7
EDITION

Study Guide
to Accompany Salkind and Frey's

Statistics for People Who (Think They) Hate Statistics

Neil J. Salkind
Bruce B. Frey
Karin Lindstrom Bremer

Study Guide to Accompany Salkind and Frey's

Statistics for People Who *(Think They)* Hate Statistics

7
EDITION

Sara Miller McCune founded SAGE Publishing in 1965 to support the dissemination of usable knowledge and educate a global community. SAGE publishes more than 1000 journals and over 800 new books each year, spanning a wide range of subject areas. Our growing selection of library products includes archives, data, case studies and video. SAGE remains majority owned by our founder and after her lifetime will become owned by a charitable trust that secures the company's continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne

Study Guide to Accompany Salkind and Frey's

Statistics for People Who *(Think They)* Hate Statistics

7
EDITION

Neil J. Salkind
Bruce B. Frey
The University of Kansas

Prepared by Karin Lindstrom Bremer
Minnesota State University, Mankato

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

FOR INFORMATION

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London, EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Printed in the United States of America

ISBN 978-1-5443-9599-9

This book is printed on acid-free paper.

Acquisitions Editor: Helen Salmon
Editorial Assistant: Megan O'Heffernan
Production Editor: Jane Martinez
Typesetter: Hurix Digital
Proofreader: Christine Dahlin
Cover Designer: Candice Harman
Marketing Manager: Shari Countryman

19 20 21 22 23 10 9 8 7 6 5 4 3 2 1

Contents

General Outline

Following is a general outline showing the sequence of content presented for each chapter.

Chapter Outline

Learning Objectives

Summary/Key Points

Key Terms

True/False Questions

Multiple-Choice Questions

Exercises

Short-Answer/Essay Questions

SPSS* Questions

Just for Fun/Challenge Yourself

Answer Key

*IBM® SPSS® Statistics is a registered trademark of International Business Machines Corporation.

1 Statistics or Sadistics? It's Up to You

CHAPTER OUTLINE

LEARNING OBJECTIVES

- Understand the purpose and scope of statistics.
- Review (briefly) the history of statistics.
- Get an introduction to descriptive and inferential statistics.
- Review the benefits of taking a statistics course.
- Learn how to use and apply this book.

SUMMARY/KEY POINTS

Introduction to Part I

- Researchers in a very wide variety of fields use statistics to make sense of the large sets of data they collect in studying a great number of interesting problems.
  - Michelle Lampl, a pediatrician and anthropologist, has studied the growth of infants, finding that some infants can grow as much as 1 inch overnight.
  - Sue Kemper, a professor of psychology, has studied the health of nuns, finding that the complexity of the nuns' writing during their early 20s is related to risk for Alzheimer's disease as many as 70 years later.
  - Aletha Huston, a researcher and teacher, has found that children who watch educational programs on television do better in school than those who don't.
- Statistics can be defined as "the science of organizing and analyzing information," making that information easier to understand.
- Statistics are used to make sense of often large and unwieldy sets of data.
- Statistics can be used in any field to answer a very wide variety of research questions and hypotheses.

A Brief History of Statistics

- Far back in human history, collecting information became an important skill.
- Once numbers became part of human language, they began to be attached to outcomes. In the 17th century, the first set of data relating to populations of people was collected.
- Once sets of data began to be collected, scientists needed to develop specific tools to answer specific questions. This led to the development of statistics.
- In the early 20th century, the simplest test for examining the differences between the averages of two groups was developed.
- The development of powerful and relatively inexpensive computers has revolutionized the field of statistics. While individuals can now conduct complex and computationally intensive statistical analyses with their own computers, they can potentially run analyses incorrectly or arrive at incorrect conclusions regarding their results.

- Today, researchers from a wide variety of fields use basically the same techniques, or statistical tests, to answer very different questions. This means that learning statistics enables you to conduct quantitative research in almost any field.

Statistics: What It Is (and Isn't)

- Statistics describes a set of tools and techniques that is used for describing, organizing, and interpreting information or data. It helps us understand the world around us.
- Descriptive statistics are used to organize and describe the characteristics of a collection of data. The collection is sometimes called a data set or just data.
- Inferential statistics are often (but not always) carried out after descriptive statistics. They are used to make inferences from a smaller group of data to a larger one. An example is using results from one kindergarten classroom to infer, or generalize, about a population of a whole kindergarten grade.
- A sample is a portion or subset of a larger population. Data from samples may be used for description only, or to generalize something about the larger population.
- A population is a full set from which a sample is taken: all the possible cases of interest. Data from a sample can be used to infer properties of a whole population.

Why Study Statistics? What Am I Doing in a Statistics Class?

- Having statistical skills puts you at an advantage when applying to graduate school or for a research or academic position.
- If not a required course for your major, a basic statistics course on your transcript sets you apart from other students.
- A statistics course can be an invigorating intellectual challenge.
- Having a knowledge of statistics makes you a better student, as it will enable you to better understand journal articles and books in your field as well as what your professors and colleagues study and discuss.
- A basic knowledge of statistics will position you well for further study if you plan to pursue a graduate degree in the social or behavioral sciences or in many other fields.

Tips for Using This Book

- Be confident: Work hard, and you'll do fine.
- Statistics is not as difficult as it's made out to be.
- Don't skip chapters: Work through them in sequence.

- Form a study group.
- Ask your professor questions.
- Do the exercises at the end of each chapter.
- Practice, practice, practice: Besides the exercises, find other opportunities to use what you've learned.
- Look for applications to make the material more real.
- Browse: Flip through the future material and review chapters.
- Have fun: Enjoy mastering a new field and acing your course.

KEY TERMS

- **Statistics**: A set of tools and techniques that is used for describing, organizing, and interpreting information or data
- **Descriptive statistics**: A set of statistical techniques and tools that is used to organize and describe data
- **Data, Data set**: A set of data points (where one data point = one observation/measurement)
- **Inferential statistics**: A set of statistical techniques and tools that is used to make inferences from a smaller group of data to a larger one
- **Sample**: A subset of the population. A researcher's goal is often to generalize findings from a sample to a population
- **Population**: All the possible subjects or cases of interest

2 Computing and Understanding Averages

Means to an End

CHAPTER OUTLINE

- ✦ Computing the Mean
- ✦ Computing the Median
- ✦ Computing the Mode
  - ✧ Apple Pie à la Bimodal
- ✦ When to Use What Measure of Central Tendency (and All You Need to Know About Scales of Measurement for Now)
  - ✧ A Rose by Any Other Name: The Nominal Level of Measurement
  - ✧ Any Order Is Fine With Me: The Ordinal Level of Measurement
  - ✧ 1 + 1 = 2: The Interval Level of Measurement
  - ✧ Can Anyone Have Nothing of Anything? The Ratio Level of Measurement
  - ✧ In Sum…
- ✦ Using SPSS to Compute Descriptive Statistics
  - ✧ The SPSS Output
- ✦ Real-World Stats
- ✦ Summary
- ✦ Time to Practice

LEARNING OBJECTIVES

- Understand averages, or measures of central tendency, one of the key components of descriptive statistics.
- Learn how to calculate the mean, median, and mode.
- Understand the distinction among these three measures of the average.
- Understand that the mean is very sensitive to outliers.
- Understand and apply scales or levels of measurement.
- Understand which measures of the average to use with different types of data.
- Learn how to use SPSS to calculate measures of central tendency.

SUMMARY/KEY POINTS

- Averages, or measures of central tendency, are used to determine the single value that best represents an entire group of scores. Popular measures of the average include the mean, median, and mode.
- The mean consists of the middle point of a set of values, and it is simply the sum of all the values divided by the number of values.
- The median consists of the middle point of a set of cases, and the mode represents the most frequent value in a set of scores.
- Only the mode can be used when determining an average for qualitative, categorical, or nominal data.
- Likewise, the median and mean can only be used with quantitative data.
- The mean can be considered the most precise measure, followed by the median and, finally, by the mode. While the mean is the most precise measure, it is sensitive to extreme scores. The median is not sensitive to extreme scores, so it better represents the center-most value of a data set that does include extreme scores.
- Scales of measurement are key to choosing the correct measure of central tendency to use. There are certain characteristics of data, specifically the scale of measurement, and each level of the scale builds upon the previous. The scale of measurement can be at the nominal, ordinal, interval, or ratio levels. The more precise levels of measurement (for example, the interval level of measurement) contain all the qualities of the scales below them.
    - **Nominal** variables are by class or category, are the least precise, and are mutually exclusive. Gender, ethnicity, and political affiliation can be nominal variables. **Ordinal** variables are measured in order or rank. The order is noted, but there is no way to tell the amount of difference between each rank. Class rankings, sports rankings, or finishers in a race could all be interval variables. **Interval** variables are frequently used in tests or assessment tools in which there is some underlying continuum that indicates the amount of a higher or lesser value. Intervals, or spaces or points, along a continuum are equal to one another. The number correct on a test could be used as an interval variable (e.g., if you get 10 correct on a 10-point vocabulary test, and someone else gets 5 correct, then you got twice as many correct). An assessment tool at the **ratio** level of measurement means that it has an absolute zero (e.g., weight or length). Ratio is the highest level of scales of measurement.

- The following are guides for when to use what scale of measurement (but there can be exceptions): Use the mode when variables are nominal. Use the median when you have extreme scores and you do not want to distort the average. Use the mean when you have data that do not include extreme scores and the variables are not categorical. The nominal level of measurement is the least precise, while the ratio level of measurement is the most precise. The "higher up" you are on the scale of measurement, the more precise, detailed, and informative your data are.
- For a sample statistic, Roman letters are used. For a population parameter, Greek letters are used.

KEY TERMS

- **Average**: The one value that best represents an entire group of scores. This can be the mean, median, or mode.
- **Measures of central tendency**: Another term for *averages*. As in the definition of average, measures of central tendency consist of the mean, median, and mode.
- **Mean**: The sum of all the values in a group, divided by the number of values in the group
  - The mean is sometimes represented using X-bar, or the letter *M*, and it is also called the typical, average, or most central score.
  - The mean is very sensitive to extreme scores.
  - When calculating the mean by hand, computing the "weighted mean" can save time when your data contain multiple instances of different values.
- **Median**: The midpoint of a set of scores
  - The median is sometimes abbreviated as *Med* or *Mdn*.
  - To compute the median, list all the values in order, from highest to lowest or lowest to highest. Next, find the middle-most score. If you have an even number of values, the median is calculated as the mean of the two middle values.
  - **Percentile points**: These are used to define the percentage of cases equal to and below a certain point in a distribution or set of scores. A score at the 25th percentile (Q_1) is at or above 25% of the other scores in the distribution. A score at the 50th percentile is often referred to as Q_2 and is the median.
  - Because the median focuses on cases, and not the values of those cases, it is much less sensitive to extreme scores, or outliers, than is the mean.
- **Mode**: The value that occurs most frequently in a set of data
  - To find the mode, first list all the values in the distribution, listing each value only once. Next, count the number of times that each value occurs. The one that occurs most often is the mode.
  - If a set of values has more than one mode, the distribution is multimodal.
  - A distribution can be multimodal even if it has multiple modes that are very similar but not exactly the same (i.e., 15 of one category and 16 of some other category).
- **Skew** (verb): When your data include too many extreme scores, the distribution of scores can become *skewed*, or significantly distorted.
- **Data points**: Individual observations in a set of data
- **Scales of measurement**: Different levels at which outcomes are measured. The four scales of measurement are nominal, ordinal, interval, and ratio.
  - **Nominal level of measurement**: The level of measurement such that outcomes can only be placed into unranked categories

- **Ordinal level of measurement**: The level of measurement such that outcomes can be rank ordered
- **Interval level of measurement**: The level of measurement such that outcomes are based on some underlying continuum that makes it possible to speak about how much greater one performance is than another
- **Ratio level of measurement**: The level of measurement such that outcomes are based on some underlying continuum that also contains a true, or absolute, zero

TRUE/FALSE QUESTIONS

1. The mode and median are both averages.
2. The mean is very sensitive to extreme scores.
3. The more precise levels of measurement (for example, the interval level of measurement) contain all the qualities of the scales below them.
4. Ordinal variables have two features. They show order or rank and they have equal distance between points along a scale.

MULTIPLE-CHOICE QUESTIONS

1. Of the set of values {170, 249, 523, 543, 572, 689, 1,050}, what is 543?
   a. The mean
   b. The median
   c. The mode
   d. The percentile
2. What is the mean of the following set of values: 1,501, 1,736, 1,930, 1,176, 446, 428, 768, and 861?
   a. 1,105.75
   b. 1,018.5
   c. 428
   d. 1,930
   e. 8
3. Your data set contains a variable on region of residence that contains the following possible responses: Northeast, South, Midwest, West, and Pacific Coast. Which of the following measures of central tendency should you use for this variable?
   a. The mean
   b. The median
   c. The mode
   d. The weighted mean
   e. Both a and b

4. This is the level of measurement where outcomes are based on some underlying continuum such that it is possible to speak about how much greater one performance is than another one.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio
5. This is the level of measurement such that outcomes can be placed only into unranked categories.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio
6. This is the level of measurement such that outcomes are based on an underlying continuum that contains a true, or absolute, zero.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio
7. This is the level of measurement such that outcomes can be rank ordered.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio
8. This is the most precise level of measurement.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio
9. This is the least precise level of measurement.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio