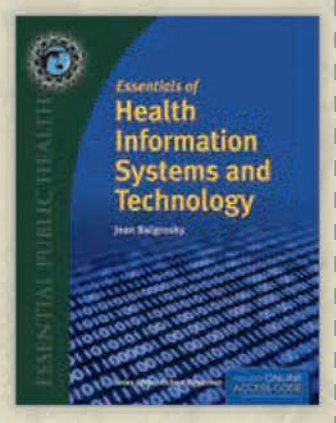


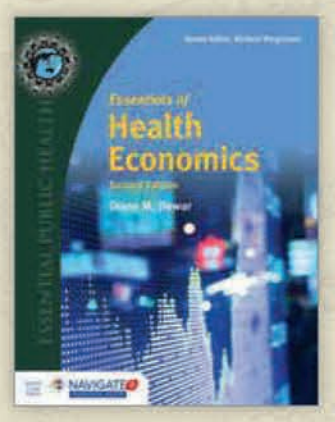


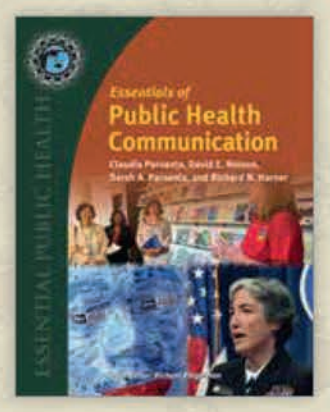


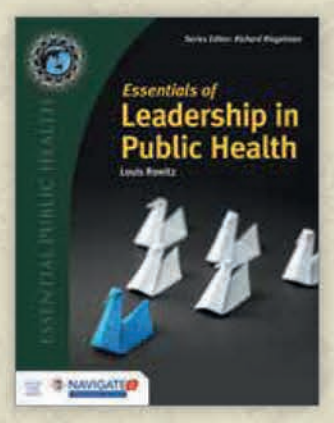


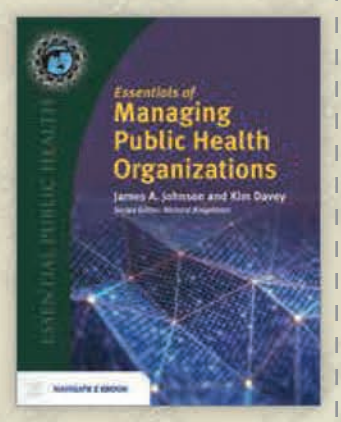


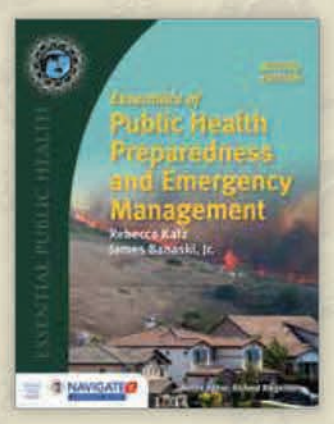


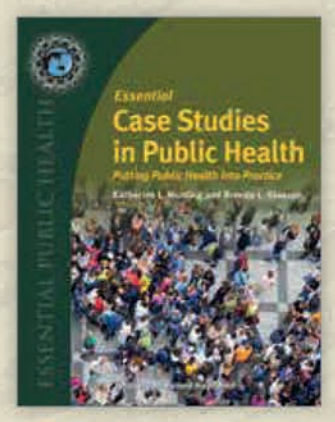


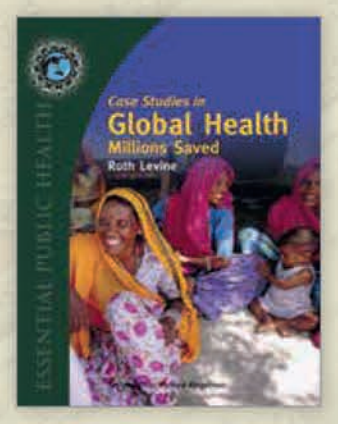












ABOUT THE EDITOR

Richard K. Riegelman, MD, MPH, PhD, is Professor of Epidemiology-Biostatistics, Medicine, and Health Policy, and Founding Dean of The George Washington University Milken Institute School of Public Health in Washington, DC. He has taken a lead role in developing the Educated Citizen and Public Health initiative which has brought together arts and sciences and public health education associations to implement the Institute of Medicine of the National Academies' recommendation that "…all undergraduates should have access to education in public health." Dr. Riegelman also led the development of The George Washington's undergraduate major and minor and currently teaches "Public Health 101" and "Epidemiology 101" to undergraduates.

www.EssentialPublicHealth.com

Background: © artconcept/Shutterstock

Editor's Preface

Understanding how to present and interpret data is the foundation for evidence-based public health. It is essential for public health practitioners, future clinicians, and health researchers to know how to use data and how to avoid being deceived by data. In *Essentials of Biostatistics in Public Health*, Lisa Sullivan ably guides students through this maze. To do so, she uses an abundance of real and relevant examples drawn from her own experience working on the Framingham Heart Study and clinical trials.

Essentials of Biostatistics in Public Health takes an intuitive, step-by-step, hands-on approach in walking students though statistical principles. It emphasizes understanding which questions to ask and knowing how to interpret statistical results appropriately.

The fourth edition of *Essentials of Biostatistics in Public Health* builds upon the success of the previous editions in presenting state-of-the-art biostatistical methods that are widely used in public health and clinical research. A new chapter on careers in biostatistics describes biostatistics is an exciting career choice. The work is impactful, there is variety in the work, the job outlook is very promising, salary projections are strong, and there are many opportunities to work with other scientists and professionals on important health issues.

The *Fourth Edition* also features new illustrative case studies related to the COVID-19 pandemic that illustrate the role that biostatisticians play in understanding the virus, how it is transmitted, who is at risk for more series consequences of infection, accuracy of testing, development and approval of vaccines, and so much more.

The strategies used in *Essentials of Biostatistics in Public Health* represent a tried-and-true, classroom-tested approach. Lisa Sullivan has more than two decades of experience teaching biostatistics to both undergraduates and graduate students. As Assistant Dean for Undergraduate Programs in Public Health at Boston University, she developed and taught undergraduate courses in biostatistics. She also served as the chair of the Department of Biostatistics. Today she is the Associate Dean for Education at Boston University School of Public Health. Her background speaks to her unique ability to combine the skills of biostatistics with the skills of education.

Dr. Sullivan has won numerous teaching awards for her skills and commitment to education in biostatistics, including the Association of Schools of Public Health Award for Teaching Excellence. She possesses a unique combination of sophisticated biostatistics expertise and a clear and engaging writing style—one that can draw students in and help them understand even the most difficult topic. Even a quick glance through *Essentials of Biostatistics in Public Health* will convince you of her skills in communication and education.

I am delighted that Dr. Sullivan has included 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
Editor, Essential Public Health Series

Author's Preface

Essentials of Biostatistics in Public Health, Fourth Edition 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, are important, and represent timely public health problems. The author aims to make the material relevant, practical, and engaging for students. Throughout the textbook, the author uses data from the Framingham Heart Study and from observational studies and 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. Illustrative case studies highlighting biostatistical applications related to COVID-19 are interspersed throughout. 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 Heart Study and large observational studies and 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.

Additional Materials Available Online

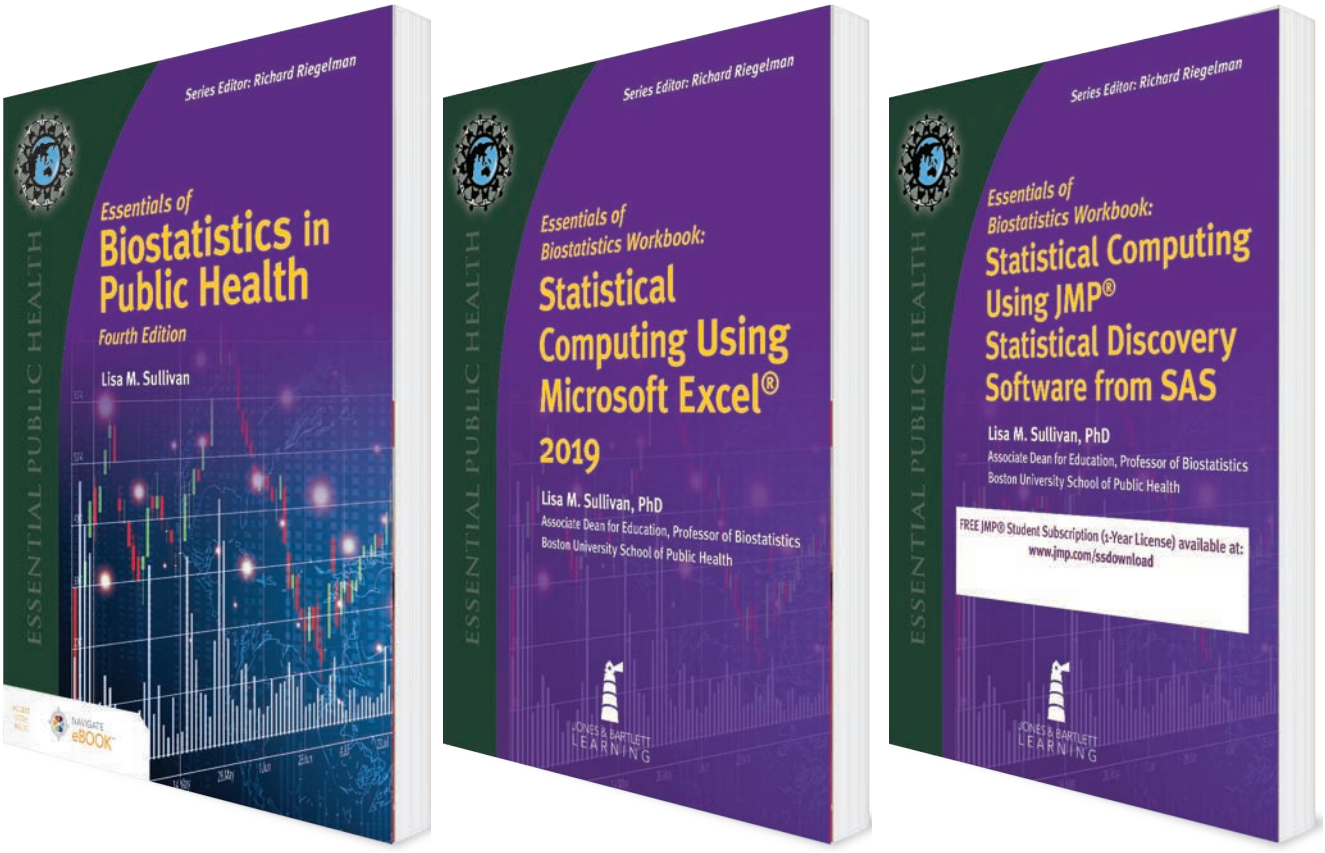
Comprehensive eBook with Workbooks

Looking for the book in digital format?

Redeem the Access Code found at the front of the book to get full access to the eBook along with **practice quizzes** and an **interactive glossary** as well as **voice notes** and **highlighting capabilities**.

Inside the eBook, you'll also get access to **two workbooks**: *Statistical Computing Using Microsoft Excel® 2019* and *Statistical Computing Using JMP® Statistical Discovery Software from SAS* as well as access to a free download of JMP Student Subscription.

The eBook Is Three Books in One:



Essentials of Biostatistics in Public Health, Fourth Edition, comes with access to two online workbooks with step-by-step instructions for using Microsoft Excel or JMP® statistical discovery software from SAS, Student Edition.

In the workbooks, you'll learn how to use either program to perform the statistical computations and analysis described in the textbook.

To gain access to the workbooks, simply redeem the access code found on the card at the front of the book at www.jblearning.com.

Workbook: Statistical Computing Using Microsoft Excel® 2019

Statistical Computing Using Microsoft Excel® 2019 has been fully updated, including new screenshots to reflect how to use Microsoft Excel 2019 for statistical computing.

Contents:

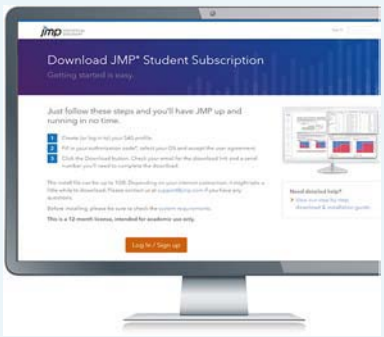
Workbook: Statistical Computing Using JMP® Statistical Discovery Software from SAS

Contents:

Get a FREE Download: JMP® Student Subscription (1-Year License)

JMP Student Subscription is an easy-to-use version of JMP® software for both Windows and Mac that provides all the statistical analysis and graphical tools covered in this textbook.

Students using this text and the accompanying workbook can get a free 1-year subscription to JMP®. Please **go to your eBook** (Access Code Card found at the front of the book) and follow the instructions to redeem access to your free JMP® student subscription.



Pedagogical Features

Key Questions

- How much variability in a measure is reasonable?
- What does it mean when you have a medical test and your result comes back and is "within normal limits"?
- Should you be concerned if you take a medical test and your result is not "normal"?

Key Questions

Key Questions prompt readers to apply critical thinking to important concepts in chapters.

Learning Objectives

Get a preview of the topics to be covered in the chapter.

LEARNING OBJECTIVES

By the end of this chapter, the reader will be able to

- Define the terms "equally likely" and "at random"
- Compute and interpret unconditional and conditional probabilities
- Evaluate and interpret independence of events
- Explain the key features of the binomial distribution model
- Calculate probabilities using the binomial formula
- Explain the key features of the normal distribution model
- Calculate probabilities using the standard normal distribution table
- Compute and interpret percentiles of the normal distribution

In the News

Hemoglobin is a protein found in red blood cells that carries oxygen. People with low levels might have anemia, and people with high levels might have heart or lung problems. In men, the "normal" range is 13.8–17.2 g/dL and in women, the "normal" range is 12.1–15.1 g/dL.[a] The American Heart Association categorizes systolic [when the heart muscle contracts] and diastolic [when the heart is at rest and refilling] blood pressures levels as normal, prehypertensive, Stage 1 hypertension, Stage 2 hypertension, and hypertensive crisis.[b] Normal systolic blood pressure is less than 120 mm Hg and normal diastolic blood pressure is less than 80 mm Hg. Cholesterol levels [e.g., total serum cholesterol, high-density lipoprotein [HDL] cholesterol and low-density lipoprotein [LDL] cholesterol] are not categorized as normal or not, but as desirable, borderline high, and high. For some of the cholesterol subfractions, different values define desirable versus borderline high categories in men versus women. Desirable total serum cholesterol for both men and women is less than 200 mg/dL.[c]

"Normal" ranges of characteristics are usually determined by measurements in a reasonably large sample of healthy volunteers and reflect the range of values that extend two standard deviations from the mean in either direction. Approximately 5% of values are outside of that range and are also observed in healthy volunteers.

[a] MedlinePlus. "Hemoglobin." Available at https://www.nlm.nih.gov/medlineplus/ency/article/003645.htm

[b] American Heart Association. "Understanding blood pressure readings." Available at http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/AboutHighBloodPressure/Understanding-Blood-Pressure-Readings_UCM_301764_Article.jsp#.V4pYkrgrIps

[c] "Understanding cholesterol and heart health." *MedlinePlus: The Magazine.* Available at https://medlineplus.gov/magazine/issues/summer12/articles/summer12pg6-7.html

In the News

An "In the News" segment highlights timely news reports on important public health issues.

Dig In
Critical thinking questions help you apply what you've learned to real life scenarios.

Dig In

- Are there any psychological risks to defining "normal ranges" for tests that, by design, include healthy values that are outside of those ranges?
- Can you think of a currently available test that is particularly problematic in terms of psychological risk? How might we minimize the risks for patients interested in taking that test?
- Can you think of a better approach to defining "normal ranges"?

Illustrative COVID-19 Case Studies
Illustrative case studies highlighting biostatistical applications related to COVID-19.

ILLUSTRATIVE CASE STUDY

COVID-19: The Role of Probability

Slowing the spread of COVID-19 infection takes coordinated effort and includes screening, testing, and contact tracing, and more recently, vaccinations. Screening usually involves a series of questions to assess an individual's risk for COVID-19 infection, and is often used for someone who might be at risk but has yet to experience symptoms. The idea is to identify potential cases early to prevent transmission. Diagnostic tests are used to confirm or rule out infection once a person begins to experience symptoms. There are a number of different kinds of screening and diagnostic tests that vary in terms of cost, accuracy, and processing time (more on tests following). Contact tracing involves identifying and informing people that they may have been exposed to someone with COVID-19 so they can carefully monitor their own symptoms. If they do begin to experience symptoms, they can be tested and quarantine, and if they test positive for COVID-19, they would then self-isolate to minimize the spread of infection to others.

There are number of tools and processes used to screen for COVID-19. The CDC has their employees screen daily

Practice Problems
Test your knowledge and apply your skills using the practice problems found at the end of each chapter.

Table 5.15 Body Mass Index (BMI) and Preterm Delivery

	BMI < 30 kg/m²	BMI 30–34.9 kg/m²	BMI 35 + kg/m²
Preterm	320	80	120
Full term	4700	480	300

13. **Table 5.15** cross-classifies pregnant women in a study by their body mass index (BMI) at 16 weeks' gestation and whether they had a preterm delivery.
    a. What is the probability that a woman delivers preterm?
    b. What is the probability a woman has a BMI less than 30 kg/m² and delivers preterm?
    c. What proportion of women with a BMI greater than 35 kg/m² delivers preterm?
    d. Are BMI and preterm delivery independent? Justify.

Acknowledgments

I am very grateful to Dr. Richard Riegelman for his unending energy and enthusiasm for this project. I thank my family for their encouragement; My parents, brothers Sean, Patrick and Andy, and Casey, Ryan, Aislinn, Abbi, and Gavin for their inspiration; most especially, Kimberly Dukes and Kevin Green for their unending love and support.

About the Author

Lisa M. Sullivan has a PhD in statistics and is Professor and former Chair of the Department of Biostatistics at the Boston University School of Public Health. She is also Associate Dean for Education. She teaches biostatistics for MPH students and lectures in biostatistical methods for clinical researchers.

From 2003 to 2015, Lisa was the principal investigator of the National Heart, Lung, and Blood Institute's *Summer Institute for Training in Biostatistics*, which was designed to promote interest in the field of biostatistics and to expose students to the many exciting career opportunities available to them. Lisa is the recipient of numerous teaching awards, including the Norman A. Scotch Award and the prestigious Metcalf Award, both for excellence in teaching at Boston University. In 2008 she won the Association of Schools of Public Health/Pfizer Excellence in Teaching Award.

In 2011, she won the American Statistical Association's Section on Teaching Statistics in the Health Sciences Outstanding Teaching Award. In 2013, she won the Mostellar Statistician of the Year Award, presented by the Boston Chapter of the American Statistical Association, and was also awarded the Massachusetts ACE National Network of Women Leaders Leadership Award. In 2020 she received the Mu Sigma Rho William D. Warde Statistics Education Award from the American Statistical Association and, in 2021, she was named a Fellow of the American Statistical Association in recognition of her "outstanding contributions and strong leadership in statistical education and collaborative research in cardiovascular disease and clinical trials, and for exceptional service in training the next generation of biostatisticians and public health practitioners."

She has authored over 200 papers and several books, including *Essentials of Biostatistics for Public Health, Fourth Edition* (Sullivan, 2022), *Teaching Public Health* (Sullivan & Galea, 2019), *Public Health: An Introduction to the Science and Practice of Population Health* (Shultz, Sullivan, & Galea, 2019), and *Biostatistics for Population Health* (Sullivan, 2020). She is currently chairing the Association of Schools and Programs of Public Health's initiative: *Framing the Future: Public Health Education for 2030.*