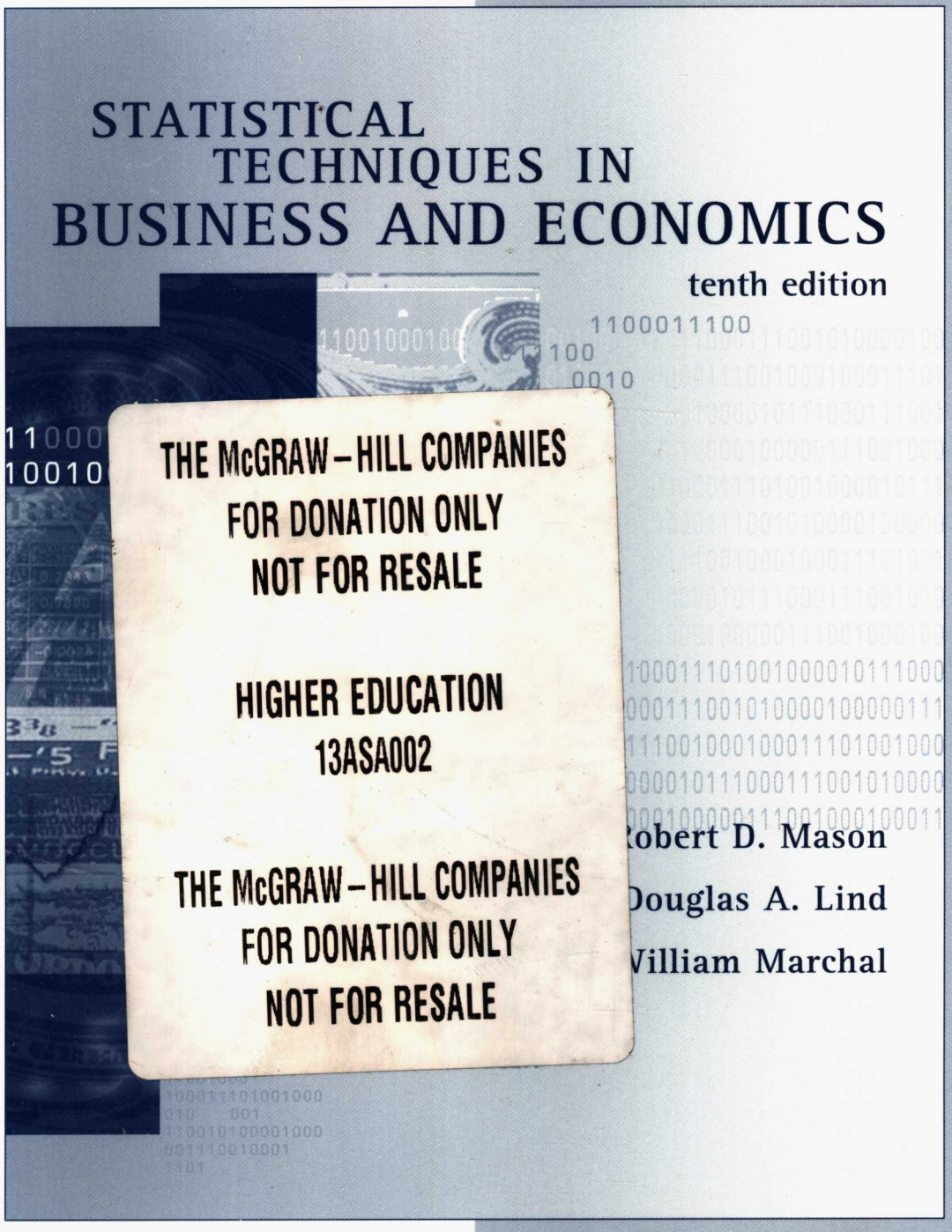
Study Guide

for use with



Prepared by
Walter H. Lange

Study Guide

for use with

Statistical Techniques in Business and Economics

Tenth Edition

Robert D. Mason
Late of The University of Toledo

Douglas A. Lind
The University of Toledo

William G. Marchal
The University of Toledo

Prepared by
Walter H. Lange
The University of Toledo

Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco St. Louis
Bangkok Bogotá Caracas Lisbon London Madrid
Mexico City Milan New Delhi Seoul Singapore Sydney Taipei Toronto

Irwin/McGraw-Hill

A Division of The McGraw-Hill Companies

Study Guide for use with
STATISTICAL TECHNIQUES IN BUSINESS AND ECONOMICS

5 6 7 8 9 0 MAZ/MAZ 9 3 2 1

ISBN 0-07-303928-4

http://www.mhhe.com

PREFACE

This study guide is especially designed to accompany the Tenth Edition of Statistical Techniques in Business and Economics by Robert D. Mason, Douglas A. Lind and William G. Marchal. It can also be used alone, or as a companion to most other introductory statistics texts. It provides a valuable source of reinforcement for the material in the text. The chapters in the text and the study guide are parallel in topics, notation, and the numbering of formulas. Students will attain the most benefit if they study the textbook first, and then read the corresponding chapter in the study guide. The major features of the study guide include:

- **Chapter Goals**. They are listed first and stress the main concepts covered and the tasks students should be able to perform after having studied the chapter. It is recommended that students refer to the goals before reading the chapter to get an overview of the material to be studied and again after completing the chapter to confirm mastery of the material.

- **Brief Introduction.** A brief Introduction follows the goals. In capsule form the material covered in previous chapters is tied with that covered in the current chapter, thus maintaining continuity throughout the book.

- **Definitions**. Key words are defined and used in their correct statistical context.

Key Words are in a text box for easy reference.

- **Formulas.** The formulas are placed in a formula box for easy reference.

Formula box is used to emphasize formulas.

- A **glossary** follows the chapter discussion. The glossary provides definitions of the key words used in the chapter and is a handy reference.

- **Chapter Problems.** Chapter problems, including solutions, come next. In this section the step-by-step method of solution is presented along with an interpretation of the results. The values are kept small to emphasize the concept.

- **Exercises.** Following the chapter problems is an exercise. The student completes the exercise and checks the answer in the answer section at the end of the guide. Thus the student can check his/her comprehension of the material as they progress through the chapter.

- **Chapter Assignments.** Chapter assignments cover the entire chapter and are intended to be completed outside the classroom. Part I of the assignment consists of multiple-choice questions, Part II is problems, with space for students to show essential work and a box for the answers. The pages are perforated, so that assignments can be torn out and handed in to the instructor for grading.

For this revised edition we wish to thank Temelon G. Rousos for his invaluable assistance in reviewing the manuscript and checking the accuracy of the solutions to the problems, the exercises, and the assignments. A special thanks to Danuta T. Lange who prepared the camera-ready copy for this publication.

Walter H. Lange
Douglas A. Lind

TABLE OF CONTENTS

CHAPTER 1
WHAT IS STATISTICS?

Chapter Goals

After completing this chapter, you will be able to:

1. Define what is meant by statistics.
2. Explain what is meant by descriptive statistics and inferential statistics.
3. Distinguish between a qualitative variable and a quantitative variable.
4. Distinguish between a discrete variable and a continuous variable.
5. Distinguish between nominal, ordinal, interval, and ratio levels of measurement.
6. Define the terms mutually exclusive and exhaustive.

Introduction

No doubt you have noticed the large number of facts and figures, often referred to as ***statistics***, that appear in the newspapers and magazines you read, the television you watch (especially sporting events) and in the grocery stores where you shop. A simple figure is called ***statistic*** (singular). A few examples:

- The best performing stock last year was Yahoo. It went up 301% in 1997. (***Smart Money***, February, 1998).
- The worst performing stock last year was Molten Metal Technology. It was down 98% in 1997. (***Smart Money***, February, 1998).
- During the summer as many as 6,000 cars a day go through the south entrance of Grand Canyon National Park. (***Highways***, February 1998).
- The International Federation of Organic Agricultural Movements is a 570-member organization in more than 100 countries. (***Organic Gardening***, March 1998).
- Stock prices recently jumped to record highs this week. The Dow Jones Industrial average closed at 8546. (***Wall Street Journal, March*** 2, 1998)

The Yahoo stock price increase of 301 percent is a statistic (singular). The Dow Jones average of 8546 is a statistic. A collection of figures is called statistics (plural). An example from the March 2, 1998 ***USA Today*** daily feature on the "Markets" is shown:

Markets	
Dow Jones industrial average	8545.72
Nasdaq composite	1770.51
T-bond, 30 year yield	5.92%
Gold, oz. Comex	$299.10
U.S. dollar, yen per dollar	126.10

You may think of statistics simply as a collection of numerical information. However, ***statistics*** has a much broader meaning.

Statistics: The science of collecting, organizing, presenting analyzing, and interpreting data to assist in making more effective decisions.

Note in this definition of statistics that the initial step is the collection of pertinent information. This information may come from newspapers or magazines, the company's human relations director, the local, state, or federal government, universities, nonprofit organizations, the United Nations, and so on. A few actual publications of the federal government and others are:

- ***Statistical Abstract of the United States***, published annually by the U.S. Department of Commerce.
- ***Monthly Labor Review***, published monthly by the U.S. Department of Labor.
- ***Survey of Current Business***, published monthly by the U.S. Department of Commerce.
- ***Social Security Bulletin***, published annually by the U.S. Social Security Administration.
- ***Crime in the United States***, published annually by the U.S. Federal Bureau of Investigation.
- ***Hospital Statistics***, published annually by the American Hospital Association.
- ***Vital Statistics of the United States***, published annually by the National Center for Health Statistics.

If the information is not available from company records or public sources, it may be necessary to conduct a ***survey.*** For example, the A.C. Nielsen Company surveys about 1200 homes on an ongoing basis to determine which TV programs are being watched, and Gallup surveys registered voters before an election to estimate the percent that will vote for a certain candidate. These firms also sample the population regarding food preference, what features in automobiles are desirable, and what appliances consumers will most likely purchase next year.

Fortune annually surveys 12,600 senior executives, outside directors, and securities analysts to evaluate the companies in their industry to find the ten most admired firms, and the least admired firms. Each executive is asked to rate a list of firms on eight attributes, namely innovativeness, quality of management, quality of products and services, long-term investment value, financial soundness, employee talent, social responsibility to the community and the environment and wise use of corporate assets. Each attribute is rated on a scale of zero (poor) to ten (excellent). The ten most admired companies are listed in the table.

Rank	Company
1	General Electric
2	Microsoft
3	Coca Cola
4	Intel
5	Hewlett-Packard
6	Southwest Airlines
7	Berkshire Hathaway
8	Disney
9	Johnson & Johnson
10	Merck

Source: ***Fortune***, March 2, 1998

Why Study Statistics?

Statistics is required for many college programs for three reasons.

1. Numerical information is everywhere. If you look in various newspapers (***USA Today, Wall Street Journal***), magazines (***Time, Business Week, Sports Illustrated, People***) you will be bombarded with numerical information. You need to be able to determine if the conclusions as reported are reasonable. Was the sample large enough? You must be able to read and interpret the charts or graphs.
2. Statistical techniques are used to make decisions that affect our lives. Insurance companies use statistics to determine the premiums you pay for automobile insurance, the Environmental Protection Agency uses various statistical tools to determine air quality in your area, and the Internal Revenue Service uses statistical surveys to determine if your tax return should be subject to an audit.

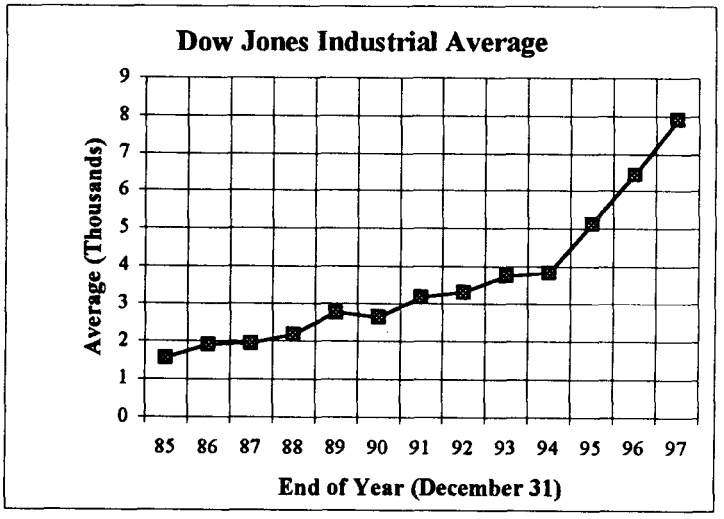
3. Knowledge of statistical methods will help you understand why decisions are made and give you a better understanding of how they affect you.

Descriptive and Inferential Statistics

The definition of statistics referred to collecting, organizing, and presenting numerical information. Data stored in a computer's memory or in a filing cabinet are of little value. Techniques are available that organize this information in a more meaningful form. Such aids are called ***descriptive statistics***.

> ***Descriptive statistics***: Methods of organizing, summarizing, and presenting data in an informative way.

A statistical tool designed to describe the movement of a series of numbers over a long period of time (such as production, imports, wages and stock market trends) is called a line chart. The line chart below, for example, depicts the upward movement of the Dow Jones average of 30 industrials year-end closing prices since 1985.



Notice how easy it is to describe the trend of stock prices: The price of the 30 industrials, as represented by the Dow, rose somewhat steadily from about 1800 in 1987 to over 8500 in early 1998.

Another descriptive measure is referred to as an average. Some examples are:

- The 1998 Toyota Sienna minivan averages 12 miles per gallon in city driving, 30 miles per gallon highway driving and 19 miles per gallon overall. (***Consumers Reports***, March, 1998)
- The average price of vehicles wholesaled at United States automobile auctions in 1997 was $9,992, down from an average price of $10,069 in 1996. (***NADA/NAAA Auction Net***, February, 1998)
- Herman Warsaw set the world record in 1985 by averaging 70 bushels of corn per acre on a 100.4 acre field. (***Farm Journal***, January, 1996)
- The median (an average) price of homes sold in 1997 in Lucas County was $64,500 (***The Blade***, January 18, 1998)

The Bureau of Labor Statistics describing the labor force in the United States reported that the average number of employed persons in 1997 was 133,900,000 and the average number of unemployed people was 5,400,000. Averages and other descriptive measures are presented in Chapter 3.

A second aspect of statistics is called ***inferential statistics***.

> ***Inferential statistics***: The methods used to find out something about a population, based on a sample.

This branch of statistics deals with problems requiring the use of a sample to infer something about the ***population***.

> ***Population***: A collection of all possible individuals, objects, or measurements of interest.

A population might consist of all the 7,425,000 people in North Carolina, or all 480,000 people in Wyoming. Or, the population might consist of all the teams in the Canadian Football League, the PE ratios for all chemical stocks, or the total assets of the 20 largest banks in the United States. A population, therefore, can be considered the total collection of people, prices, ages, square footage of homes being constructed in Flint, Michigan in 1998, and so on.

To infer something about a population, we usually take a ***sample*** from the population.

> ***Sample***: A portion, or part, of the population of interest.

A sample might consist of 2,000 people out of the 32,268,000 people in California, 12 headlights selected from a production run of 1000 for a life test, or the three scoops of grain selected at random to be tested for moisture content from a 15-ton truckload of grain. If we found that the three scoops of grain consisted of 9.50 percent moisture, we would infer that all the grain in the 15-ton load had 9.50 percent moisture. We start our discussion of inferential statistics in Chapter 5.

Types of Variables

There are two types of variables, ***quantitative*** and ***qualitative***.

> ***Qualitative variable***: A variable that has the characteristic of being nonnumeric.

A classification of students at your university by the state of birth, gender, or college affiliation (Business, Education, Liberal Arts, etc.) is an example of a qualitative variable.

> ***Quantitative variable***: A variable being studied that can be reported numerically.

Examples of quantitative variables include: the balance in your checking account, the ages of the members of the United States Congress, the speeds of automobiles traveling along I-70 in Kansas, the number of customers served in the Commodore Barber Shop last week, or the number of new single family homes constructed by Reynolds Construction Company last year in Erie, Pennsylvania.

There are two types of quantitative variables, ***discrete*** and ***continuous***.

> ***Discrete variable***: A quantitative variable that can only assume certain values. There is usually a "gap" between the values.

Examples of discrete variables are: the number of children in a family, the number of customers in a carpet store in an hour, or the number of commercials aired last hour on radio station WEND. A family can have two or three children, but not 2.445, or WEND can air five or six commercials, but not 5.75. Usually discrete variables result from counting.

> ***Continuous variable:*** A quantitative variable that can assume any value within a range.

Examples of continuous variables are: the amount of snow for the winter of 1998-99 in Toronto, Ontario, the pressure in a tire, or a person's weight. Typically, continuous variables are the result of measuring something. We can measure the pressure in a tire, or the amount of snow in Toronto.

Levels of Measurement

Data may be classified into four categories or levels of measurement. These categories are nominal, ordinal, interval, and ratio.

Nominal Level Data

When data can only be classified into categories, we refer to it as being the ***nominal*** level of measurement.

> ***Nominal level***: A level of measurement in which the data are sorted into categories with no particular order to the categories.

At this level the categories have no particular order or rank and are ***mutually exclusive.***

> ***Mutually exclusive***: An individual, object, or measurement is included in only one category.

For example, the Office of Special Education, U.S. Department of Education, gave these counts of the number of handicapped children 3 to 21 years old who were in special education programs.

Type of Handicap	Number Served
Visual impairments	23,000
Serious emotional impairment	401,000
Speech impairments	996,000
Deaf-blindness	2,000
Learning disabilities	2,354,000
Mental retardation	519,000
Hearing impairments	60,000
Orthopedic impairments	52,000
All others	718,000
Total	5,125,000

The data is nominal level of measurement because it can only be classified into categories and it is immaterial what order the type of handicap is listed. Mental retardation could be listed first, serious emotional impairment second, and so on. The categories are mutually exclusive meaning that the type of handicap a child has can be counted into only one category. And such categories are said to be ***exhaustive***.

> ***Exhaustive:*** Each individual, object, or measurement must appear in a category.

This means that a handicapped child enrolled in the program must appear in one of the categories. Chapter 14 deals with data that is nominal level of measurement.

To summarize, nominal level data has these properties:

- Data categories are mutually exclusive, so an object belongs to only one category.
- Data categories have no logical order.

Ordinal Level Data

The ***ordinal*** level of measurement implies some sort of ranking.

> ***Ordinal level***: A level of measurement that presumes that one category is ranked higher than another category.

An example of ordinal level of measurement follows: undergraduate students in a basic statistics class were classified according to class rank.

Class Rank	Number
Freshman	13
Sophomore	17
Junior	9
Senior	5

Note that the categories are mutually exclusive meaning that a student can only be counted in one category. A student cannot be a freshman and a sophomore at the same time. Also, the categories are exhaustive meaning that a student must appear in one of the categories. In addition, a ranking of students is implied meaning that juniors are ranked "higher" than sophomores. Chapter 15 deals with tests involving ordinal level of measurement.

To summarize, ordinal level data has these properties:

- Data categories are mutually exclusive and exhaustive.
- Data categories have some logical order.
- Data categories are ranked or ordered according to the particular trait they possess.

Interval Level Data

The ***interval level*** of measurement is the next highest level.

> ***Interval level:*** Includes the ranking characteristics of the ordinal scale and, in addition, the distance between values is a constant size.

Temperature on the Fahrenheit scale is an example. Suppose the high temperature for the last three days was 85, 73, and 78 degrees Fahrenheit. We can easily put the readings in a rank order, but in addition we can study the difference between readings. Why is this so? One degree on the Fahrenheit temperature scale is a constant unit of measure for all three days. Note in this example that the zero point is just another point on the scale. It does not represent the absence of temperature, just that it is cold! Test scores are another example of the interval scale of measurement.

In addition to the constant difference characteristic, interval scaled data have all the features of nominal and ordinal measurements. Temperatures are mutually exclusive, that is, the high temperature yesterday cannot be both 88 and 85 degrees. The "greater than" feature of ordinal data permits the ranking of daily high temperatures.

The properties of the interval scale are:

- Data categories are mutually exclusive.
- Data categories are scaled according to the amount of the characteristic they possess.
- Equal differences in the characteristic are represented by equal differences in the numbers assigned to the categories.

Ratio Level Data

The ***ratio level*** of measurement is the highest level of measurement.

> ***Ratio level***: Has all the characteristics of the interval scale, but additionally there is a meaningful zero point and the ratio of two values is meaningful.

Weight, height, and money are examples of the ratio scale of measurement. If you have $20 and your friend has $10, then you have twice as much money as your friend. The zero point represents the absence of money. That is, the zero point is fixed and represents the absence of the characteristic being measured. If you have zero dollars, you have none of the characteristic being measured.

The properties of the ratio level are:

- Data categories are mutually exclusive.
- Data categories are scaled according to the amount of the characteristic they posses.
- Equal differences in the characteristic are represented by equal differences in the numbers assigned to the categories.
- The point 0 reflects the absence of the characteristic.

Glossary

Statistics: The science of collecting, organizing, presenting, analyzing, and interpreting data to assist in making more effective decisions.

Descriptive statistics: Methods of organizing, summarizing, and presenting data in an informative way.

Inferential statistics: The methods used to find out something about a population, based on a sample.

Population: A collection of all possible individuals, objects, or measurements of interest.

Sample: A portion, or part, of the population of interest.

Qualitative variable: A variable that has the characteristic of being nonnumeric.

Quantitative variable: A variable being studied that can be reported numerically.

Discrete variable: A quantitative variable that can only assume certain values. There is usually a "gap" between the values.

Continuous variable: A quantitative variable that can assume any value within a range.

Nominal level: A level of measurement in which the data are sorted into categories with no particular order to the categories.

Mutually exclusive: An individual or object can be included in only one category.

Exhaustive: Each individual, object, or measurement must appear in a category.

Ordinal level: A level of measurement that presumes that one category is ranked higher than another category.

Interval level: Includes the ranking characteristics of the ordinal scale and, in addition, the distance between values is a constant size.

Ratio level: Has all the characteristics of the interval scale, but additionally there is a meaningful zero point and the ratio of two values is meaningful.

CHAPTER 1 ASSIGNMENT

WHAT IS STATISTICS?

Name _____ Section _____ Score _____

Part I Classify the following sets of data as qualitative or quantitative.

_____ 1. The height of each member of a basketball team

_____ 2. The religious affiliations of school faculty members

_____ 3. Scores of students on the first statistics exam

_____ 4. The Olympic track and field world records

_____ 5. The color of Labrador puppies in a litter

Part II Classify the following sets of data as continuous or discrete.

_____ 6. The number of sit-ups

_____ 7. The acceleration of an automobile

_____ 8. The number of pairs of shoes sold at a shoe store

_____ 9. The temperature of an oven

_____ 10. The diving depth of a submarine

Part III Identify the measurement scale for each of the following.

_____ 11. Classification of automobile make

_____ 12. The temperature readings in Nome, Alaska

_____ 13. College major

_____ 14. Number of traffic fatalities

_____ 15. Military rank

_____ 16. Time required to complete a crossword puzzle

_____ 17. Order of finish in the 1996 Glass City Marathon

_____ 18. Number of people at a business meeting

_____ 19. Years in which Huntington Bank stock split

_____ 20. The hair color of employees at Jones City Hardware

Part IV Select the correct answer and write the appropriate letter in the space provided.

_____ 21. The collection of all possible individuals, objects, or measurements is called

a. a sample.
b. a ratio measurement.
c. a population.
d. an inference.

_____ 22. Techniques used to organize, summarize, and present the data that have been collected are called

a. populations.
b. samples.
c. inferential statistics.
d. descriptive statistics.

_____ 23. An individual, measurement, or object that can appear in only one category is said to be

a. mutually exclusive.
b. exhaustive.
c. inferential.
d. descriptive.

_____ 24. Techniques used to determine something about a population, based on a sample, are called

a. descriptive statistics.
b. inferential statistics.
c. populations.
d. samples.

_____ 25. A difference between the interval scale and the ratio scale is

a. the interval scale cannot be ranked.
b. the zero point on the interval scale is arbitrary.
c. the ratio scale does not meet the exhaustive criteria.
d. the interval scale does not meet the mutually exclusive criteria.

CHAPTER 2
DESCRIBING DATA: FREQUENCY DISTRIBUTIONS AND GRAPHIC PRESENTATION

Chapter Goals

After completing this chapter, you will be able to:

1. Organize data into a frequency distribution.
2. Portray a frequency distribution in a histogram, a frequency polygon, and a cumulative frequency polygon.
3. Develop a stem-and-leaf display.
4. Present data using such graphic techniques as line charts, bar charts, and pie charts.

Introduction

This chapter begins our study of ***descriptive statistics.*** Recall from Chapter 1 that when using descriptive statistics we merely describe a set of data. For example, we want to describe the entry level salary for a select group of professions. We find that the entry level salary for accountants is $28,000, for systems analysts $30,000, for infectious disease specialists $70,000, and so on. This unorganized data provides little insight into the pattern of entry level salaries which makes conclusions difficult.

This chapter presents a technique that is used to organize raw data into some meaningful form. It is called a ***frequency distribution.*** Then, to better understand the main features of the data, we portray the frequency distribution will in the form of a frequency polygon, a histogram, or a cumulative frequency polygon.

Frequency Distributions

A ***frequency distribution*** is a useful statistical tool for organizing a mass of data into some meaningful form.

> ***Frequency Distribution***: A grouping of data into categories showing the number of observations in each mutually exclusive category.

As noted, a frequency distribution is used to summarize and organize large amounts of data.

The steps to follow in developing a frequency distribution are:

1. Decide on the number of classes or the class interval.
2. Tally the observations into the appropriate classes.
3. Count the number of tallies in each class.

As an example, the lengths of service, in years, of a sample of eleven employees are given.

Length of Service (in years)					
4	3	2	10	6	6
5	8	4	8	4	

The eleven observations are referred to as **raw data** or ungrouped data. To organize the lengths of service into a frequency distribution we first set up groups called **classes.** We decided to use classes 1 up to 3, 3 up to 5, and so on. Then we **tally** the lengths of service into the appropriate classes. Finally, we count the number of tallies in each class as shown.

Lengths of service	Tallies	Number of employees
1 up to 3 years	/	1
3 up to 5 years	////	4
5 up to 7 years	///	3
7 up to 9 years	//	2
9 up to 11 yrs.	/	1
Total		11

How many classes should there be? A common guideline is from 5 to 15. Having too few or too many classes gives little insight into the data. The size of the class interval may be a value such as 3, 5, 10, 15, 20, 50, 100, 1,000, and so on.

Class Interval: The size or width of the class.

The class interval can be approximated by text formula [2-1]

$$\text{Class Interval}\,(i) = \frac{\text{highest value} - \text{lowest value}}{\text{number of classes}} \qquad [2-1]$$

Each class has a lower class limit and an upper class limit. The lower limit of the first class is usually slightly below the smallest value and is a multiple of the class interval.

In the previous example, the smallest number of years of service is 2. Therefore, we selected 1, which is slightly below 2, as the lower limit of the first class. The lower limit of the second class is 3 years, and so on.

The number of tallies that occurs in each class is called the ***class frequency***.

Class frequency: The number of observations in each class.

In the example, the class frequency of the lowest class is 1. For the next higher class it is 4. The class midpoint divides a class into two equal parts.

Class midpoint: The point halfway between the upper and lower limit of a class.

Note that the class midpoint is also called the ***class mark***.

In the example, the class midpoint of the 5 up to 7 class is 6 found by (5 + 7)/2. The class interval is the distance between the lower limit of two consecutive classes. It is 2 found by subtracting 1 (the lower limit of the first class) from 3 (the lower limit of the second class).