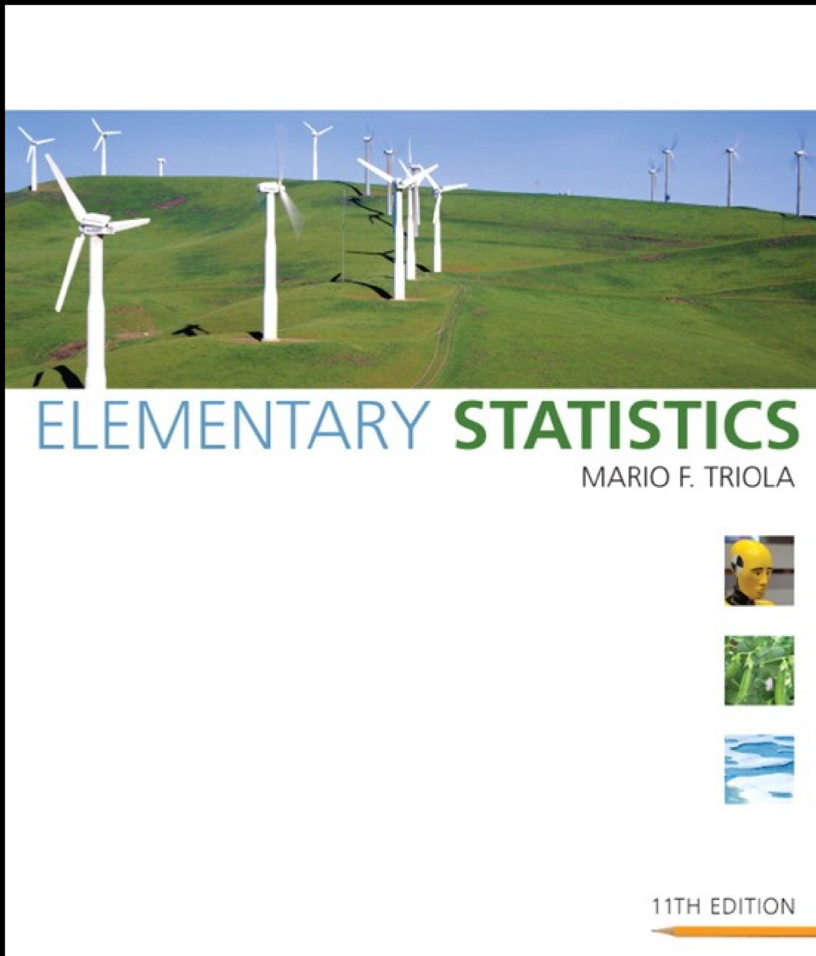
Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 1
Introduction to Statistics

1-1 Review and Preview

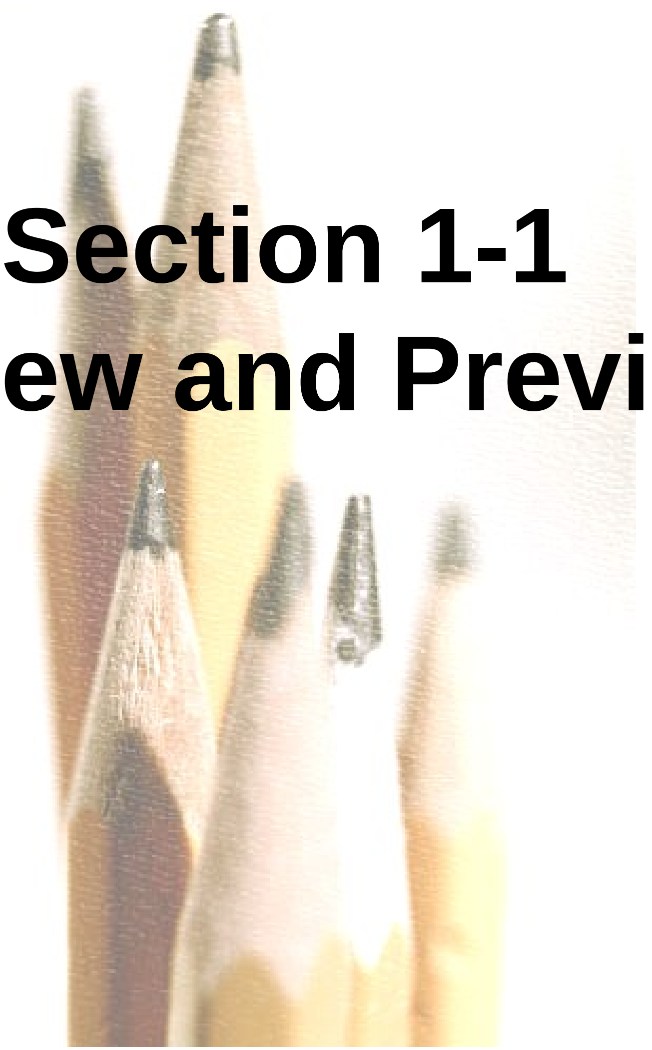
1-2 Statistical Thinking

1-3 Types of Data

1-4 Critical Thinking

1-5 Collecting Sample Data

Section 1-1
Review and Preview



Preview

Polls, studies, surveys and other data collecting tools collect data from a small part of a larger group so that we can learn something about the larger group. This is a common and important goal of statistics: Learn about a large group by examining data from some of its members.

Preview

In this context, the terms sample and population have special meaning. Formal definitions for these and other basic terms will be given here.

In this section we will look at some of the ways to describe data.

Data

- **Data**

  collections of observations (such as measurements, genders, survey responses)

Statistics

- **Statistics**

  is the science of planning studies and experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting, and drawing conclusions based on the data

Population

- **Population**

  the complete collection of all individuals (scores, people, measurements, and so on) to be studied; the collection is complete in the sense that it includes *all* of the individuals to be studied

Census versus Sample

- **Census**

  Collection of data from *every* member of a population

- **Sample**

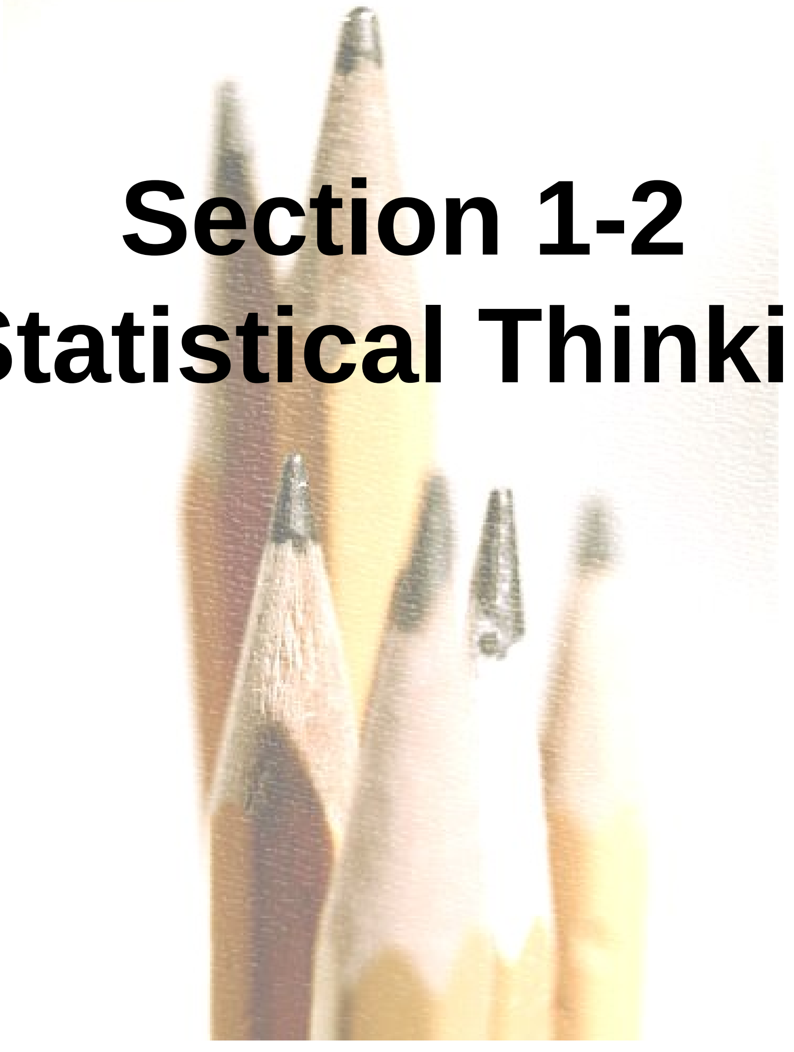
  ***Subcollection* of members selected from a population**

Chapter Key Concepts

- **Sample data must be collected in an appropriate way, such as through a process of *random* selection.**

- **If sample data are not collected in an appropriate way, the data may be so completely useless that no amount of statistical torturing can salvage them.**

Section 1-2
Statistical Thinking



Key Concept

This section introduces basic principles of statistical thinking used throughout this book. Whether conducting statistical analysis of data that we have collected, or analyzing a statistical analysis done by someone else, we should not rely on blind acceptance of mathematical calculation. We should consider these factors:

Key Concept (continued)

- **Context of the data**
- **Source of the data**
- **Sampling method**
- **Conclusions**
- **Practical implications**

Context

- What do the values represent?
- Where did the data come from?
- Why were they collected?
- An understanding of the context will directly affect the statistical procedure used.

Source of data

- Is the source objective?
- Is the source biased?
- Is there some incentive to distort or spin results to support some self-serving position?
- Is there something to gain or lose by distorting results?
- Be vigilant and skeptical of studies from sources that may be biased.

Sampling Method

- Does the method chosen greatly influence the validity of the conclusion?
- Voluntary response (or self-selected) samples often have bias (those with special interest are more likely to participate). These samples' results are not necessarily valid.
- Other methods are more likely to produce good results.

Conclusions

- **Make statements that are clear to those without an understanding of statistics and its terminology.**
- **Avoid making statements not justified by the statistical analysis.**

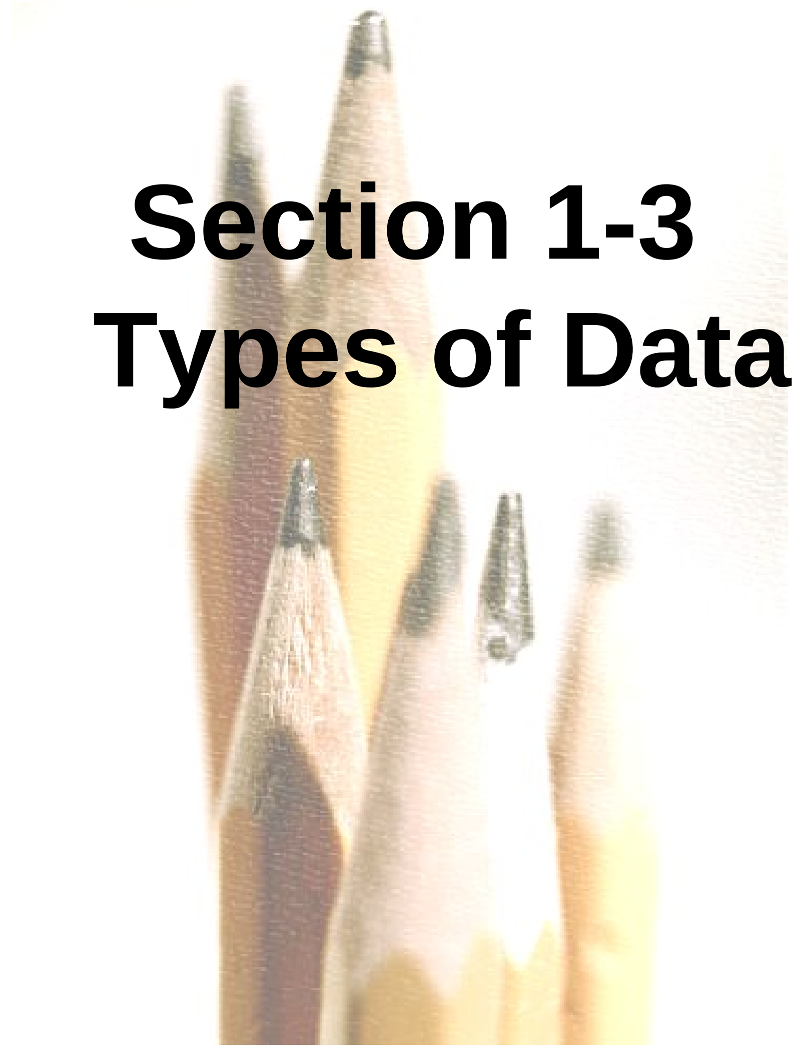
Practical Implications

- **State practical implications of the results.**
- **There may exist some *statistical significance* yet there may be NO *practical significance*.**
- **Common sense might suggest that the finding does not make enough of a difference to justify its use or to be practical.**

Statistical Significance

- Consider the likelihood of getting the results by chance.
- If results could easily occur by chance, then they are ***not statistically significant***.
- If the likelihood of getting the results is so small, then the results are ***statistically significant***.

Section 1-3
Types of Data



Key Concept

The subject of statistics is largely about using sample data to make inferences (or generalizations) about an entire population. It is essential to know and understand the definitions that follow.

Parameter

- **Parameter**

  a numerical measurement describing some characteristic of a population.

population

↕

parameter

Statistic

- **Statistic**

  a numerical measurement describing some characteristic of a sample.

sample

⇕

statistic

Quantitative Data

- **Quantitative (or numerical) data**

consists of *numbers* representing counts or measurements.

Example: The weights of supermodels

Example: The ages of respondents

Categorical Data

- **Categorical (or qualitative or attribute) data**

  consists of names or labels (representing categories)

Example: The genders (male/female) of professional athletes

Example: Shirt numbers on professional athletes uniforms - substitutes for names.

Working with Quantitative Data

Quantitative data can further be described by distinguishing between discrete and continuous types.

Discrete Data

- **Discrete data**

  result when the number of possible values is either a finite number or a ‘countable’ number

  (i.e. the number of possible values is

  0, 1, 2, 3, . . .)

  Example: The number of eggs that a hen lays

Continuous Data

- **Continuous (numerical) data**

  result from infinitely many possible values that correspond to some continuous scale that covers a range of values without gaps, interruptions, or jumps

  Example: The amount of milk that a cow produces; e.g. 2.343115 gallons per day

Levels of Measurement

Another way to classify data is to use levels of measurement. Four of these levels are discussed in the following slides.

Nominal Level

- **Nominal level of measurement**

  characterized by data that consist of names, labels, or categories only, and the data cannot be arranged in an ordering scheme (such as low to high)

  Example: Survey responses yes, no, undecided

Ordinal Level

- **Ordinal level of measurement**

involves data that can be arranged in some order, but differences between data values either cannot be determined or are meaningless

Example: Course grades A, B, C, D, or F

Interval Level

- **Interval level of measurement**

  like the ordinal level, with the additional property that the difference between any two data values is meaningful, however, there is no natural zero starting point (where none of the quantity is present)

  Example: Years 1000, 2000, 1776, and 1492

Ratio Level

- **Ratio level of measurement**

  the interval level with the additional property that there is also a natural zero starting point (where zero indicates that none of the quantity is present); for values at this level, differences and ratios are meaningful

  Example: Prices of college textbooks ($0 represents no cost, a $100 book costs twice as much as a $50 book)

Summary - Levels of Measurement

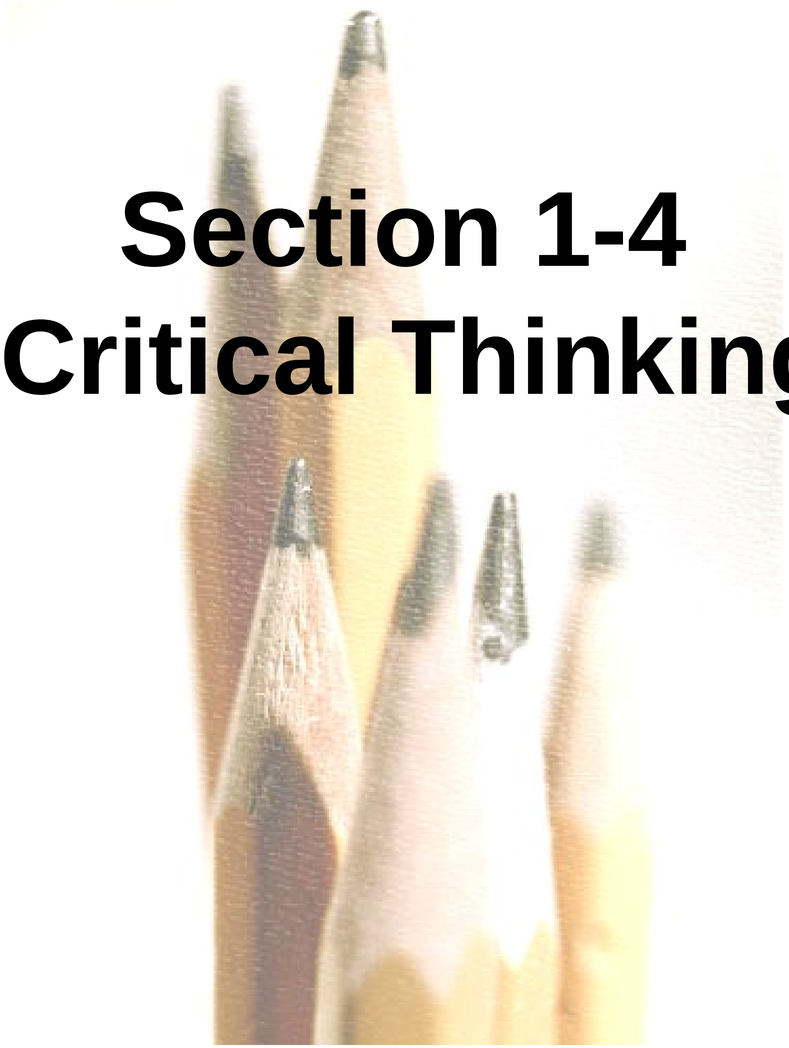
- **Nominal** - categories only
- **Ordinal** - categories with some order
- **Interval** - differences but no natural starting point
- **Ratio** - differences and a natural starting point

Recap

In this section we have looked at:

- **Basic definitions and terms describing data**
- **Parameters versus statistics**
- **Types of data (quantitative and qualitative)**
- **Levels of measurement**

Section 1-4
Critical Thinking



Key Concepts

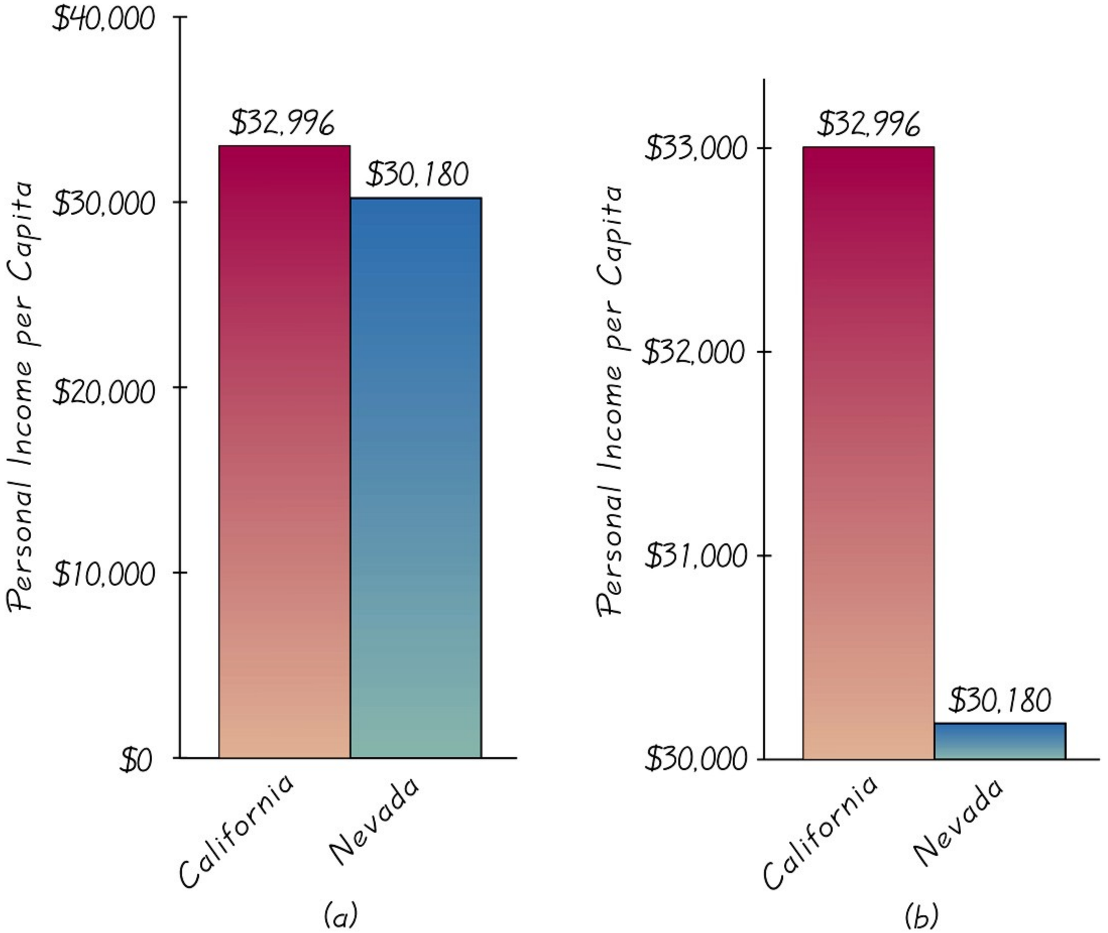
- Success in the introductory statistics course typically requires more common sense than mathematical expertise.
- Improve skills in interpreting information based on data.
- This section is designed to illustrate how common sense is used when we think critically about data and statistics.
- Think carefully about the context, source, method, conclusions and practical implications.

Misuses of Statistics

1. Evil intent on the part of dishonest people.
2. Unintentional errors on the part of people who don't know any better.

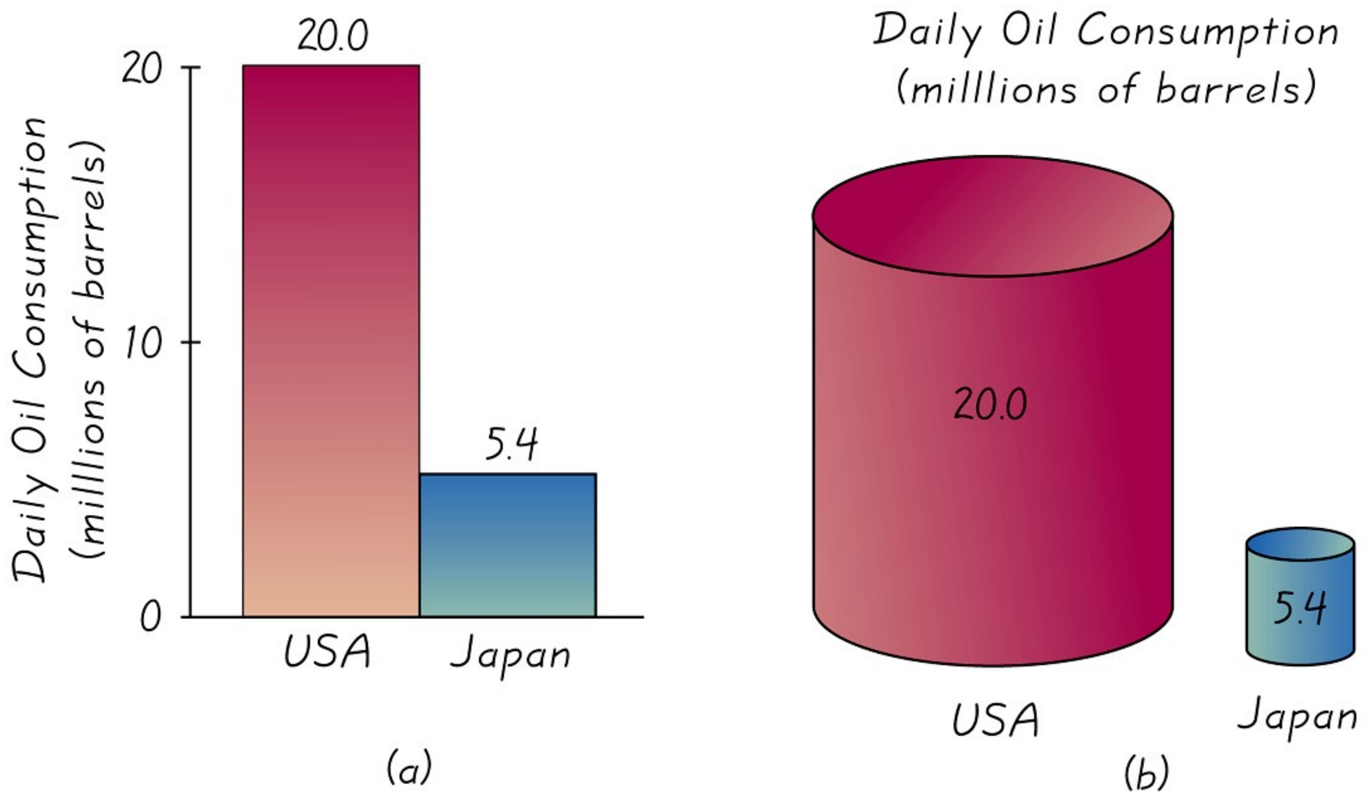
We should learn to distinguish between statistical conclusions that are likely to be valid and those that are seriously flawed.

Graphs



To correctly interpret a graph, you must analyze the numerical information given in the graph, so as not to be misled by the graph's shape. READ labels and units on the axes!

Pictographs



Part (b) is designed to exaggerate the difference by increasing each dimension in proportion to the actual amounts of oil consumption.

Bad Samples

- **Voluntary response sample (or self-selected sample)**

one in which the respondents themselves decide whether to be included

In this case, valid conclusions can be made only about the specific group of people who agree to participate and not about the population.

Correlation and Causality

- **Concluding that one variable *causes* the other variable when in fact the variables are linked**

Two variables may seemed linked, smoking and pulse rate, this relationship is called correlation. Cannot conclude the one causes the other.
***Correlation does not imply causality*.**

Small Samples

Conclusions should not be based on samples that are far too small.

Example: Basing a school suspension rate on a sample of only three students

Percentages

Misleading or unclear percentages are sometimes used. For example, if you take 100% of a quantity, you take it all. If you have improved 100%, then are you perfect?! 110% of an effort does not make sense.

Loaded Questions

If survey questions are not worded carefully, the results of a study can be misleading. Survey questions can be "loaded" or intentionally worded to elicit a desired response.

Too little money is being spent on "welfare" versus too little money is being spent on "assistance to the poor." Results: 19% versus 63%

Order of Questions

Questions are unintentionally loaded by such factors as the order of the items being considered.

Would you say traffic contributes more or less to air pollution than industry?
Results: traffic - 45%; industry - 27%

When order reversed.
Results: industry - 57%; traffic - 24%

Nonresponse

Occurs when someone either refuses to respond to a survey question or is unavailable.

People who refuse to talk to pollsters have a view of the world around them that is markedly different than those who will let poll-takers into their homes.

Missing Data

Can dramatically affect results.

Subjects may drop out for reasons unrelated to the study.

People with low incomes are less likely to report their incomes.

US Census suffers from missing people (tend to be homeless or low income).

Self-Interest Study

Some parties with interest to promote will sponsor studies.

Be wary of a survey in which the sponsor can enjoy monetary gain from the results.

When assessing validity of a study, always consider whether the sponsor might influence the results.

Precise Numbers

Because as a figure is precise, many people incorrectly assume that it is also *accurate*.

A precise number can be an estimate, and it should be referred to that way.

Deliberate Distortion

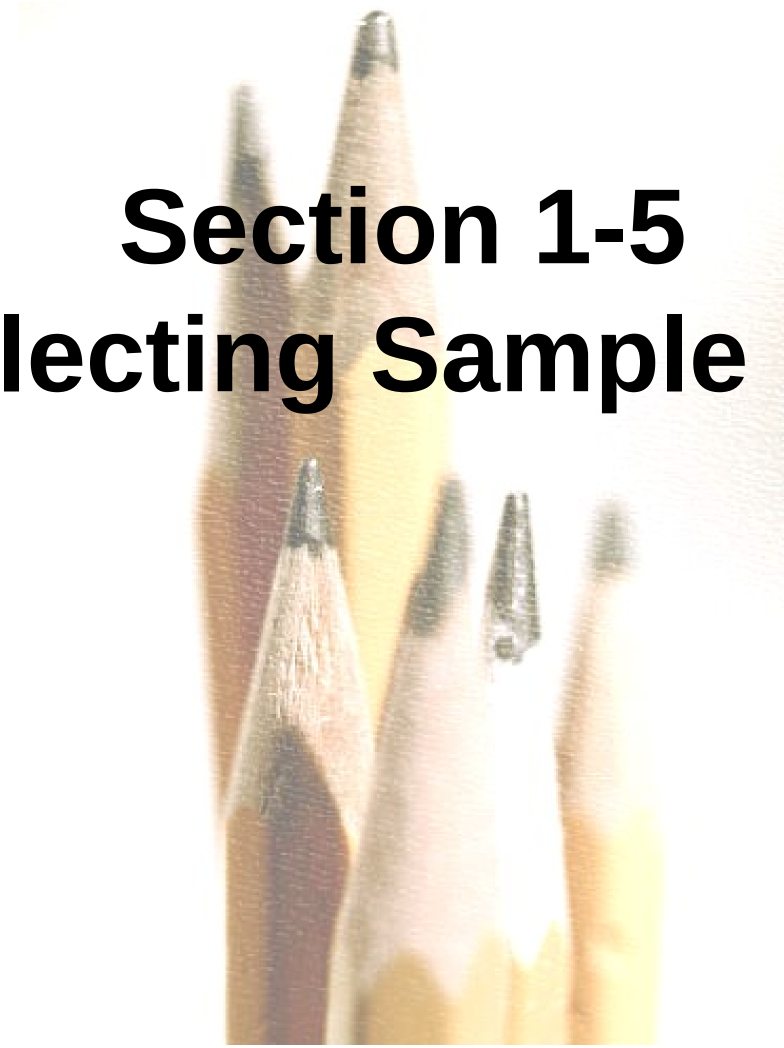
Some studies or surveys are distorted on purpose. The distortion can occur within the context of the data, the source of the data, the sampling method, or the conclusions.

Recap

In this section we have:

- **Reviewed misuses of statistics**
- **Illustrated how common sense can play a big role in interpreting data and statistics**

Section 1-5
Collecting Sample Data



Key Concept

- **If sample data are not collected in an appropriate way, the data may be so completely useless that no amount of statistical torturing can salvage them.**
- **Method used to collect sample data influences the quality of the statistical analysis.**
- **Of particular importance is *simple random sample*.**

Basics of Collecting Data

Statistical methods are driven by the data that we collect. We typically obtain data from two distinct sources: *observational studies* and *experiment*.

Observational Study

- **Observational study**

  observing and measuring specific characteristics without attempting to modify the subjects being studied

Experiment

- **Experiment**

  apply some treatment and then observe its effects on the subjects; (subjects in experiments are called experimental units)

Simple Random Sample

- **Simple Random Sample**

of n subjects selected in such a way that every possible sample of the same size n has the same chance of being chosen

Random & Probability Samples

- ## Random Sample

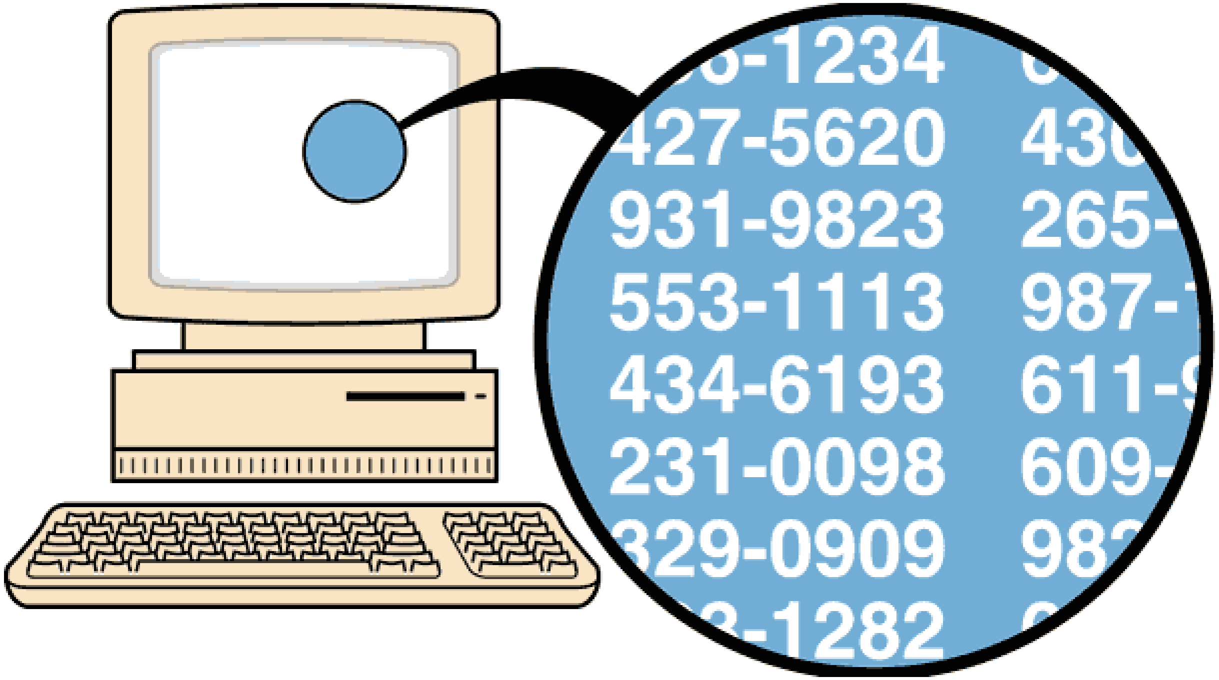
  members from the population are selected in such a way that each individual member in the population has an equal chance of being selected

- ## Probability Sample

  selecting members from a population in such a way that each member of the population has a known (but not necessarily the same) chance of being selected

Random Sampling

selection so that each individual member has an equal chance of being selected

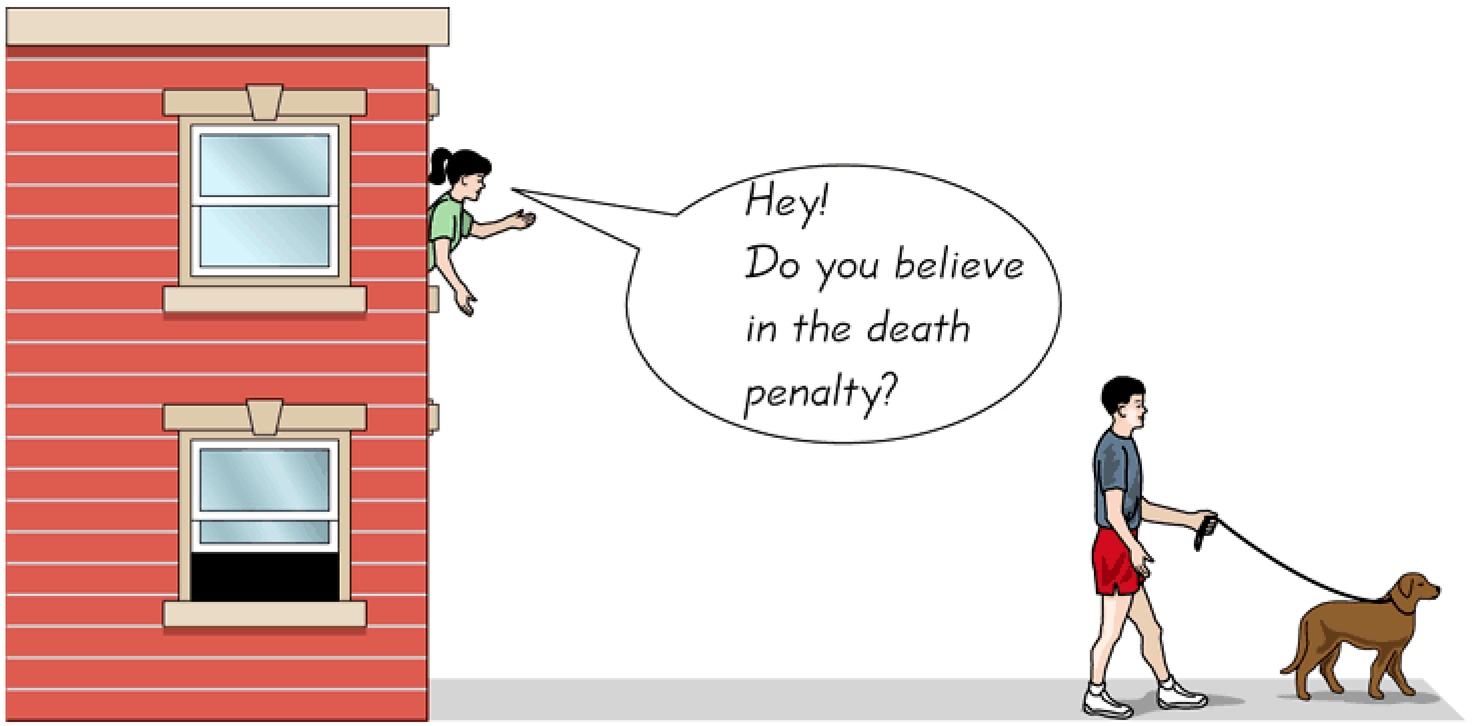


Systematic Sampling

Select some starting point and then select every *k*th element in the population

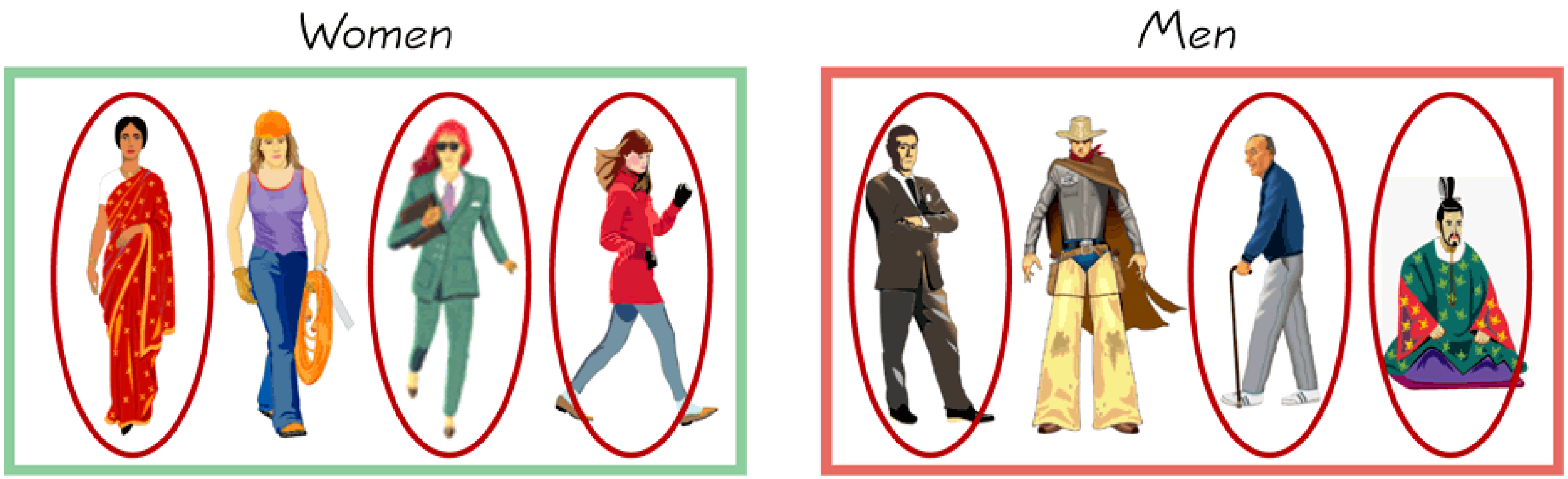
Convenience Sampling

use results that are easy to get



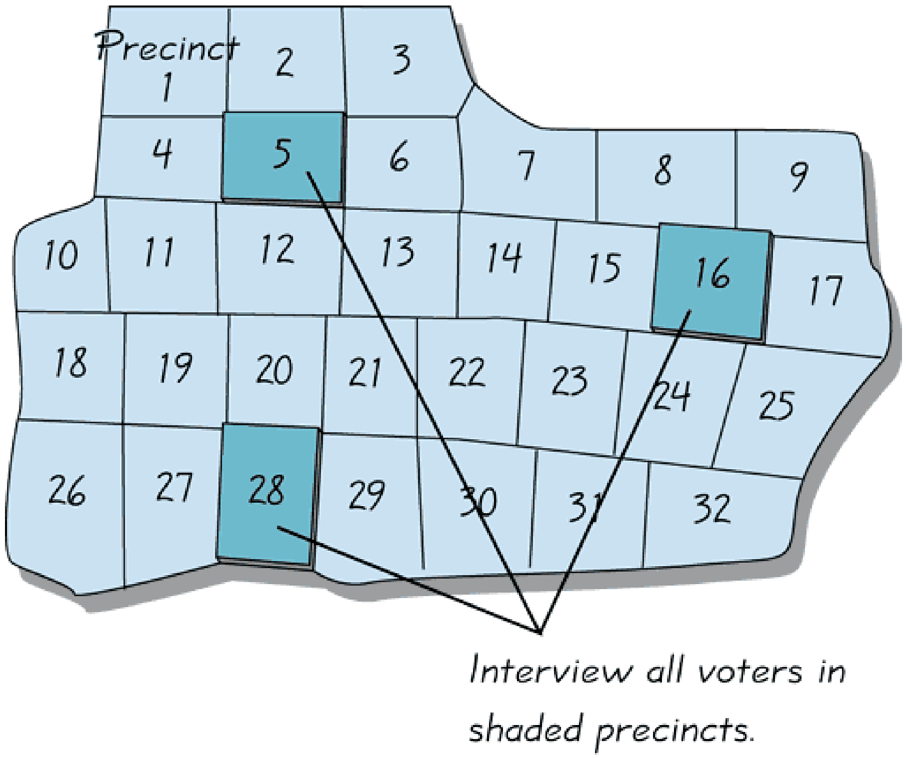
Stratified Sampling

subdivide the population into at least two different subgroups that share the same characteristics, then draw a sample from each subgroup (or stratum)



Cluster Sampling

divide the population area into sections (or clusters); randomly select some of those clusters; choose all members from selected clusters



Multistage Sampling

Collect data by using some combination of the basic sampling methods

In a multistage sample design, pollsters select a sample in different stages, and each stage might use different methods of sampling

Methods of Sampling - Summary

- Random
- Systematic
- Convenience
- Stratified
- Cluster
- Multistage

Beyond the Basics of Collecting Data

Different types of observational studies and experiment design

Types of Studies

- **Cross sectional study**

  data are observed, measured, and collected at one point in time

- **Retrospective (or case control) study**

  data are collected from the past by going back in time (examine records, interviews, …)

- **Prospective (or longitudinal or cohort) study**

  data are collected in the future from groups sharing common factors (called cohorts)

Randomization

- **Randomization**

  is used when subjects are assigned to different groups through a process of random selection. The logic is to use chance as a way to create two groups that are similar.

Replication

- **Replication**

  is the repetition of an experiment on more than one subject. Samples should be large enough so that the erratic behavior that is characteristic of very small samples will not disguise the true effects of different treatments. It is used effectively when there are enough subjects to recognize the differences from different treatments.

Use a sample size that is large enough to let us see the true nature of any effects, and obtain the sample using an appropriate method, such as one based on *randomness*.

Blinding

- **Blinding**

  is a technique in which the subject doesn't know whether he or she is receiving a treatment or a placebo. Blinding allows us to determine whether the treatment effect is significantly different from a placebo effect, which occurs when an untreated subject reports improvement in symptoms.

Double Blind

- **Double-Blind**

Blinding occurs at two levels:

(1) The subject doesn't know whether he or she is receiving the treatment or a placebo

(2) The experimenter does not know whether he or she is administering the treatment or placebo

Confounding

- **Confounding**

  occurs in an experiment when the experimenter is not able to distinguish between the effects of different factors.

  Try to plan the experiment so that confounding does not occur.

Controlling Effects of Variables

- **Completely Randomized Experimental Design**
  assign subjects to different treatment groups through a process of random selection
- **Randomized Block Design**
  a block is a group of subjects that are similar, but blocks differ in ways that might affect the outcome of the experiment
- **Rigorously Controlled Design**
  carefully assign subjects to different treatment groups, so that those given each treatment are similar in ways that are important to the experiment
- **Matched Pairs Design**
  compare exactly two treatment groups using subjects matched in pairs that are somehow related or have similar characteristics

Summary

Three very important considerations in the design of experiments are the following:

1. **Use *randomization* to assign subjects to different groups**

2. **Use replication by repeating the experiment on enough subjects so that effects of treatment or other factors can be clearly seen.**

3. ***Control the effects of variables* by using such techniques as blinding and a completely randomized experimental design**

Errors

No matter how well you plan and execute the sample collection process, there is likely to be some error in the results.

- **Sampling error**

  the difference between a sample result and the true population result; such an error results from chance sample fluctuations

- **Nonsampling error**

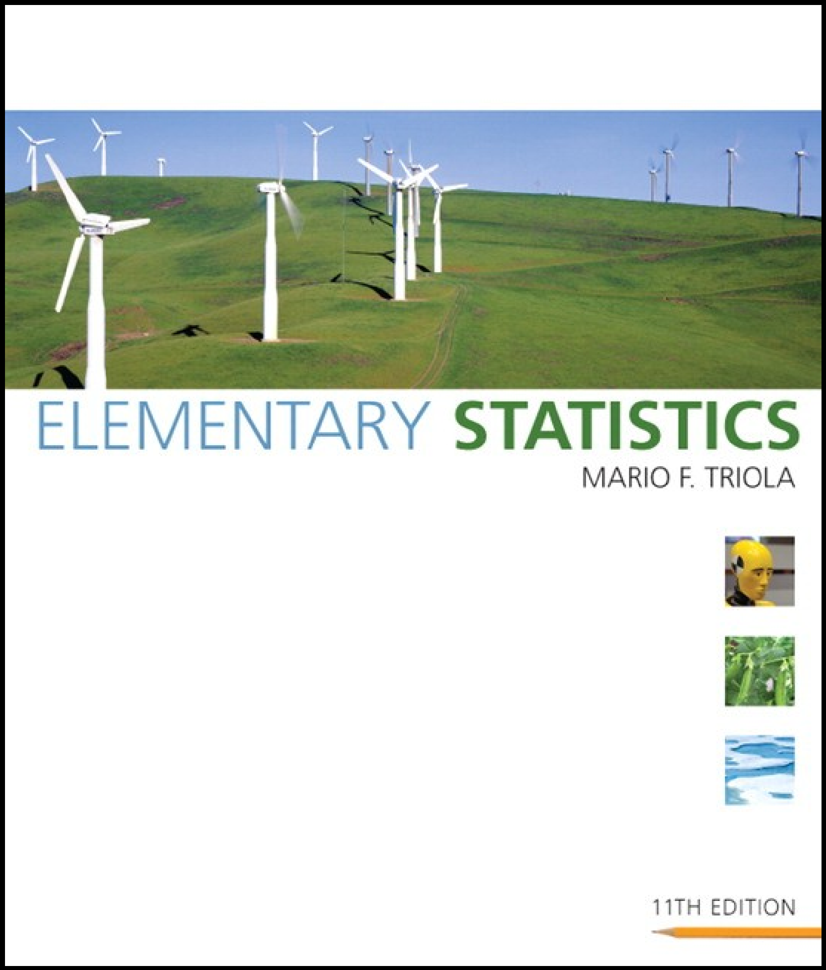
  sample data incorrectly collected, recorded, or analyzed (such as by selecting a biased sample, using a defective instrument, or copying the data incorrectly)

Recap

In this section we have looked at:

- **Types of studies and experiments**
- **Controlling the effects of variables**
- **Randomization**
- **Types of sampling**
- **Sampling errors**

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 2
Summarizing and Graphing Data

2-1 Review and Preview

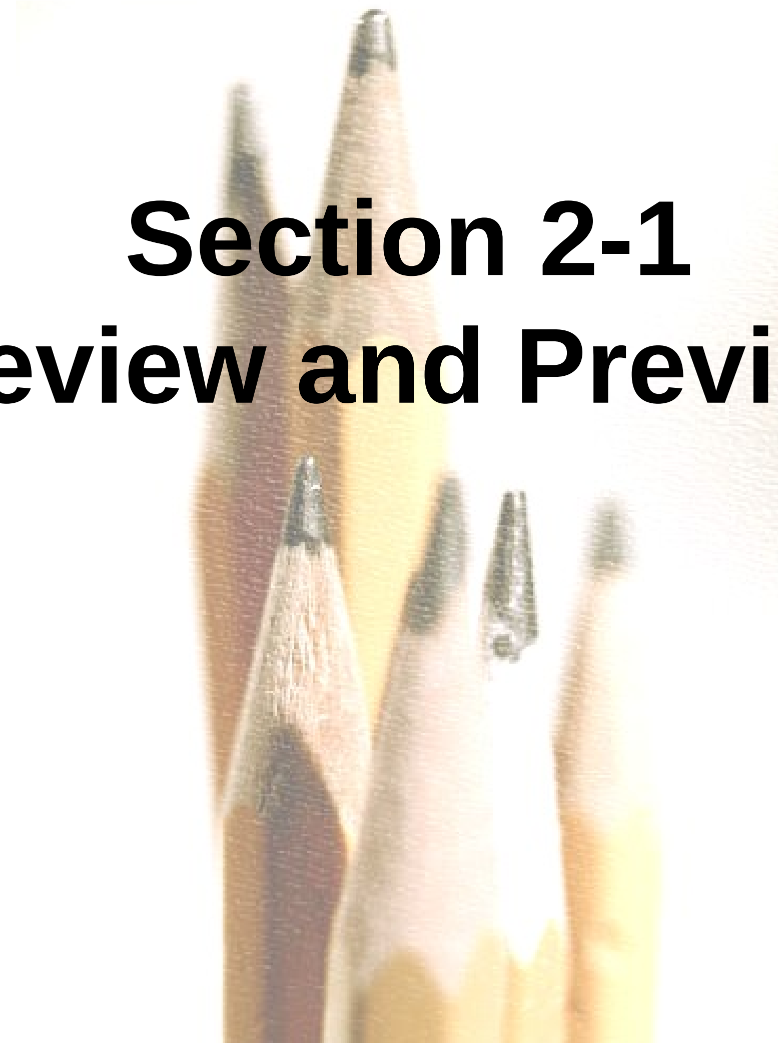
2-2 Frequency Distributions

2-3 Histograms

2-4 Statistical Graphics

2-5 Critical Thinking: Bad Graphs

Section 2-1
Review and Preview



Preview
Important Characteristics of Data

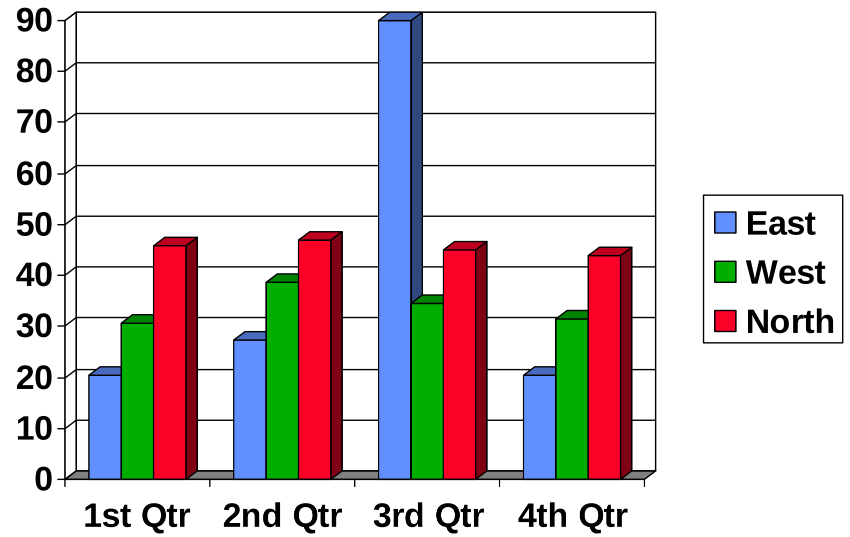
1. **Center**: A representative or average value that indicates where the middle of the data set is located.

2. **Variation**: A measure of the amount that the data values vary.

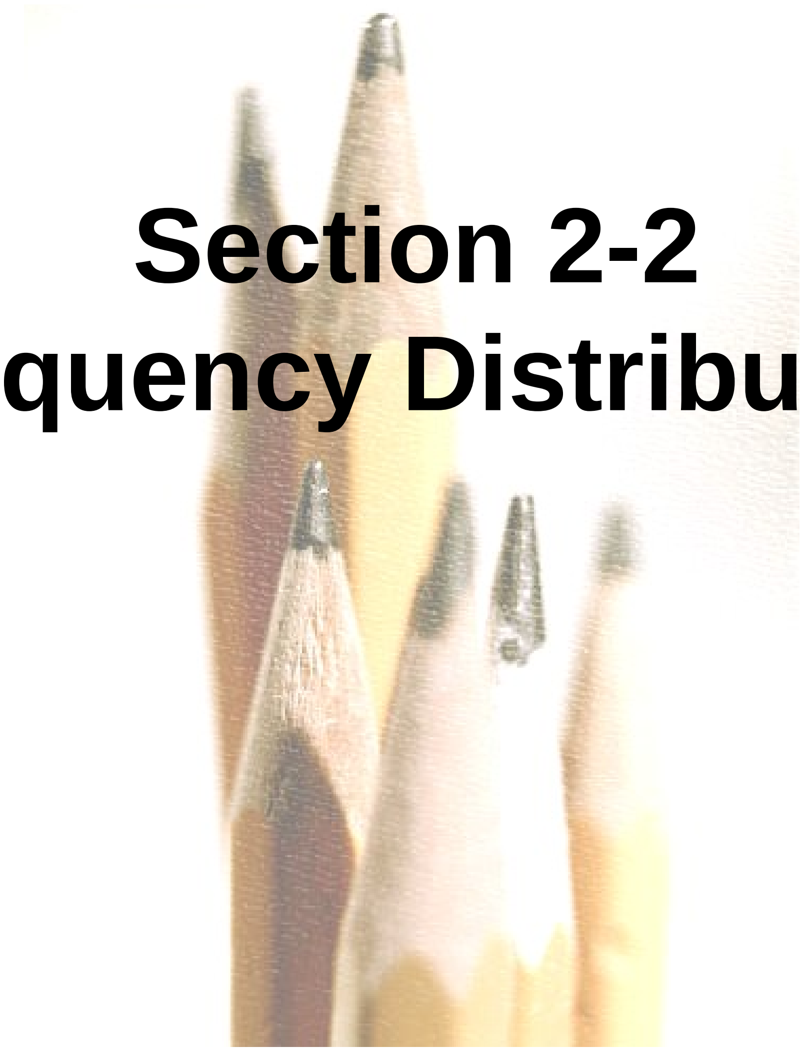
3. **Distribution**: The nature or shape of the spread of data over the range of values (such as bell-shaped, uniform, or skewed).

4. **Outliers**: Sample values that lie very far away from the vast majority of other sample values.

5. **Time**: Changing characteristics of the data over time.



Section 2-2
Frequency Distributions



Key Concept

When working with large data sets, it is often helpful to organize and summarize data by constructing a table called a frequency distribution, defined later. Because computer software and calculators can generate frequency distributions, the details of constructing them are not as important as what they tell us about data sets. It helps us understand the nature of the *distribution* of a data set.

Definition

- **Frequency Distribution (or Frequency Table)**

  shows how a data set is partitioned among all of several categories (or classes) by listing all of the categories along with the number of data values in each of the categories.

Pearson Education

Pulse Rates of Females and Males

Original Data

Table 2-1 Pulse Rates (beats per minute) of Females and Males

Females																			
76	72	88	60	72	68	80	64	68	68	80	76	68	72	96	72	68	72	64	80
64	80	76	76	76	80	104	88	60	76	72	72	88	80	60	72	88	88	124	64
Males																			
68	64	88	72	64	72	60	88	76	60	96	72	56	64	60	64	84	76	84	88
72	56	68	64	60	68	60	60	56	84	72	84	88	56	64	56	56	60	64	72

Frequency Distribution
Pulse Rates of Females

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60–69	12
70–79	14
80–89	11
90–99	1
100–109	1
110–119	0
120–129	1

The *frequency* for a particular class is the number of original values that fall into that class.

Frequency Distributions

Definitions

Lower Class Limits

are the smallest numbers that can actually belong to different classes

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60-69	12
70-79	14
80-89	11
90-99	1
100-109	1
110-119	0
120-129	1

Lower Class Limits

Upper Class Limits

are the largest numbers that can actually belong to different classes

Upper Class Limits

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60-69	12
70-79	14
80-89	11
90-99	1
100-109	1
110-119	0
120-129	1

Class Boundaries

are the numbers used to separate classes, but without the gaps created by class limits

Class Boundaries

59.5
69.5
79.5
89.5
99.5
109.5
119.5
129.5

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60–69	12
70–79	14
80–89	11
90–99	1
100–109	1
110–119	0
120–129	1

Class Midpoints

are the values in the middle of the classes and can be found by adding the lower class limit to the upper class limit and dividing the sum by two

Class Midpoints

Table 2-2 Pulse Rates of Females

	Pulse Rate	Frequency
64.5	60–69	12
74.5	70–79	14
84.5	80–89	11
94.5	90–99	1
104.5	100–109	1
114.5	110–119	0
124.5	120–129	1

Class Width

is the difference between two consecutive lower class limits or two consecutive lower class boundaries

Table 2-2 Pulse Rates of Females

Class Width

	Pulse Rate	Frequency
	60–69	12
10	70–79	14
10	80–89	11
10	90–99	1
10	100–109	1
10	110–119	0
10	120–129	1

Reasons for Constructing Frequency Distributions

1. Large data sets can be summarized.

2. We can analyze the nature of data.

3. We have a basis for constructing important graphs.

Pearson Education

Constructing A Frequency Distribution

1. Determine the number of classes (should be between 5 and 20).

2. Calculate the class width (round up).

$$\text{class width} \approx \frac{(\text{maximum value}) - (\text{minimum value})}{\text{number of classes}}$$

3. Starting point: Choose the minimum data value or a convenient value below it as the first lower class limit.

4. Using the first lower class limit and class width, proceed to list the other lower class limits.

5. List the lower class limits in a vertical column and proceed to enter the upper class limits.

6. Take each individual data value and put a tally mark in the appropriate class. Add the tally marks to get the frequency.

Relative Frequency Distribution

includes the same class limits as a frequency distribution, but the frequency of a class is replaced with a relative frequencies (a proportion) or a percentage frequency (a percent)

$$\text{relative frequency} = \frac{\text{class frequency}}{\text{sum of all frequencies}}$$

$$\text{percentage frequency} = \frac{\text{class frequency}}{\text{sum of all frequencies}} \times 100\%$$

Relative Frequency Distribution

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60-69	12
70-79	14
80-89	11
90-99	1
100-109	1
110-119	0
120-129	1

Total Frequency = 40

Table 2-3 Relative Frequency Distribution of Pulse Rates of Females

Pulse Rate	Relative Frequency
60-69	30% *
70-79	35%
80-89	27.5%
90-99	2.5%
100-109	2.5%
110-119	0
120-129	2.5%

$* \; 12/40 \times 100 = 30\%$

Cumulative Frequency Distribution

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60–69	12
70–79	14
80–89	11
90–99	1
100–109	1
110–119	0
120–129	1

Table 2-4 Cumulative Frequency Distribution of Pulse Rates of Females

Pulse Rate	Cumulative Frequency
Less than 70	12
Less than 80	26
Less than 90	37
Less than 100	38
Less than 110	39
Less than 120	39
Less than 130	40

Cumulative Frequencies

Frequency Tables

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60–69	12
70–79	14
80–89	11
90–99	1
100–109	1
110–119	0
120–129	1

Table 2-3 Relative Frequency Distribution of Pulse Rates of Females

Pulse Rate	Relative Frequency
60–69	30%
70–79	35%
80–89	27.5%
90–99	2.5%
100–109	2.5%
110–119	0
120–129	2.5%

Table 2-4 Cumulative Frequency Distribution of Pulse Rates of Females

Pulse Rate	Cumulative Frequency
Less than 70	12
Less than 80	26
Less than 90	37
Less than 100	38
Less than 110	39
Less than 120	39
Less than 130	40

Critical Thinking Interpreting Frequency Distributions

In later chapters, there will be frequent reference to data with a normal distribution. One key characteristic of a normal distribution is that it has a "bell" shape.

- **The frequencies start low, then increase to one or two high frequencies, then decrease to a low frequency.**
- **The distribution is approximately symmetric, with frequencies preceding the maximum being roughly a mirror image of those that follow the maximum.**

Gaps

- **Gaps**

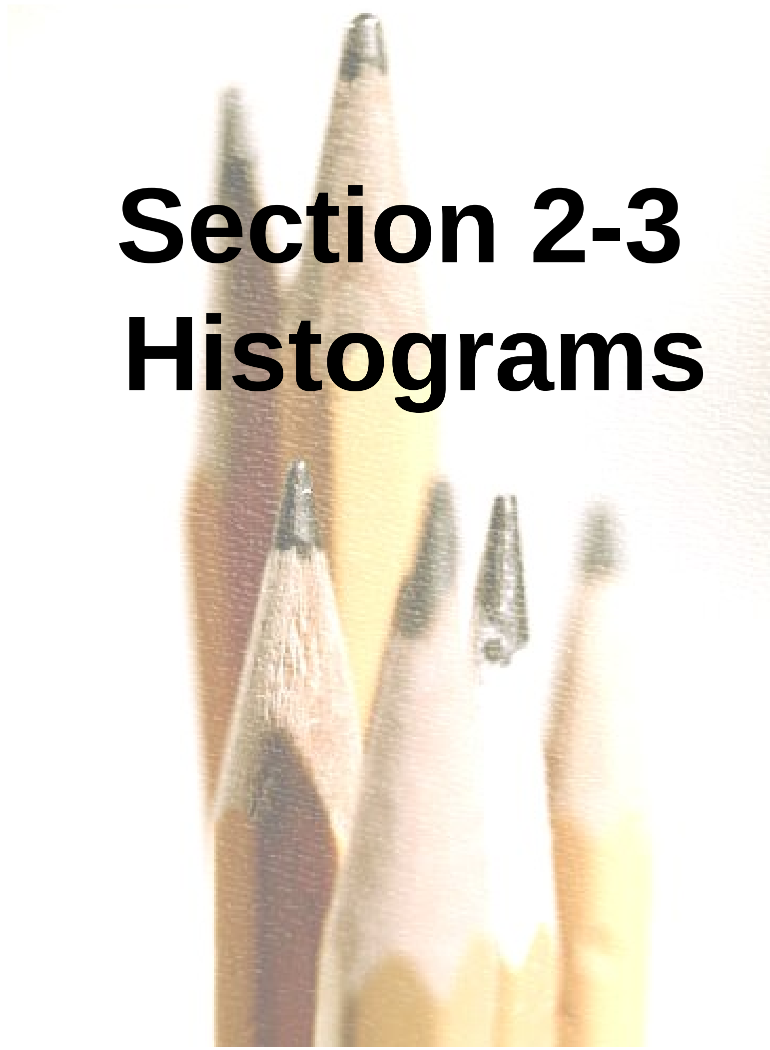
  The presence of gaps can show that we have data from two or more different populations. **However, the converse is not true, because data from different populations do not necessarily result in gaps.**

Pearson Education

Recap

In this Section we have discussed

- Important characteristics of data
- Frequency distributions
- Procedures for constructing frequency distributions
- Relative frequency distributions
- Cumulative frequency distributions



Section 2-3
Histograms

Key Concept

We use a visual tool called a histogram to analyze the shape of the distribution of the data.

Histogram

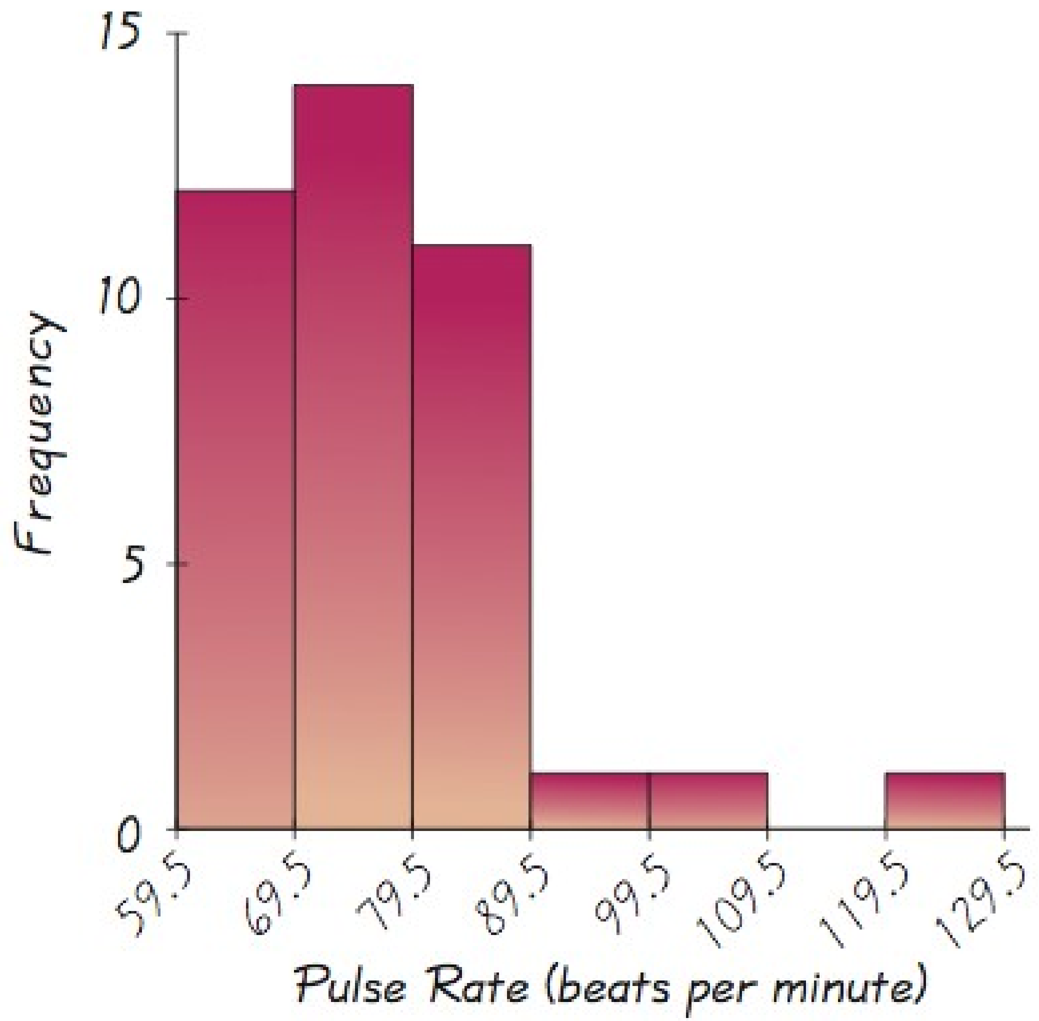
A graph consisting of bars of equal width drawn adjacent to each other (without gaps). The horizontal scale represents the classes of quantitative data values and the vertical scale represents the frequencies. The heights of the bars correspond to the frequency values.

Histogram

Basically a graphic version of a frequency distribution.

Table 2-2 Pulse Rates of Females

Pulse Rate	Frequency
60–69	12
70–79	14
80–89	11
90–99	1
100–109	1
110–119	0
120–129	1



Histogram

The bars on the horizontal scale are labeled with one of the following:

(1) Class boundaries

(2) Class midpoints

(3) Lower class limits (introduces a small error)

Horizontal Scale for Histogram: Use class boundaries or class midpoints.

Vertical Scale for Histogram: Use the class frequencies.

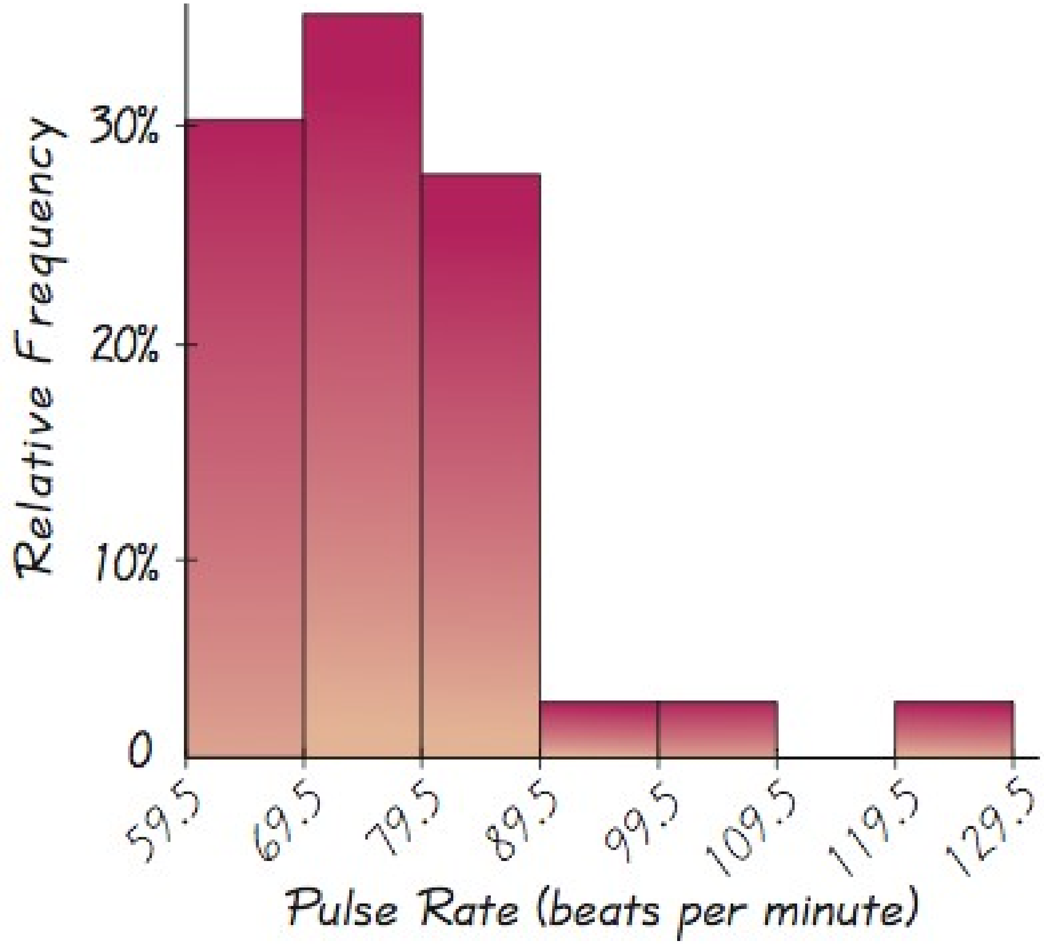
Pearson Education

Relative Frequency Histogram

Has the same shape and horizontal scale as a histogram, but the vertical scale is marked with relative frequencies instead of actual frequencies

Table 2-3 Relative Frequency Distribution of Pulse Rates of Females

Pulse Rate	Relative Frequency
60–69	30%
70–79	35%
80–89	27.5%
90–99	2.5%
100–109	2.5%
110–119	0
120–129	2.5%



Critical Thinking
Interpreting Histograms

Objective is not simply to construct a histogram, but rather to *understand* something about the data.

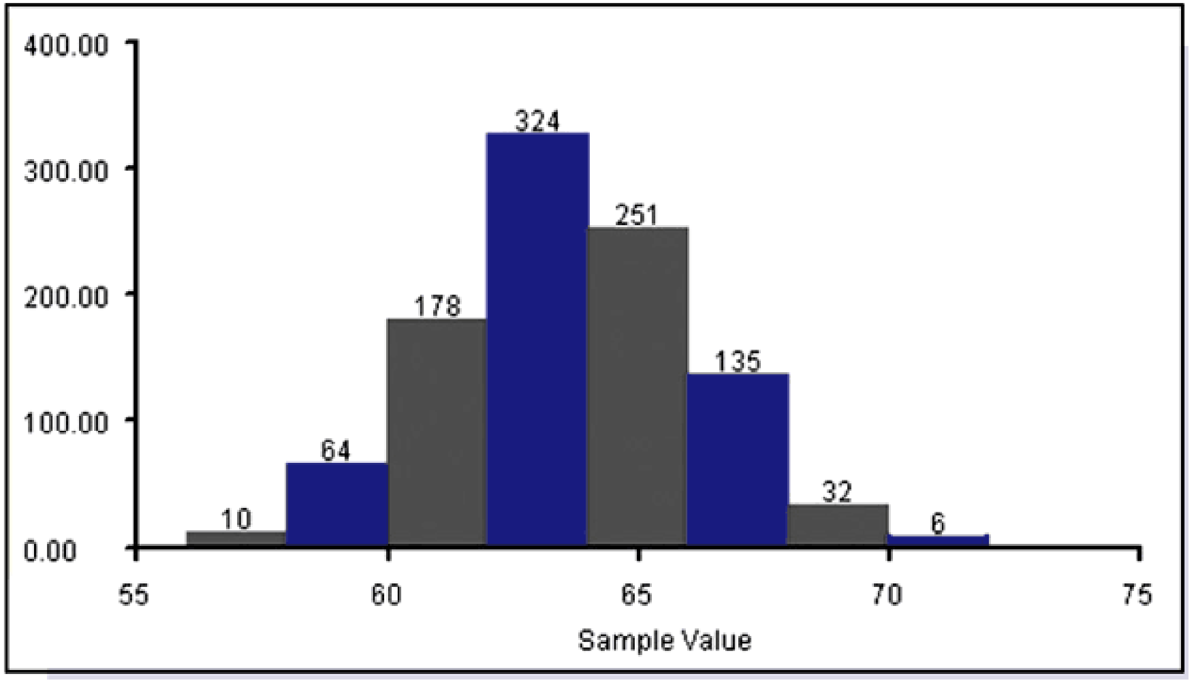
When graphed, a normal distribution has a "bell" shape. Characteristic of the bell shape are

(1) The frequencies increase to a maximum, and then decrease, and

(2) symmetry, with the left half of the graph roughly a mirror image of the right half.

The histogram on the next slide illustrates this.

Critical Thinking
Interpreting Histograms

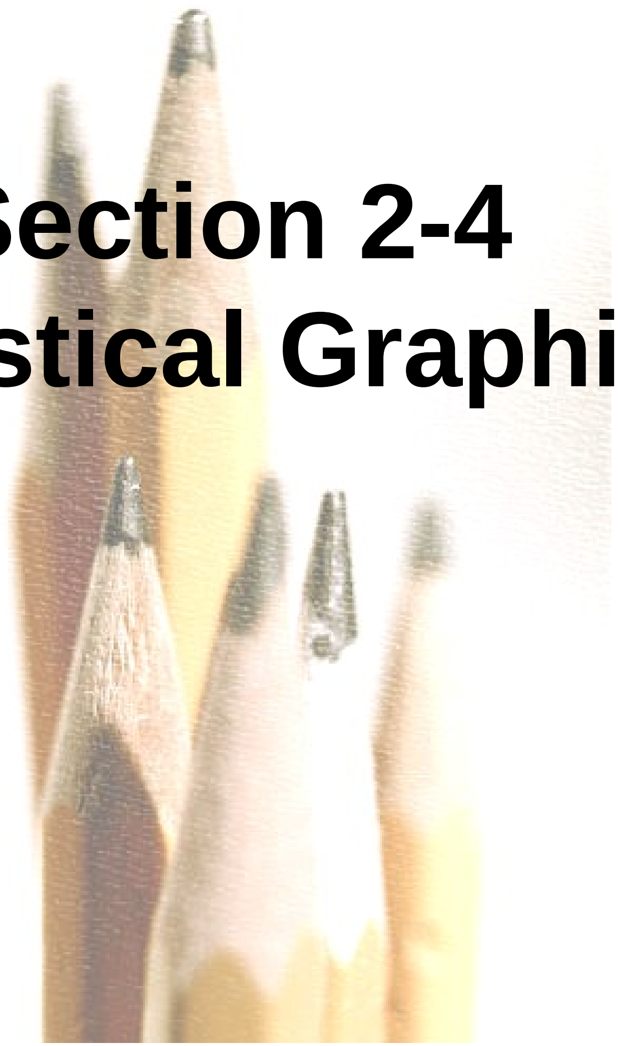


Recap

In this Section we have discussed

- **Histograms**
- **Relative Frequency Histograms**

Section 2-4
Statistical Graphics



Key Concept

This section discusses other types of statistical graphs.

Our objective is to identify a suitable graph for representing the data set. The graph should be effective in revealing the important characteristics of the data.

Frequency Polygon

Uses line segments connected to points directly above class midpoint values

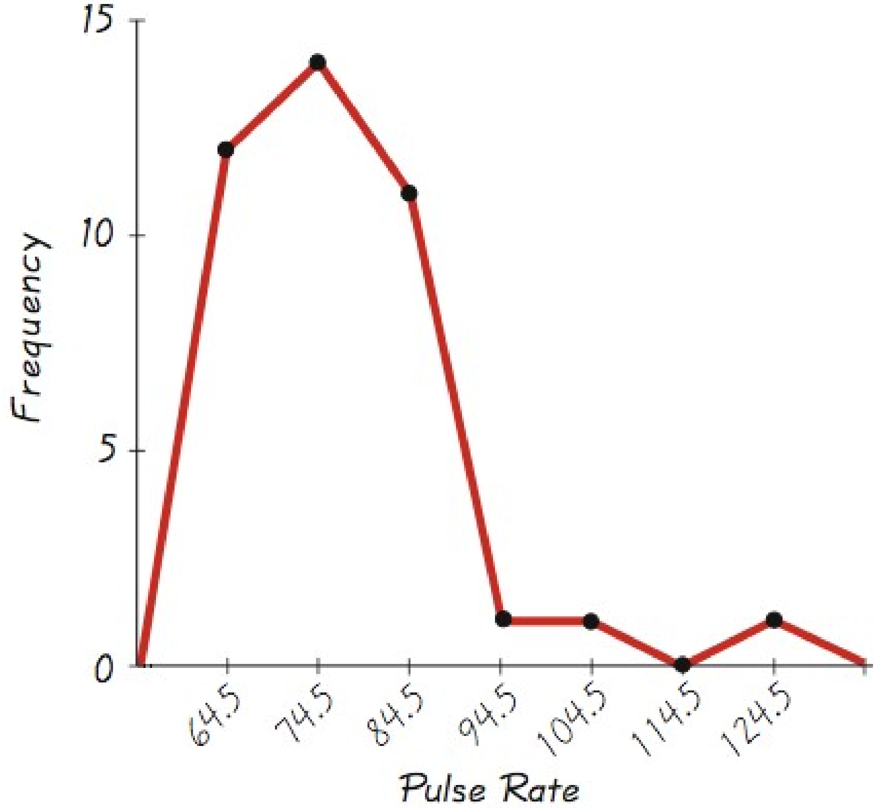


Figure 2-5 Frequency Polygon: Pulse Rates of Women

Relative Frequency Polygon

Uses relative frequencies (proportions or percentages) for the vertical scale.

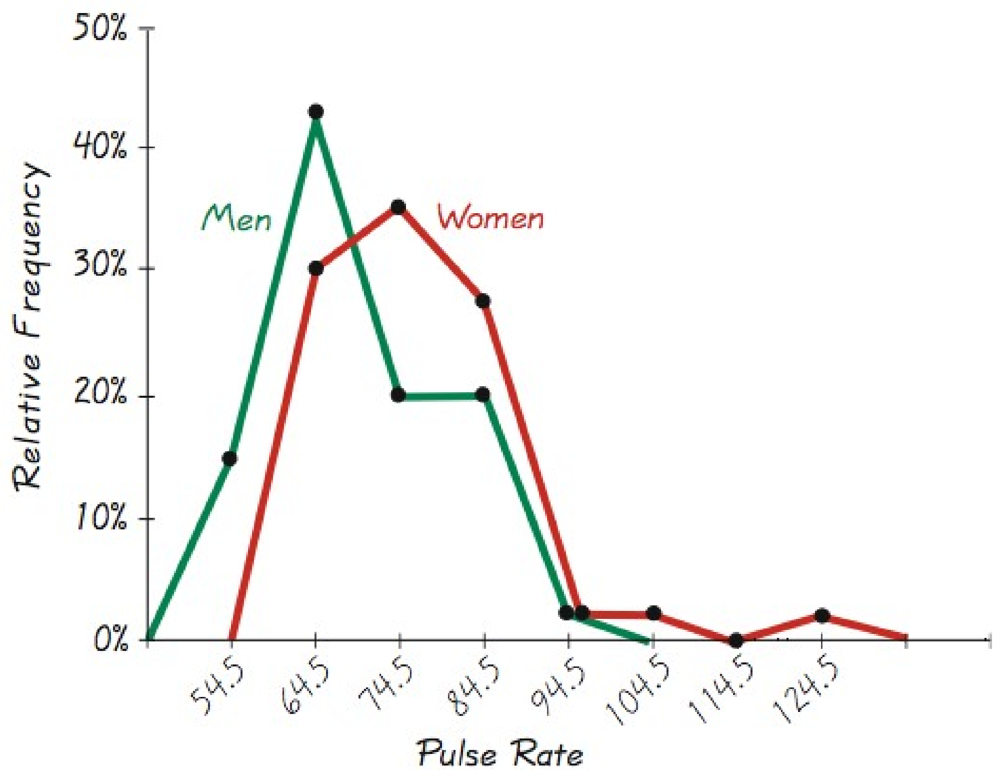


Figure 2-6 Relative Frequency Polygons: Pulse Rates of Women and Men

Ogive

A line graph that depicts cumulative frequencies

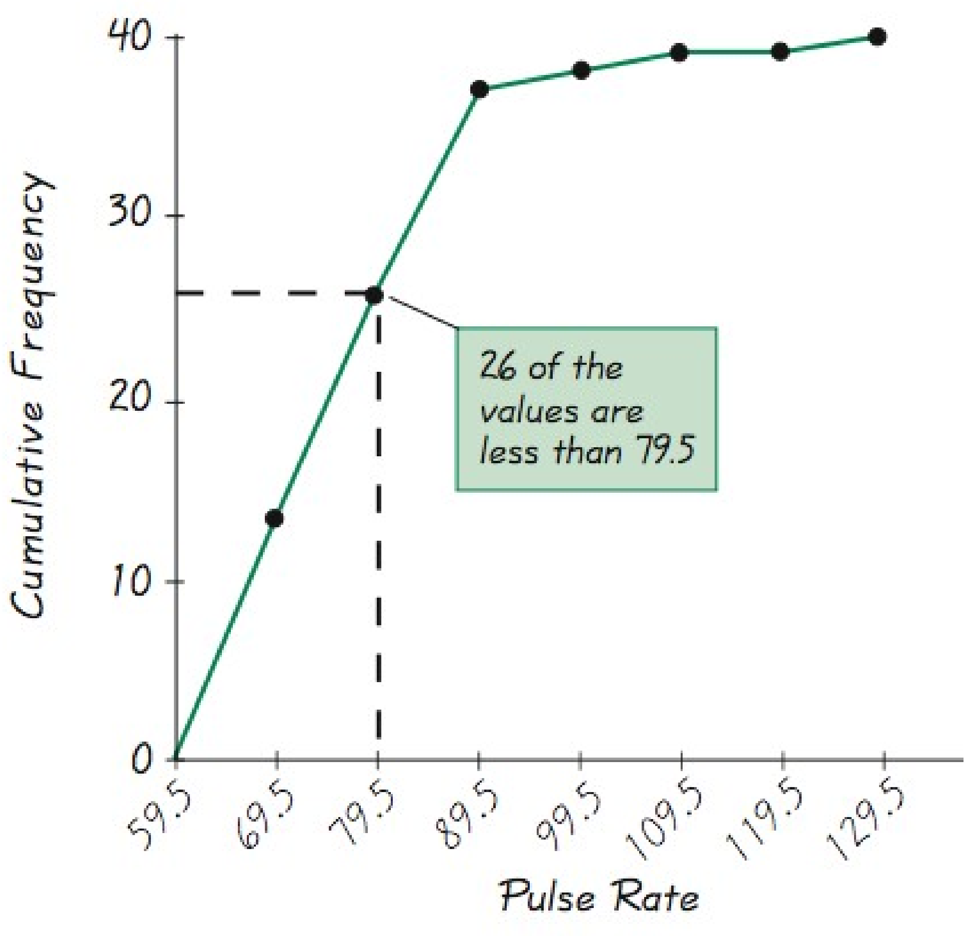
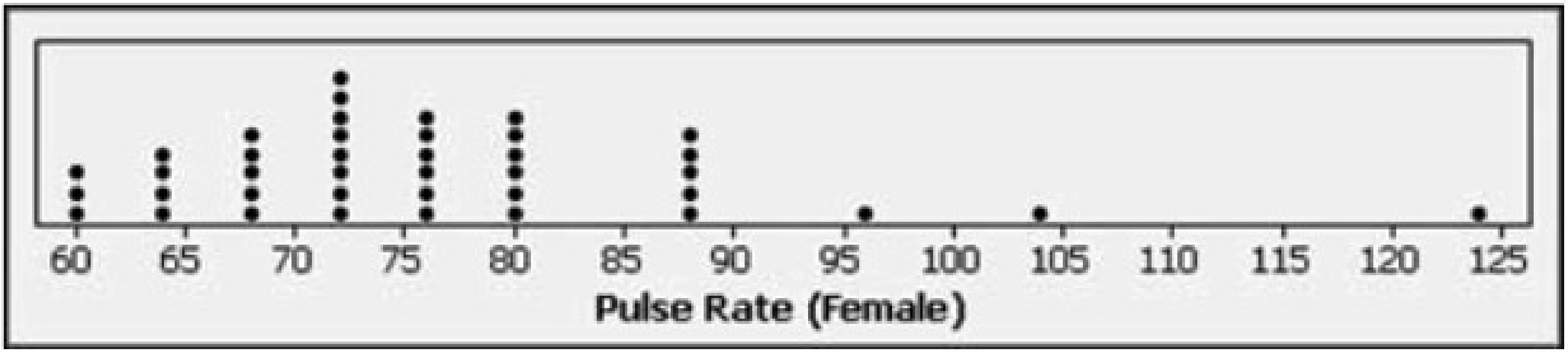


Figure 2-7 Ogive

Dot Plot

Consists of a graph in which each data value is plotted as a point (or dot) along a scale of values. Dots representing equal values are stacked.



Stemplot (or Stem-and-Leaf Plot)

Represents quantitative data by separating each value into two parts: the stem (such as the leftmost digit) and the leaf (such as the rightmost digit)

Stemplot

Stem (tens)	Leaves (units)	
6	000444488888	← Data values are 60, 60, 60, 64, . . . , 68.
7	2222222666666	
8	00000088888	
9	6	← Data value is 96.
10	4	← Data value is 104.
11		
12	4	

Pulse Rates of Females

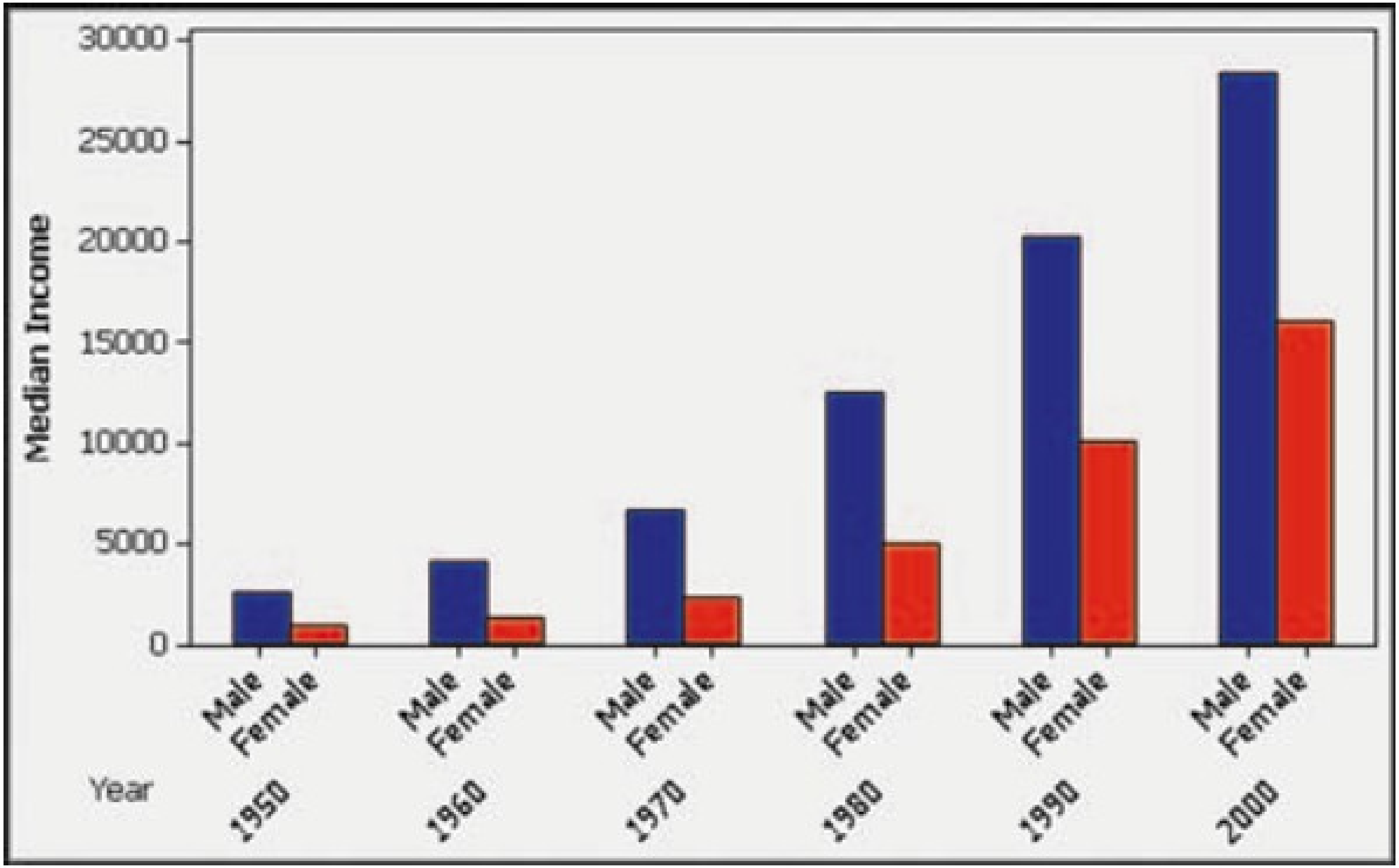
Bar Graph

Uses bars of equal width to show frequencies of categories of qualitative data. Vertical scale represents frequencies or relative frequencies. Horizontal scale identifies the different categories of qualitative data.

A *multiple bar graph* has two or more sets of bars, and is used to compare two or more data sets.

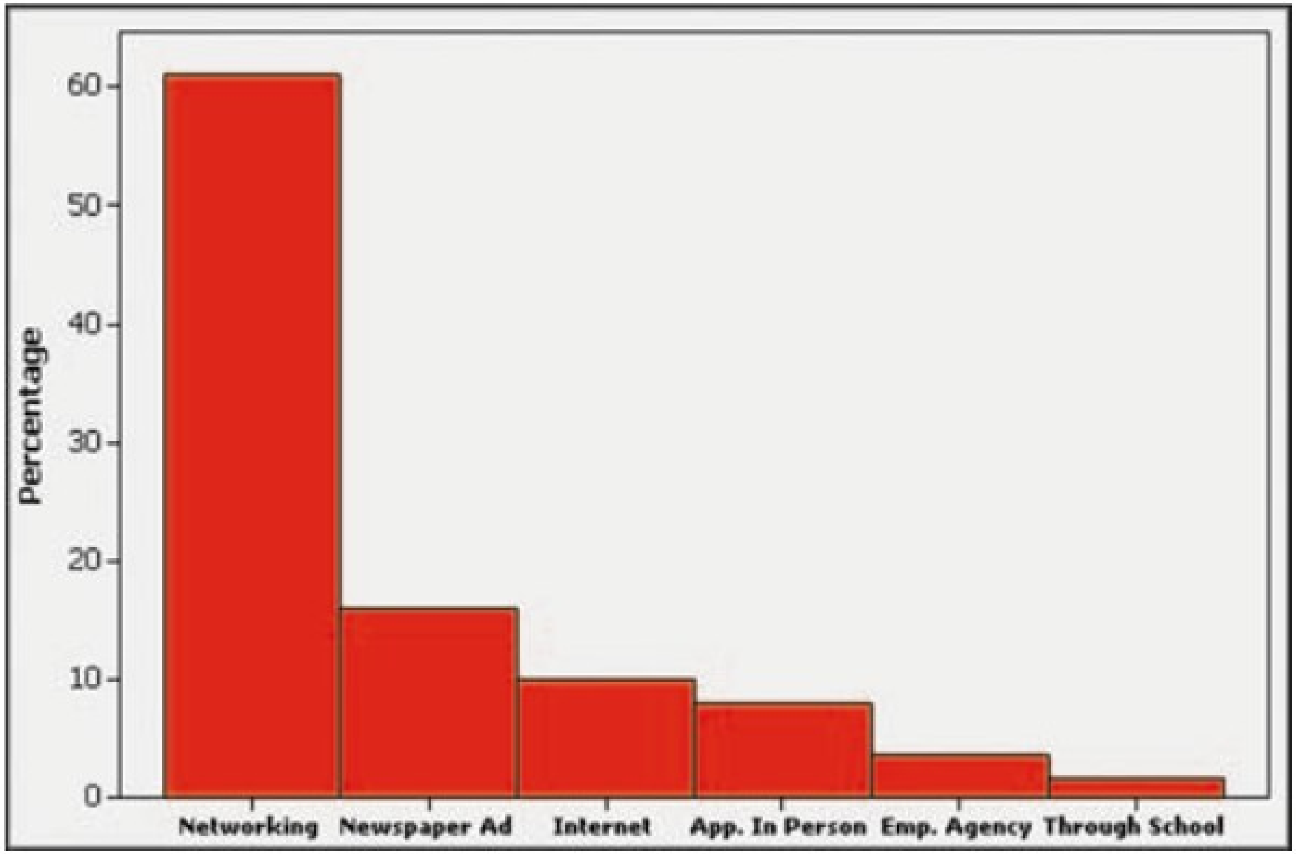
Multiple Bar Graph

Median Income of Males and Females



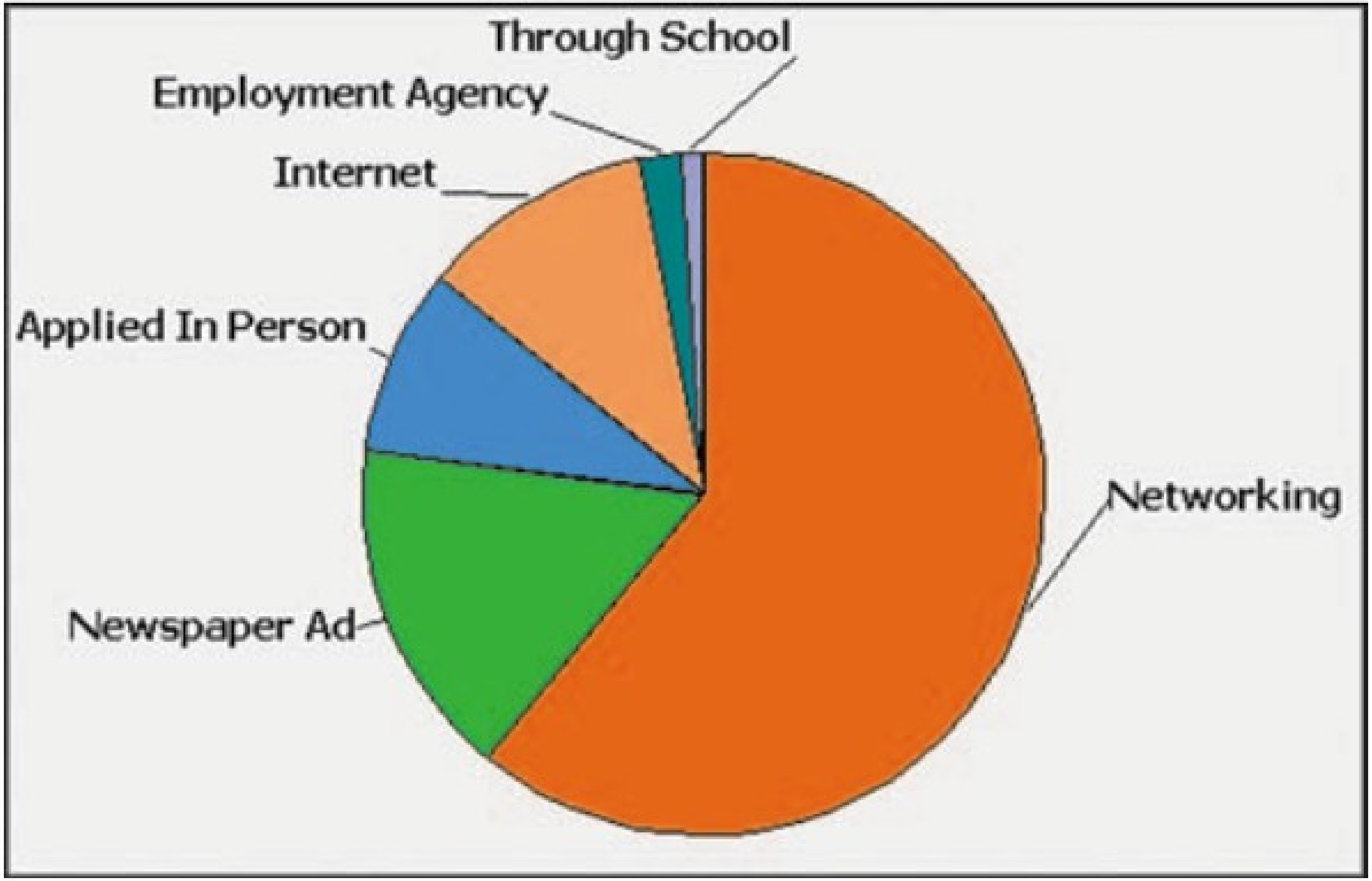
Pareto Chart

A bar graph for qualitative data, with the bars arranged in descending order according to frequencies



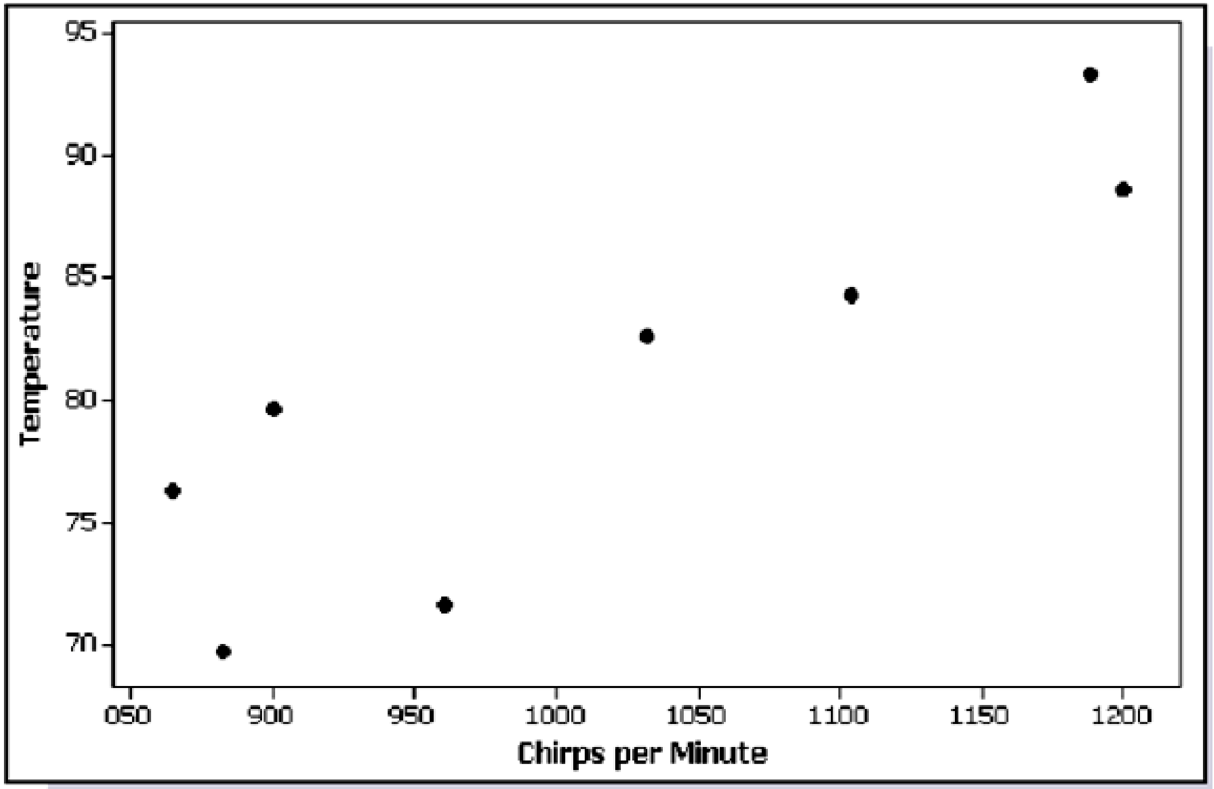
Pie Chart

A graph depicting qualitative data as slices of a circle, size of slice is proportional to frequency count



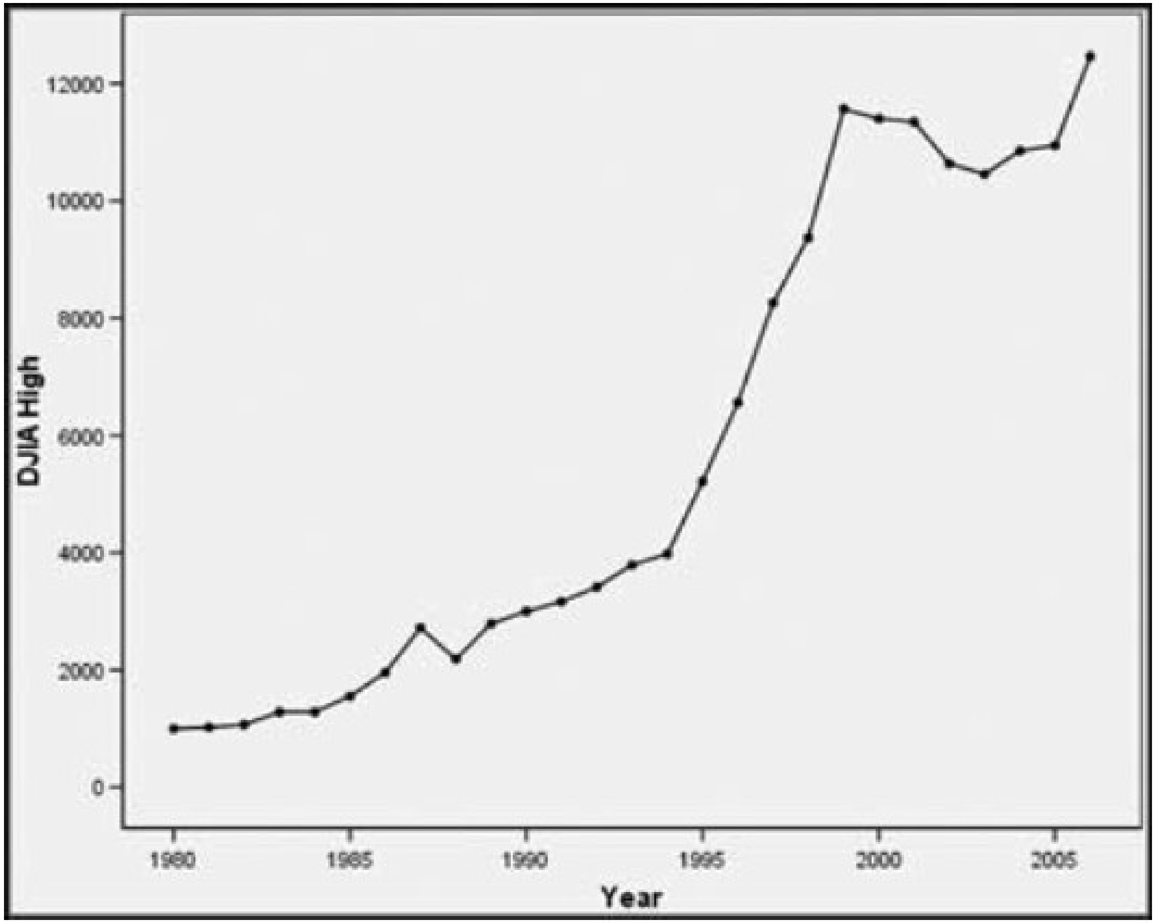
Scatter Plot (or Scatter Diagram)

A plot of paired (*x*,*y*) data with a horizontal *x*-axis and a vertical *y*-axis. Used to determine whether there is a relationship between the two variables



Time-Series Graph

Data that have been collected at different points in time: *time-series data*



Important Principles Suggested by Edward Tufte

For small data sets of 20 values or fewer, use a table instead of a graph.

A graph of data should make the viewer focus on the true nature of the data, not on other elements, such as eye-catching but distracting design features.

Do not distort data, construct a graph to reveal the true nature of the data.

Almost all of the ink in a graph should be used for the data, not the other design elements.

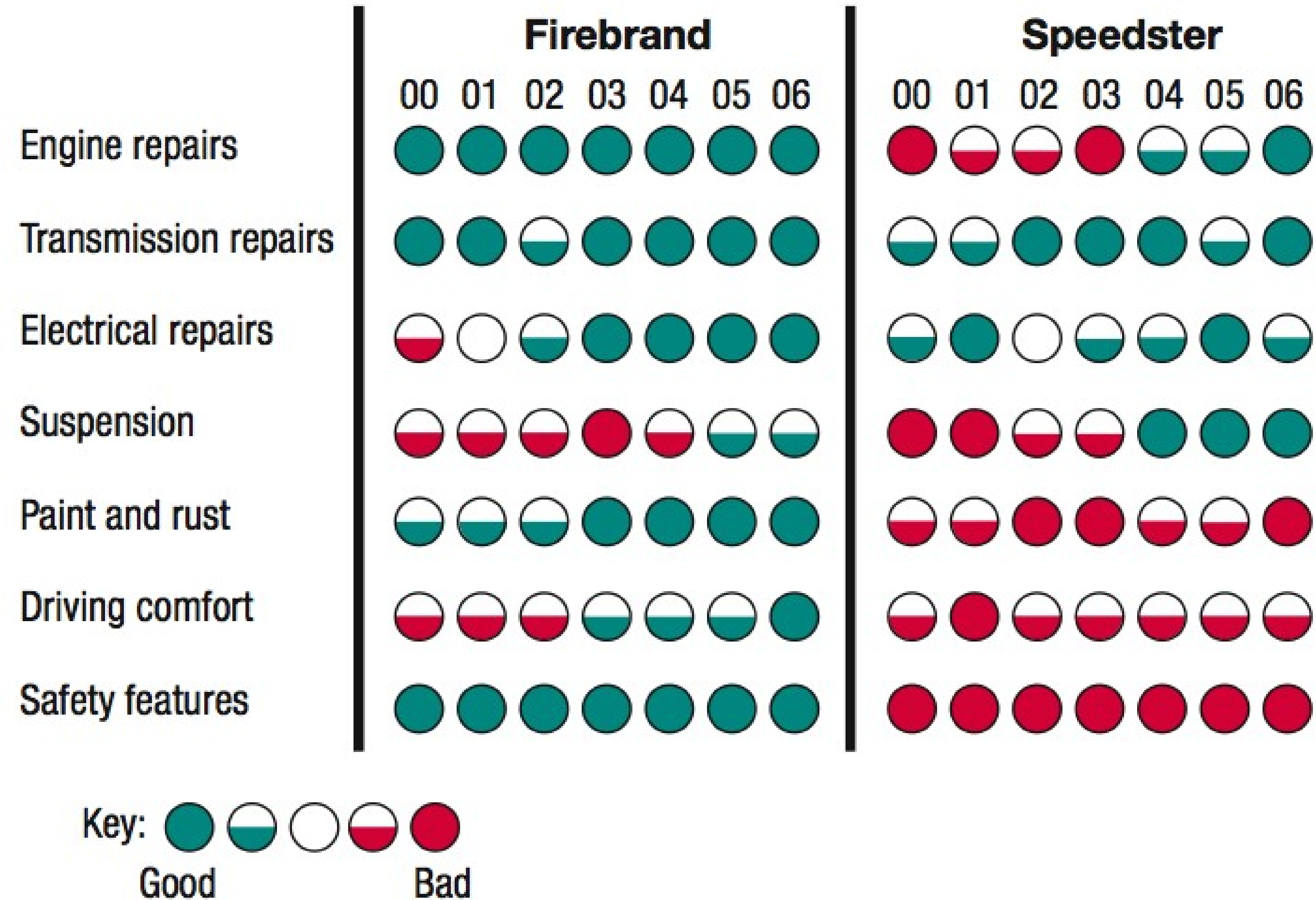
Important Principles Suggested by Edward Tufte

Don’t use screening consisting of features such as slanted lines, dots, cross-hatching, because they create the uncomfortable illusion of movement.

Don’t use area or volumes for data that are actually one-dimensional in nature. (Don’t use drawings of dollar bills to represent budget amounts for different years.)

Never publish pie charts, because they waste ink on nondata components, and they lack appropriate scale.

Car Reliability Data



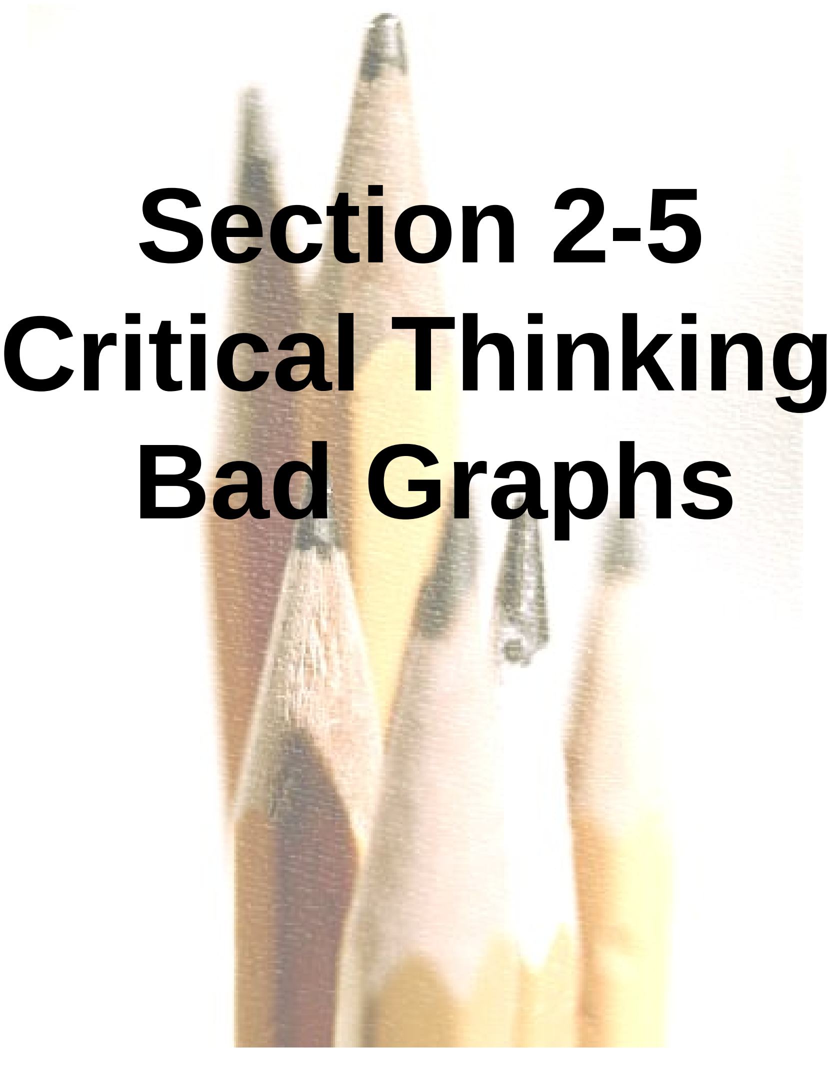
Recap

In this section we saw that graphs are excellent tools for describing, exploring and comparing data.

***Describing data*: Histogram - consider distribution, center, variation, and outliers.**

***Exploring data*: features that reveal some useful and/or interesting characteristic of the data set.**

***Comparing data*: Construct similar graphs to compare data sets.**



Section 2-5
Critical Thinking:
Bad Graphs

Key Concept

Some graphs are bad in the sense that they contain errors.

Some are bad because they are technically correct, but misleading.

It is important to develop the ability to recognize bad graphs and identify exactly how they are misleading.

Nonzero Axis

Are misleading because one or both of the axes begin at some value other than zero, so that differences are exaggerated.

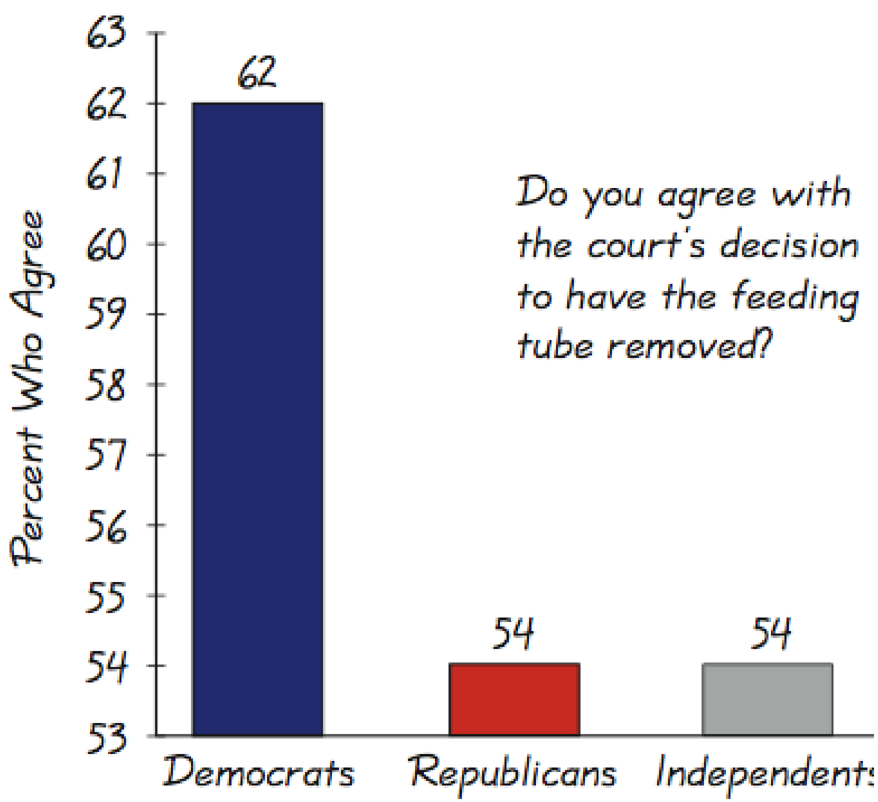


Figure 2-1 Survey Results by Party

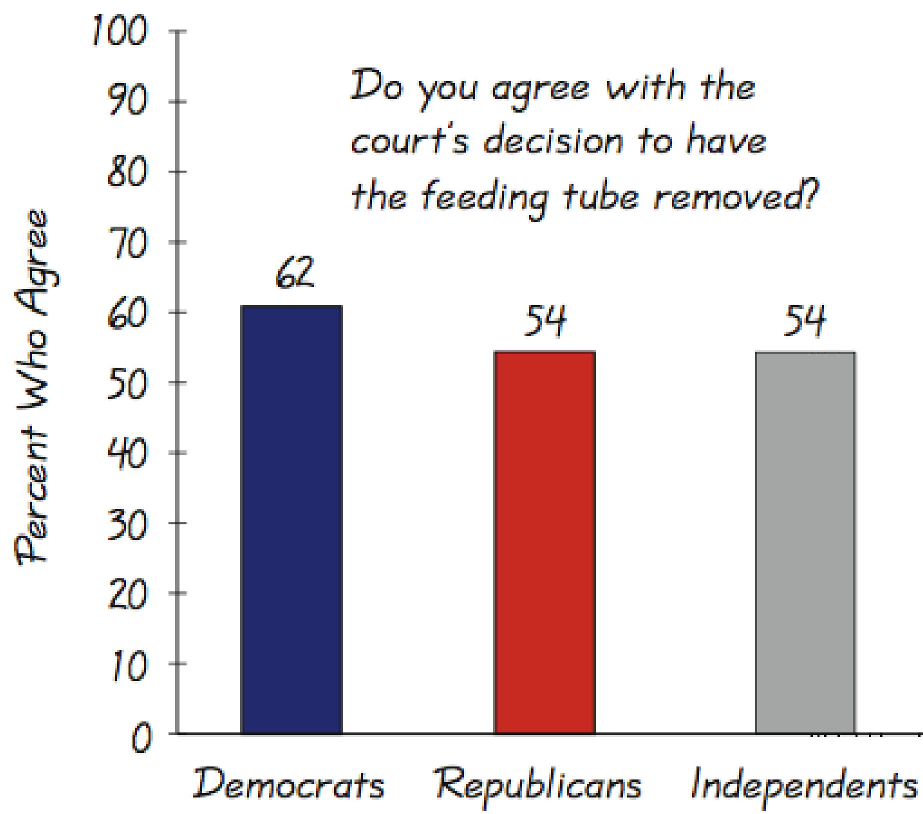


Figure 2-9 Survey Results by Party

Pictographs

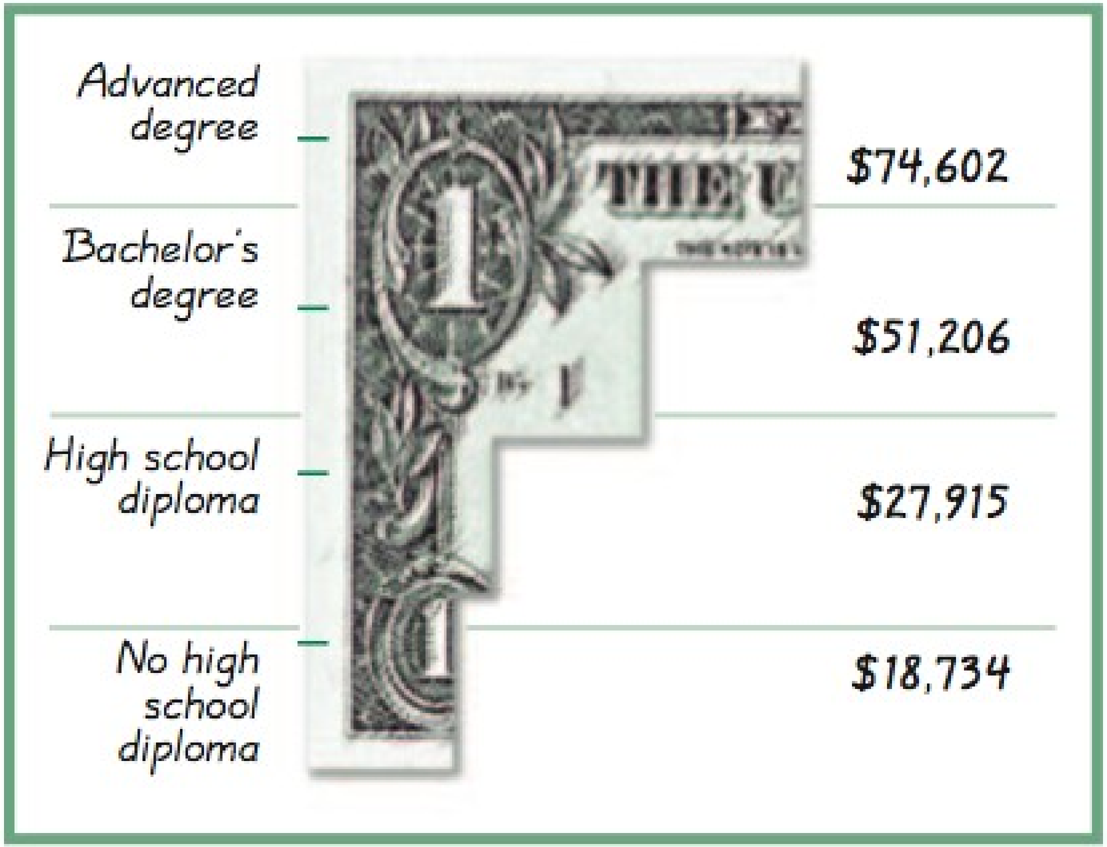
are drawings of objects. Three-dimensional objects - money bags, stacks of coins, army tanks (for army expenditures), people (for population sizes), barrels (for oil production), and houses (for home construction) are commonly used to depict data.

These drawings can create false impressions that distort the data.

If you double each side of a square, the area does not merely double; it increases by a factor of four;if you double each side of a cube, the volume does not merely double; it increases by a factor of eight.

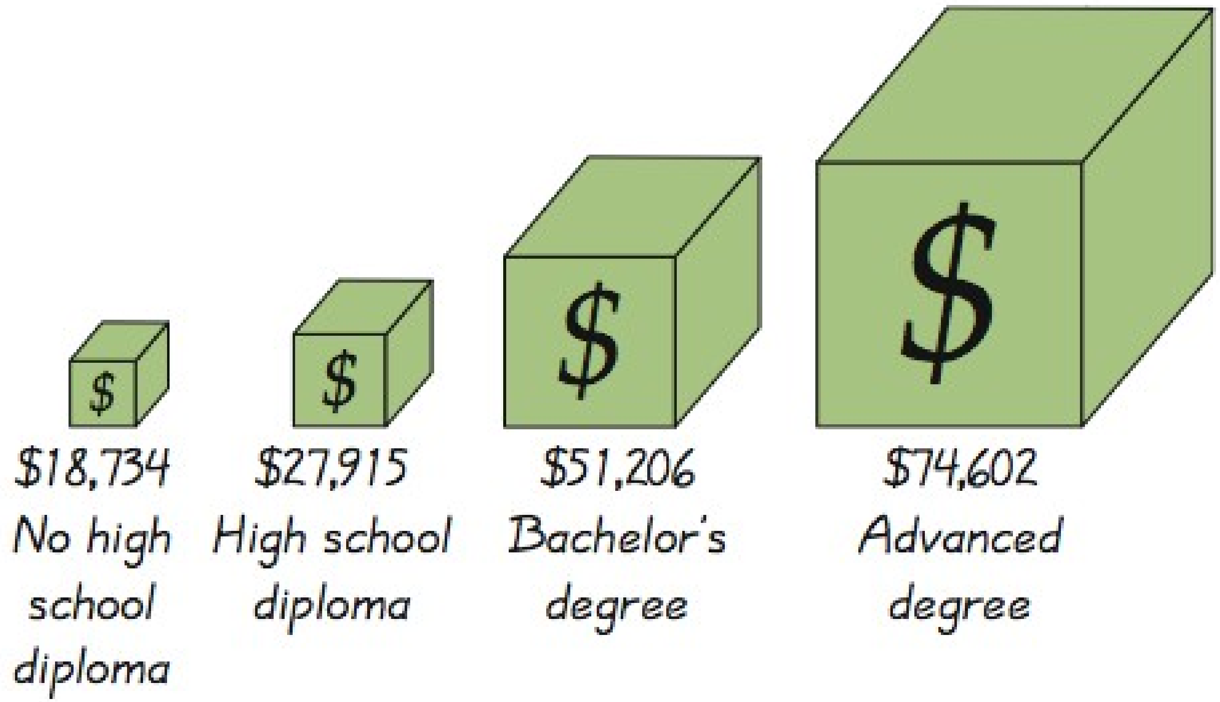
Pictographs using areas and volumes can therefore be very misleading.

Annual Incomes of Groups with Different Education Levels



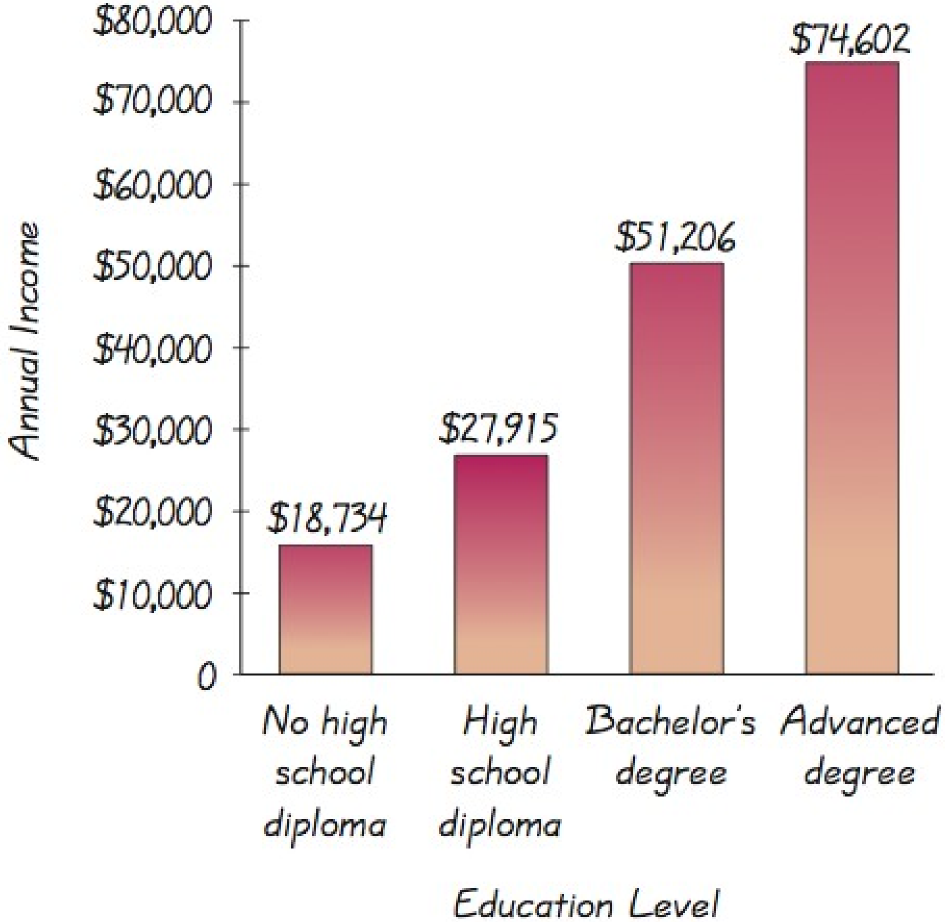
Bars have same width, too busy, too difficult to understand.

Annual Incomes of Groups with Different Education Levels



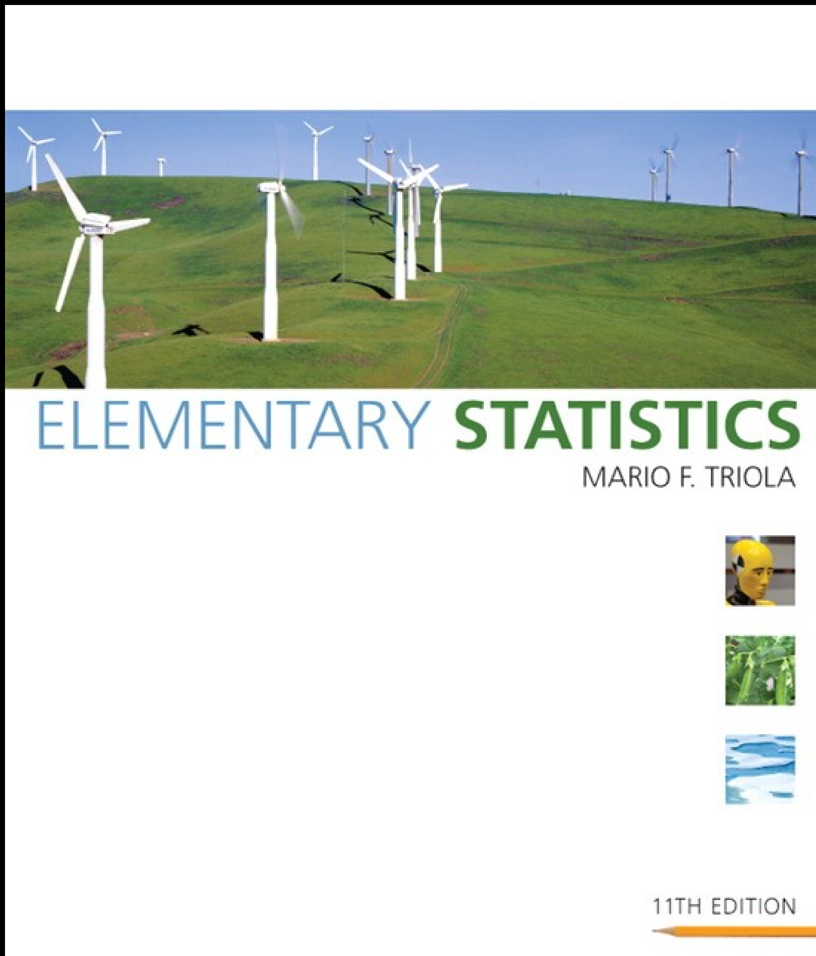
Misleading. Depicts one-dimensional data with three-dimensional boxes. Last box is 64 times as large as first box, but income is only 4 times as large.

Annual Incomes of Groups with Different Education Levels



Fair, objective, unencumbered by distracting features.

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 3
Statistics for Describing, Exploring, and Comparing Data

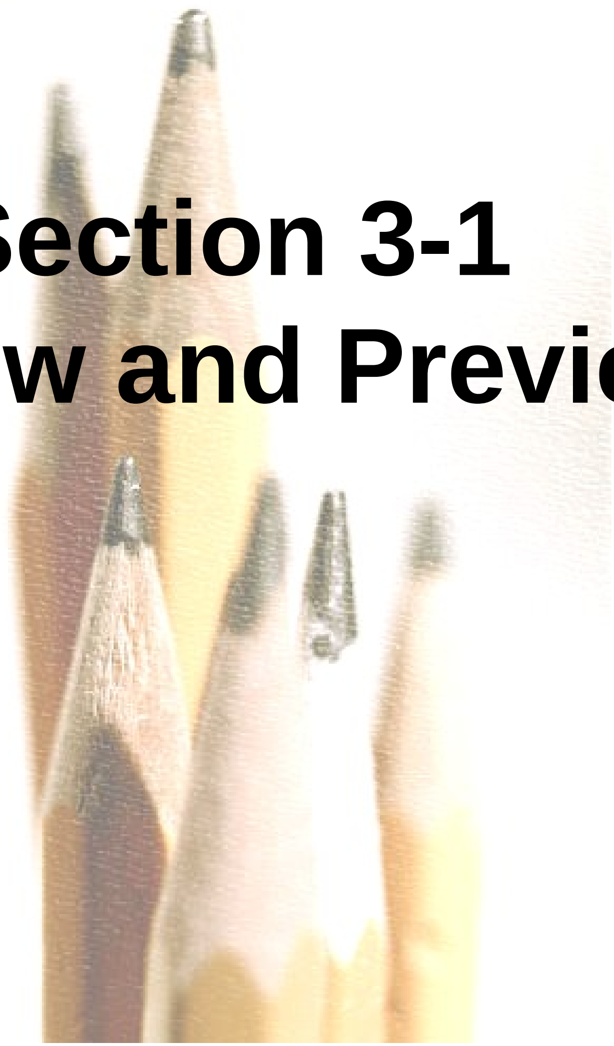
3-1 Review and Preview

3-2 Measures of Center

3-3 Measures of Variation

3-4 Measures of Relative Standing and Boxplots

Section 3-1
Review and Preview



Created by Tom Wegleitner, Centreville, Virginia

Review

- **Chapter 1**
  Distinguish between population and sample, parameter and statistic
  Good sampling methods: *simple random sample*, collect in appropriate ways

- **Chapter 2**
  Frequency distribution: summarizing data
  Graphs designed to help understand data
  Center, variation, distribution, outliers, changing characteristics over time

Preview

- **Important Statistics**

  Mean, median, standard deviation, variance

- **Understanding and Interpreting**

  these important statistics

Preview

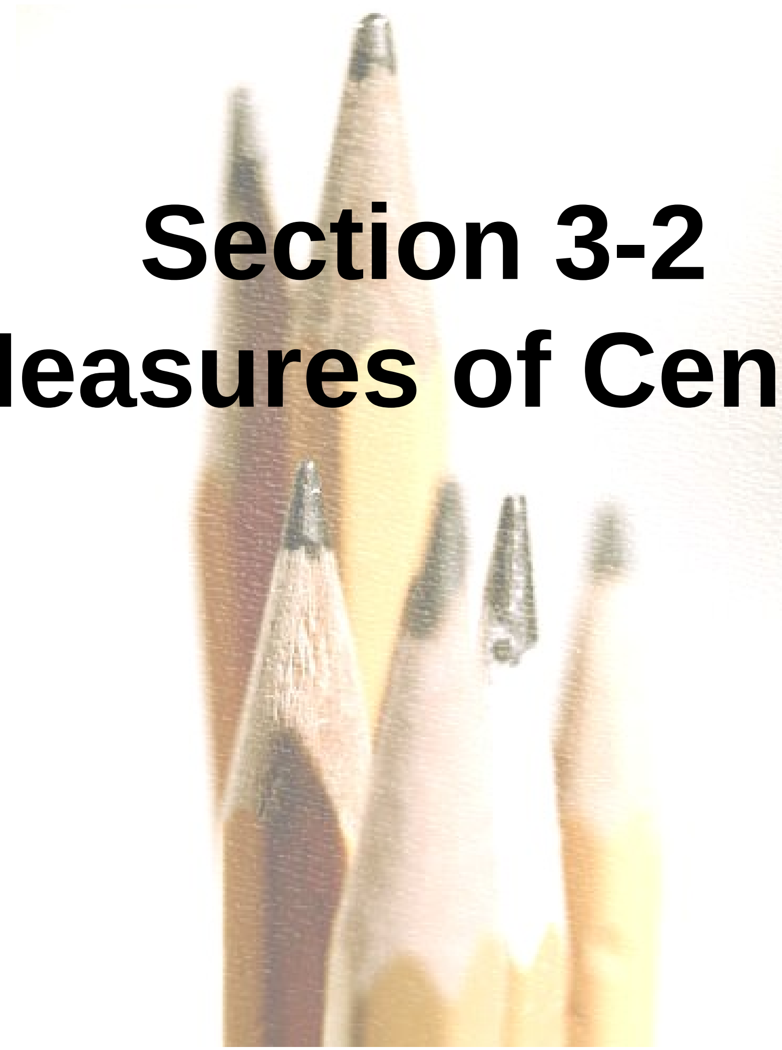
- **Descriptive Statistics**

  In this chapter we'll learn to summarize or describe the important characteristics of a known set of data

- **Inferential Statistics**

  In later chapters we'll learn to use sample data to make inferences or generalizations about a population

Section 3-2
Measures of Center



Key Concept

Characteristics of center. Measures of center, including mean and median, as tools for analyzing data. Not only determine the value of each measure of center, but also interpret those values.

Part 1

Basics Concepts of Measures of Center

Measure of Center

- **Measure of Center**

  the value at the center or middle of a data set

Arithmetic Mean

- **Arithmetic Mean (Mean)**

  the measure of center obtained by adding the values and dividing the total by the number of values

What most people call an *average*.

Notation

Σ denotes the **sum** of a set of values.

x is the **variable** usually used to represent the individual data values.

n represents the **number of data values** in a **sample**.

N represents the **number of data values** in a **population**.

Notation

$\bar{x}$ is pronounced 'x-bar' and denotes the mean of a set of sample values

$$\bar{x} = \frac{\Sigma x}{n}$$

μ is pronounced 'mu' and denotes the mean of all values in a population

$$\mu = \frac{\Sigma x}{N}$$

Mean

- **Advantages**

  Is relatively reliable, means of samples drawn from the same population don't vary as much as other measures of center

  Takes every data value into account

- **Disadvantage**

  Is sensitive to every data value, one extreme value can affect it dramatically; is not a *resistant* measure of center

Median

- **Median**

  the middle value when the original data values are arranged in order of increasing (or decreasing) magnitude

- **often denoted by $\tilde{x}$ (pronounced ‘x-tilde’)**

- **is not affected by an extreme value - is a resistant measure of the center**

Finding the Median

First *sort* the values (arrange them in order), the follow one of these

1. **If the number of data values is odd, the median is the number located in the exact middle of the list.**

2. **If the number of data values is even, the median is found by computing the mean of the two middle numbers.**

5.40	1.10	0.42	0.73	0.48	1.10
0.42	0.48	0.73	1.10	1.10	5.40

(in order - even number of values – no exact middle shared by two numbers)

$$\frac{0.73 + 1.10}{2}$$

MEDIAN is 0.915

5.40	1.10	0.42	0.73	0.48	1.10	0.66
0.42	0.48	0.66	0.73	1.10	1.10	5.40

(in order - odd number of values)

exact middle **MEDIAN is 0.73**

Mode

- **Mode**
  the value that occurs with the greatest frequency

- **Data set can have one, more than one, or no mode**

Bimodal	**two data values occur with the same greatest frequency**
Multimodal	**more than two data values occur with the same greatest frequency**
No Mode	**no data value is repeated**

Mode is the only measure of central tendency that can be used with nominal data

Mode - Examples

a. 5.40 1.10 0.42 0.73 0.48 1.10	⇦ Mode is 1.10
b. 27 27 27 55 55 55 88 88 99	⇦ Bimodal - 27 & 55
c. 1 2 3 6 7 8 9 10	⇦ No Mode

Definition

- **Midrange**
  the value midway between the maximum and minimum values in the original data set

$$\text{Midrange} = \frac{\text{maximum value} + \text{minimum value}}{2}$$

Midrange

- **Sensitive to extremes**
  because it uses only the maximum and minimum values, so rarely used

- **Redeeming Features**

  (1) very easy to compute

  (2) reinforces that there are several ways to define the center

  (3) Avoids confusion with median

Round-off Rule for Measures of Center

Carry one more decimal place than is present in the original set of values.

Critical Thinking

Think about whether the results are reasonable.

Think about the method used to collect the sample data.

Part 2

Beyond the Basics of Measures of Center

Mean from a Frequency Distribution

Assume that all sample values in each class are equal to the class midpoint.

Mean from a Frequency Distribution

use class midpoint of classes for variable *x*

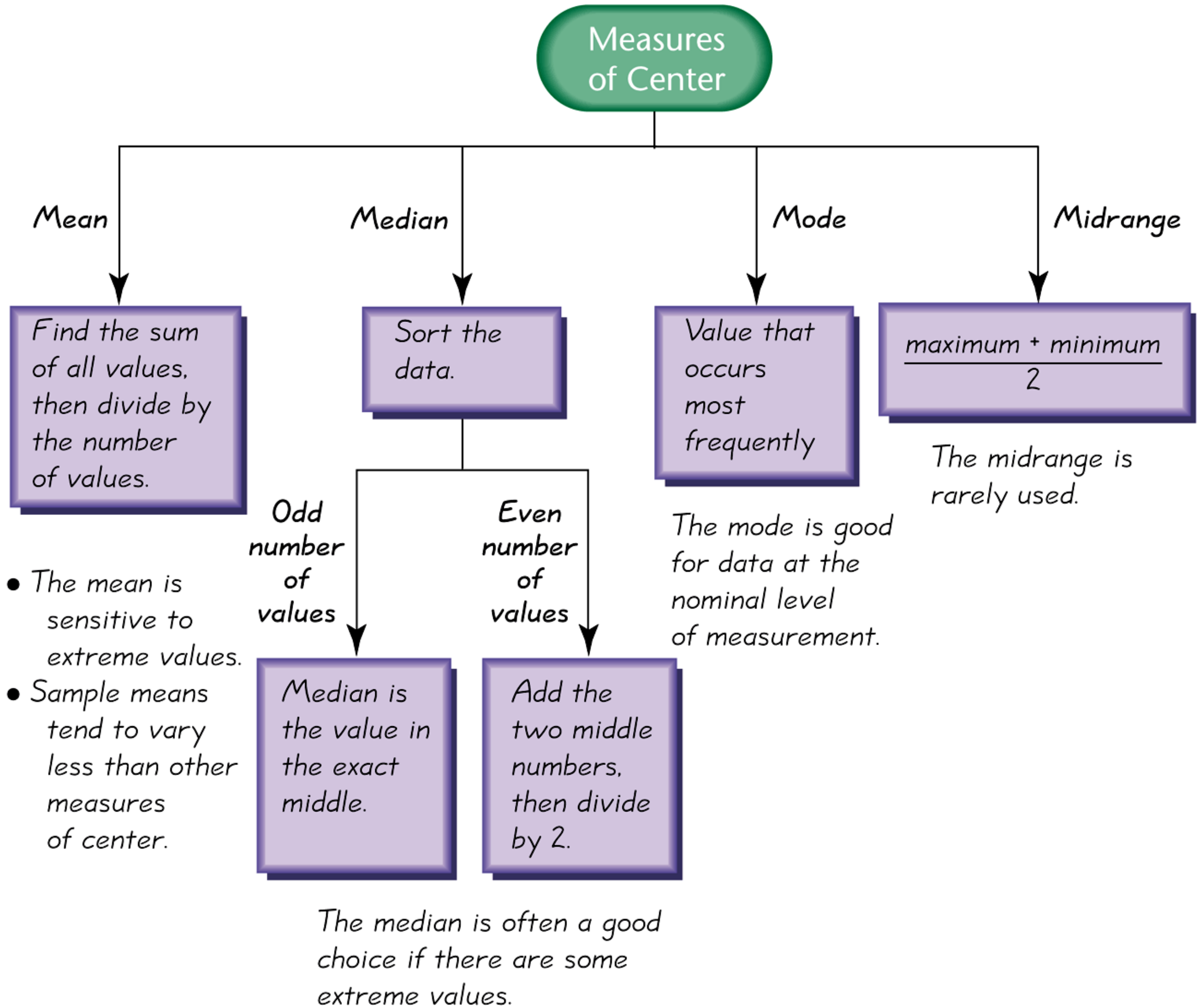
$$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f}$$

Weighted Mean

When data values are assigned different weights, we can compute a **weighted mean**.

$$\bar{x} = \frac{\Sigma (w \cdot x)}{\Sigma w}$$

Best Measure of Center



Skewed and Symmetric

- **Symmetric**

  distribution of data is symmetric if the left half of its histogram is roughly a mirror image of its right half

- **Skewed**

  distribution of data is skewed if it is not symmetric and extends more to one side than the other

Skewed Left or Right

- **Skewed to the left**

  (also called negatively skewed) have a longer left tail, mean and median are to the left of the mode

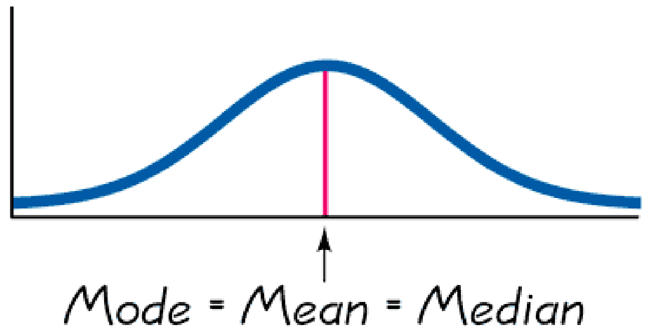
- **Skewed to the right**

  (also called positively skewed) have a longer right tail, mean and median are to the right of the mode

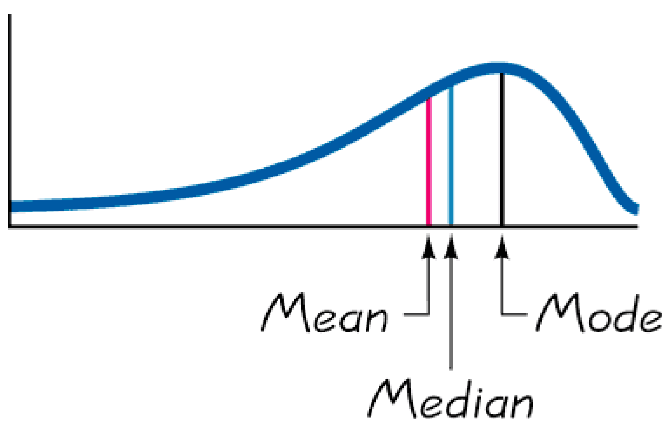
Shape of the Distribution

The mean and median cannot always be used to identify the shape of the distribution.

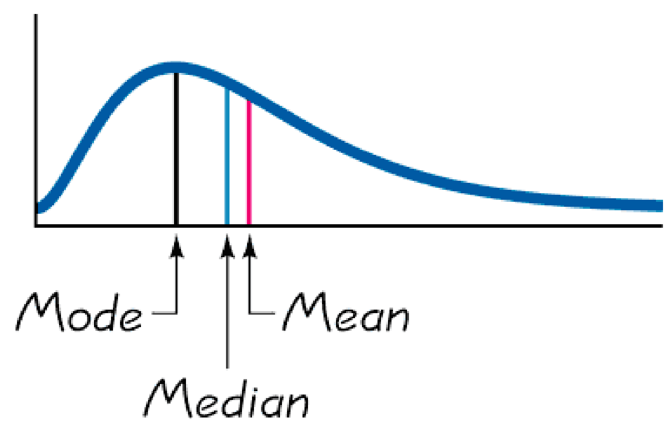
Skewness



(b) Symmetric



(a) Skewed to the Left (Negatively)



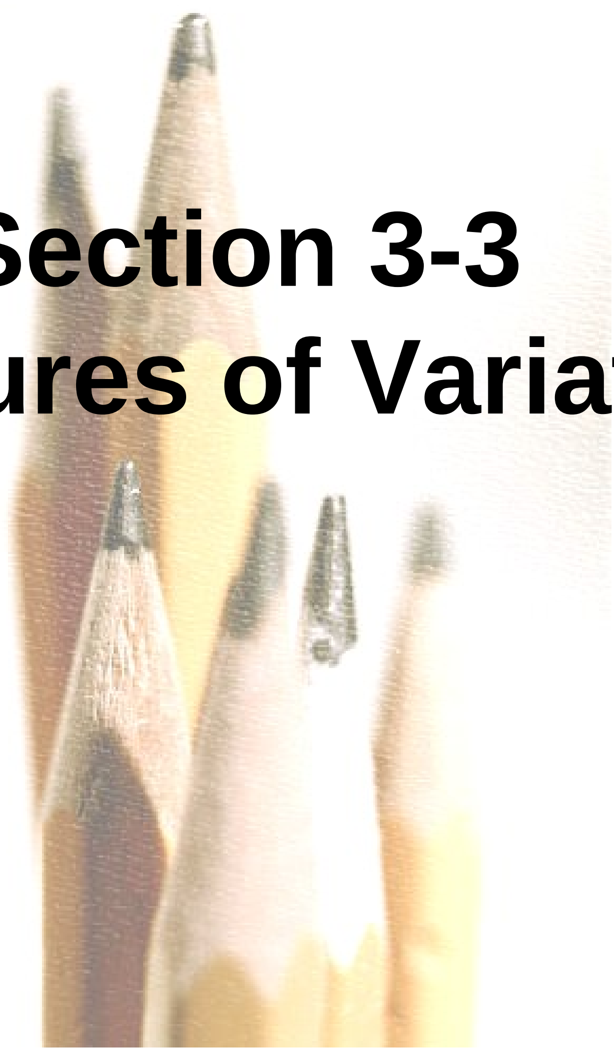
(c) Skewed to the Right (Positively)

Recap

In this section we have discussed:

- **Types of measures of center**
  - **Mean**
  - **Median**
  - **Mode**
- **Mean from a frequency distribution**
- **Weighted means**
- **Best measures of center**
- **Skewness**

Section 3-3
Measures of Variation



Key Concept

Discuss characteristics of variation, in particular, measures of variation, such as standard deviation, for analyzing data.

Make understanding and interpreting the standard deviation a priority.

Part 1

Basics Concepts of Measures of Variation

Definition

The **range** of a set of data values is the difference between the maximum data value and the minimum data value.

Range = (maximum value) – (minimum value)

It is very sensitive to extreme values; therefore not as useful as other measures of variation.

Round-Off Rule for Measures of Variation

When rounding the value of a measure of variation, carry one more decimal place than is present in the original set of data.

Round only the final answer, not values in the middle of a calculation.

Definition

The standard deviation of a set of sample values, denoted by *s*, is a measure of variation of values about the mean.

Pearson Education

Sample Standard Deviation Formula

$$s = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n - 1}}$$

Sample Standard Deviation (Shortcut Formula)

$$s = \sqrt{\frac{n\Sigma(x^2) - (\Sigma x)^2}{n\,(n-1)}}$$

Standard Deviation - Important Properties

- The standard deviation is a measure of variation of all values from the **mean**.
- The value of the standard deviation s is usually positive.
- The value of the standard deviation s can increase dramatically with the inclusion of one or more outliers (data values far away from all others).
- The units of the standard deviation s are the same as the units of the original data values.

Comparing Variation in Different Samples

It’s a good practice to compare two sample standard deviations only when the sample means are approximately the same.

When comparing variation in samples with very different means, it is better to use the coefficient of variation, which is defined later in this section.

Population Standard Deviation

$$\sigma = \sqrt{\frac{\Sigma (x - \mu)^2}{N}}$$

This formula is similar to the previous formula, but instead, the population mean and population size are used.

Variance

- The **variance** of a set of values is a measure of variation equal to the square of the standard deviation.

- Sample variance: s^2 - Square of the sample standard deviation s

- Population variance: σ^2 - Square of the population standard deviation σ

Unbiased Estimator

The sample variance s^2 is an unbiased estimator of the population variance σ^2, which means values of s^2 tend to target the value of σ^2 instead of systematically tending to overestimate or underestimate σ^2.

Variance - Notation

s = *sample* standard deviation

s^2 = *sample* variance

σ = *population* standard deviation

σ^2 = *population* variance

Part 2

Beyond the Basics of Measures of Variation

Range Rule of Thumb

is based on the principle that for many data sets, the vast majority (such as 95%) of sample values lie within two standard deviations of the mean.

Range Rule of Thumb for Interpreting a Known Value of the Standard Deviation

Informally define *usual* values in a data set to be those that are typical and not too extreme. Find rough estimates of the minimum and maximum "usual" sample values as follows:

Minimum "usual" value $=$ (mean) – 2 $\times$ (standard deviation)

Maximum "usual" value $=$ (mean) + 2 $\times$ (standard deviation)

Range Rule of Thumb for Estimating a Value of the Standard Deviation *s*

To roughly estimate the standard deviation from a collection of known sample data use

$$s \approx \frac{\text{range}}{4}$$

where

range = (maximum value) – (minimum value)

Properties of the Standard Deviation

- Measures the variation among data values
- Values close together have a small standard deviation, but values with much more variation have a larger standard deviation
- Has the same units of measurement as the original data

Pearson Education

Properties of the Standard Deviation

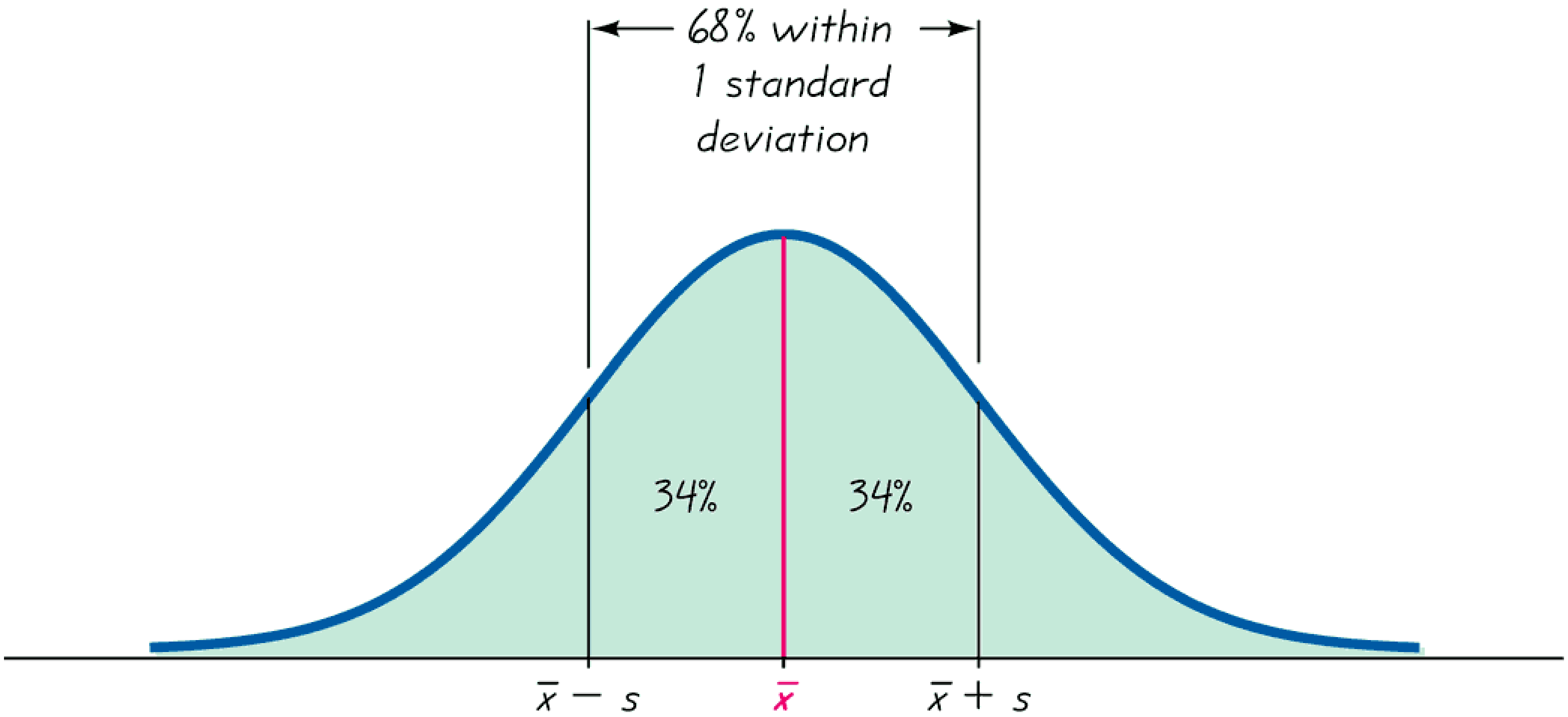
- **For many data sets, a value is *unusual* if it differs from the mean by more than two standard deviations**
- **Compare standard deviations of two different data sets only if the they use the same scale and units, and they have means that are approximately the same**

Empirical (or 68-95-99.7) Rule

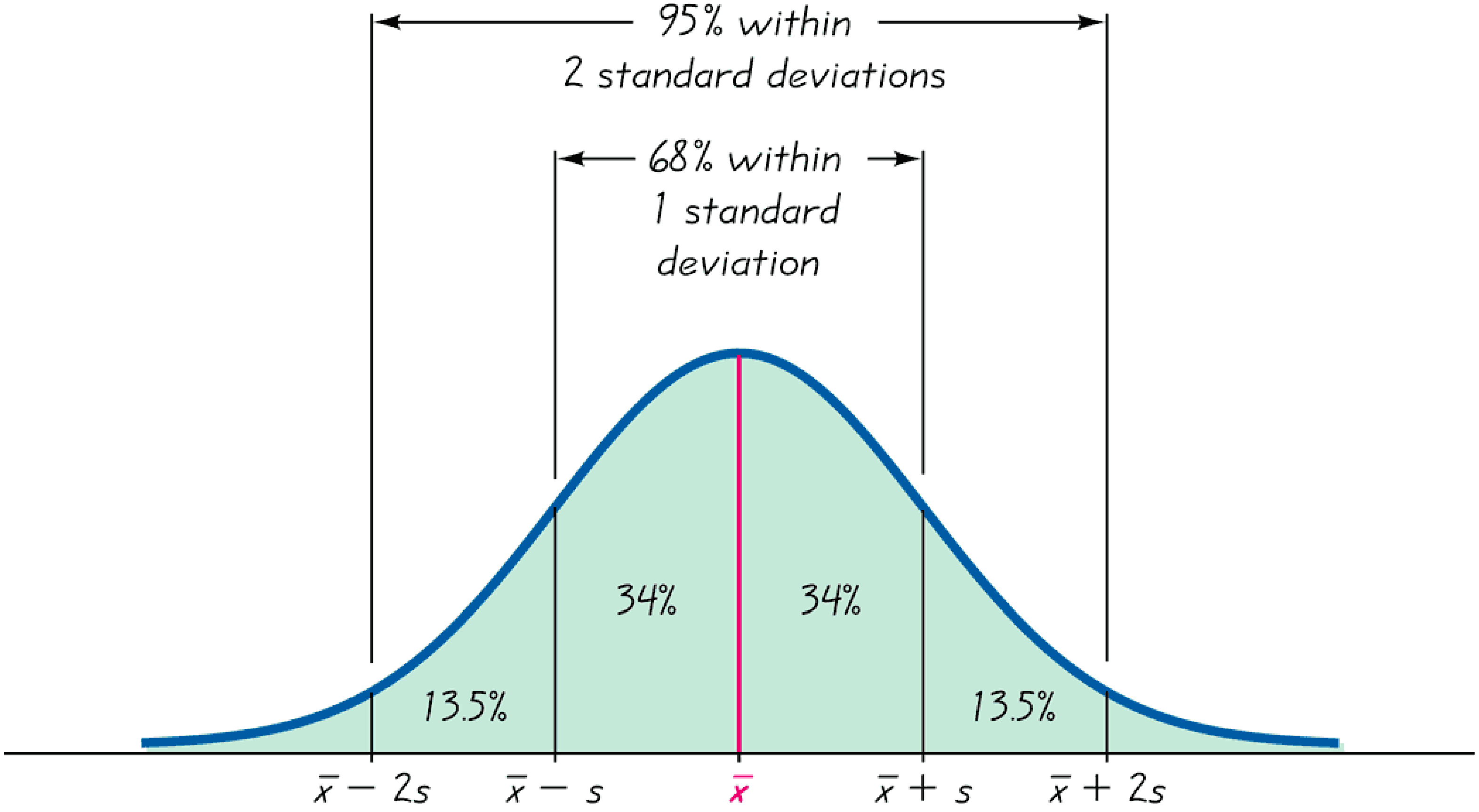
For data sets having a distribution that is approximately bell shaped, the following properties apply:

- About 68% of all values fall within 1 standard deviation of the mean.
- About 95% of all values fall within 2 standard deviations of the mean.
- About 99.7% of all values fall within 3 standard deviations of the mean.

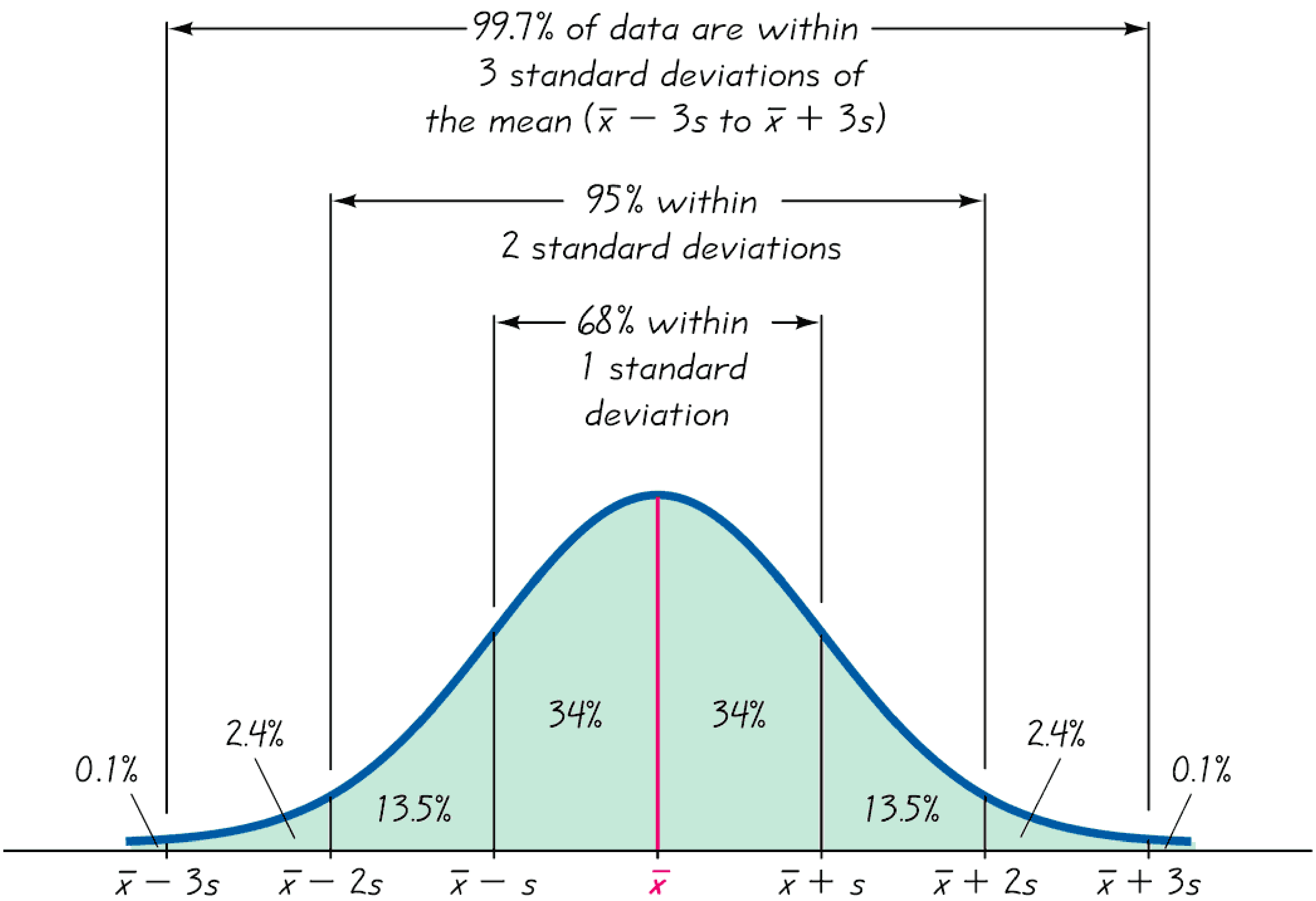
The Empirical Rule



The Empirical Rule



The Empirical Rule



Chebyshev's Theorem

The proportion (or fraction) of any set of data lying within K standard deviations of the mean is always at least $1–1/K^2$, where K is any positive number greater than 1.

- For $K = 2$, at least 3/4 (or 75%) of all values lie within 2 standard deviations of the mean.
- For $K = 3$, at least 8/9 (or 89%) of all values lie within 3 standard deviations of the mean.

Rationale for using *n* – 1 versus *n*

There are only *n* – 1 independent values. With a given mean, only *n* – 1 values can be freely assigned any number before the last value is determined.

Dividing by *n* – 1 yields better results than dividing by *n*. It causes s^2 to target σ^2 whereas division by *n* causes s^2 to underestimate σ^2.

Pearson Education

Coefficient of Variation

The **coefficient of variation** (or **CV**) for a set of nonnegative sample or population data, expressed as a percent, describes the standard deviation relative to the mean.

Sample

$$CV = \frac{s}{\bar{x}} \bullet 100\%$$

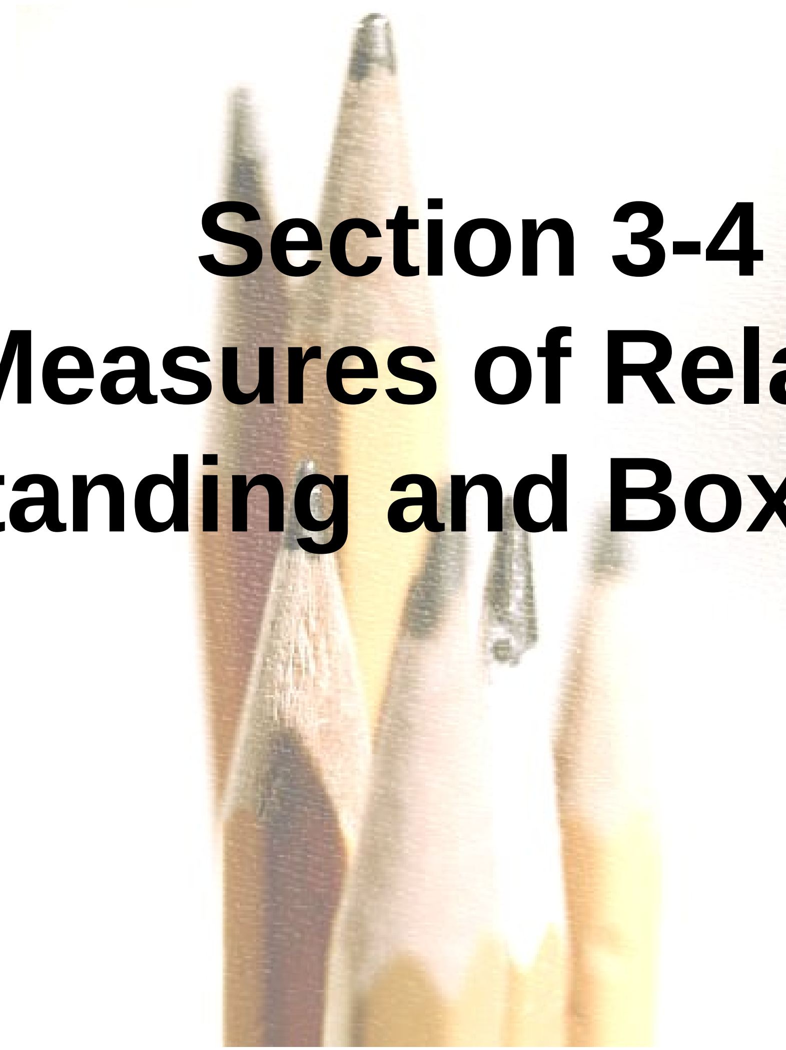
Population

$$CV = \frac{\sigma}{\mu} \bullet 100\%$$

Recap

In this section we have looked at:

- Range
- Standard deviation of a sample and population
- Variance of a sample and population
- Range rule of thumb
- Empirical distribution
- Chebyshev's theorem
- Coefficient of variation (CV)



Section 3-4
Measures of Relative Standing and Boxplots

Key Concept

This section introduces measures of relative standing, which are numbers showing the location of data values relative to the other values within a data set. They can be used to compare values from different data sets, or to compare values within the same data set. The most important concept is the *z* score. We will also discuss percentiles and quartiles, as well as a new statistical graph called the boxplot.

Part 1

Basics of *z* Scores, Percentiles, Quartiles, and Boxplots

Z score

- *z* **Score** **(or standardized value)**

  the number of standard deviations that a given value *x* is above or below the mean

Measures of Position *z* Score

Sample

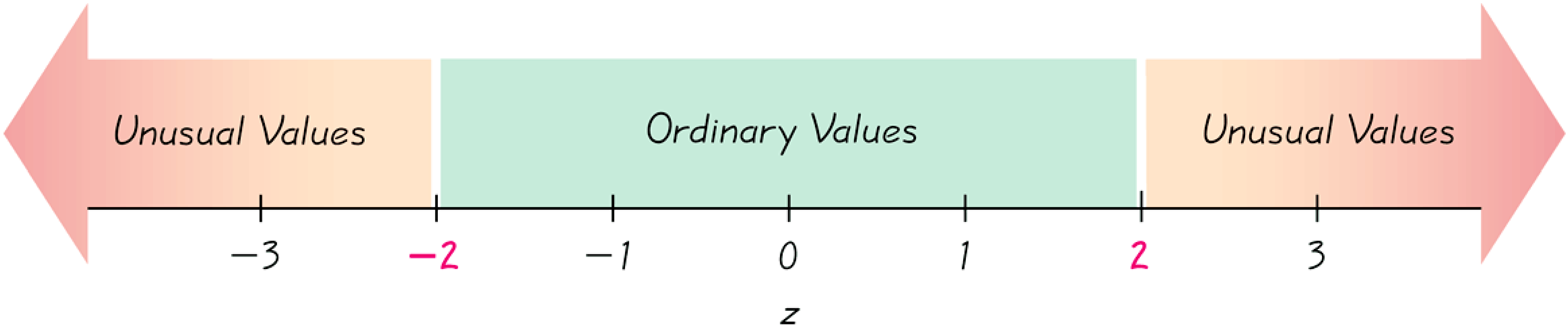
$$z = \frac{x - \bar{x}}{s}$$

Population

$$z = \frac{x - \mu}{\sigma}$$

Round *z* scores to 2 decimal places

Interpreting *Z* Scores



Whenever a value is less than the mean, its corresponding *z* score is negative

Ordinary values: $-2 \le z \text{ score} \le 2$

Unusual Values: $z \text{ score} < -2$ or $z \text{ score} > 2$

Percentiles

are measures of location. There are 99 **percentiles** denoted P_1, P_2, . . . P_{99}, which divide a set of data into 100 groups with about 1% of the values in each group.

Finding the Percentile of a Data Value

$$\text{Percentile of value } x = \frac{\text{number of values less than } x}{\text{total number of values}} \cdot 100$$

Converting from the *k*th Percentile to the Corresponding Data Value

Notation

$$L = \frac{k}{100} \cdot n$$

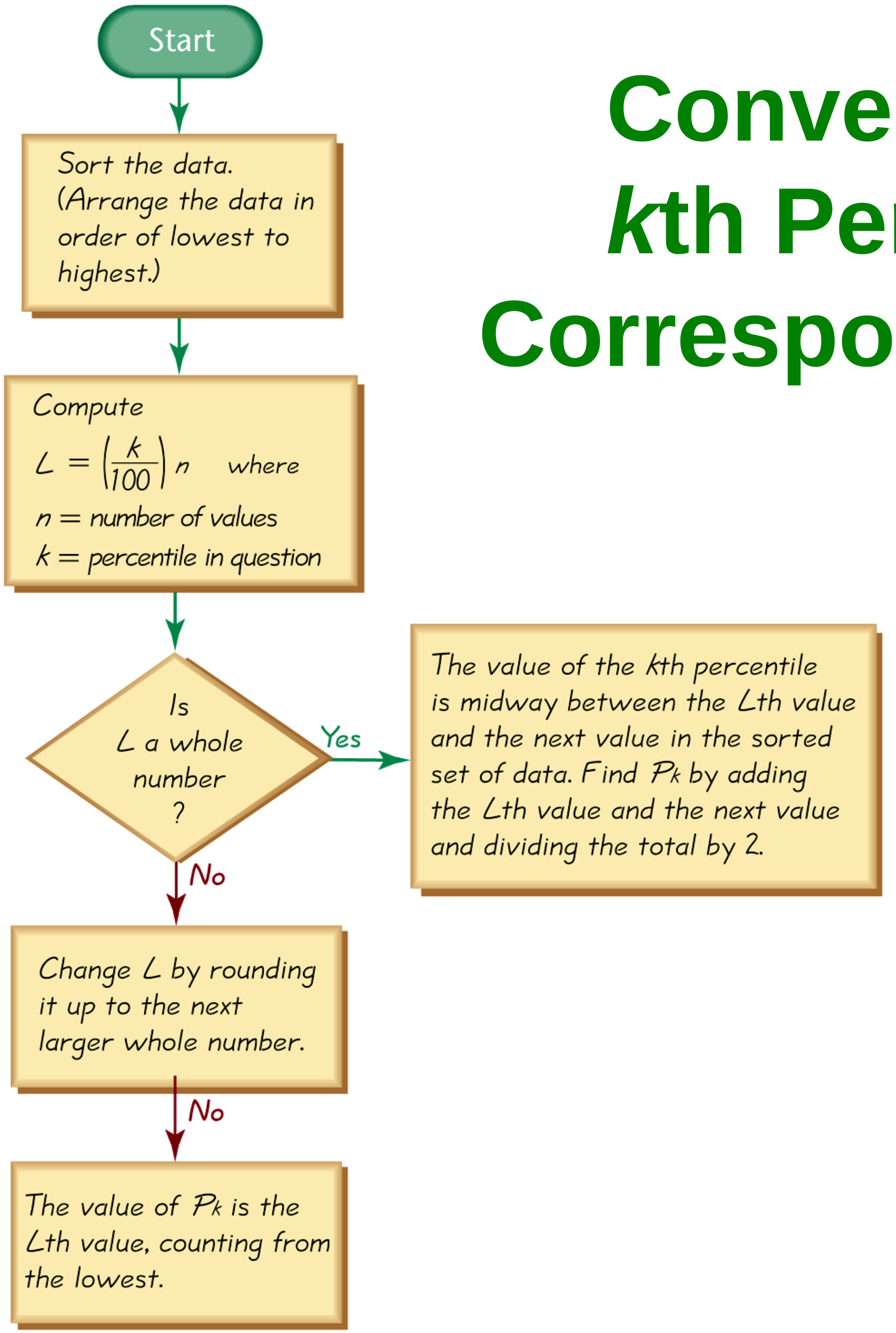
n total number of values in the data set

k percentile being used

L locator that gives the **position** of a value

P_k *k*th percentile

Converting from the *k*th Percentile to the Corresponding Data Value



Quartiles

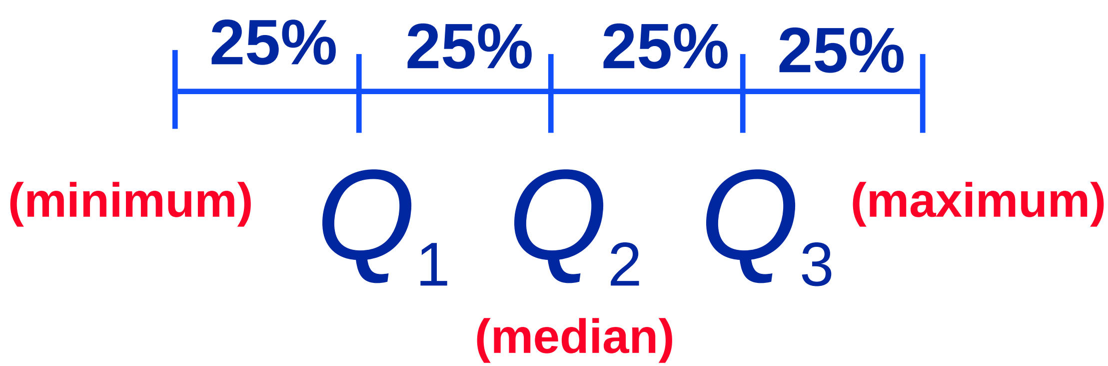
Are measures of location, denoted Q_1, Q_2, and Q_3, which divide a set of data into four groups with about 25% of the values in each group.

- **Q_1 (First Quartile) separates the bottom 25% of sorted values from the top 75%.**
- **Q_2 (Second Quartile) same as the median; separates the bottom 50% of sorted values from the top 50%.**
- **Q_3 (Third Quartile) separates the bottom 75% of sorted values from the top 25%.**

Quartiles

Q_1, Q_2, Q_3

divide ranked scores into four equal parts



Some Other Statistics

- **Interquartile Range (or IQR):** $Q_3 - Q_1$
- **Semi-interquartile Range:** $\dfrac{Q_3 - Q_1}{2}$
- **Midquartile:** $\dfrac{Q_3 + Q_1}{2}$
- **10 - 90 Percentile Range:** $P_{90} - P_{10}$

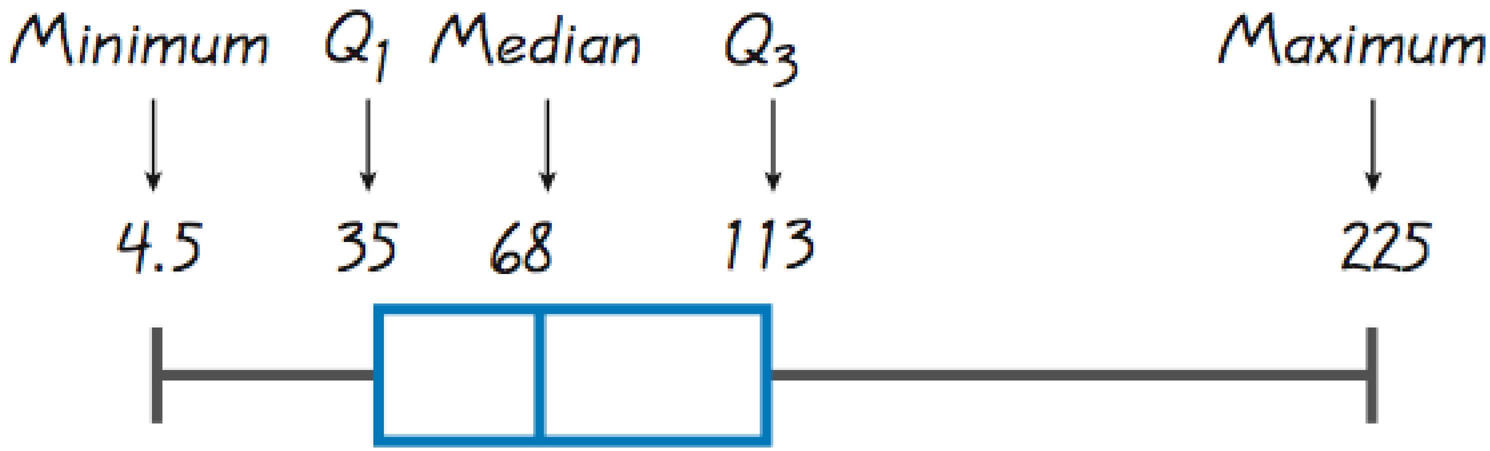
5-Number Summary

- For a set of data, the **5-number summary** consists of the minimum value; the first quartile Q_1; the median (or second quartile Q_2); the third quartile, Q_3; and the maximum value.

Boxplot

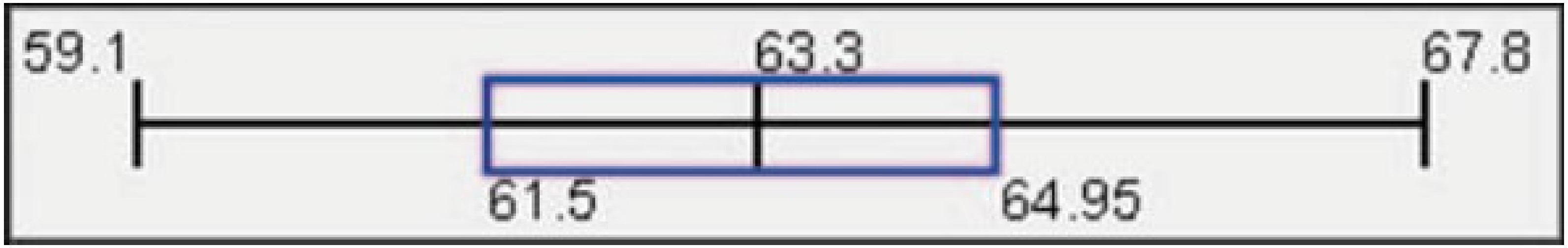
- A **boxplot** (or **box-and-whisker-diagram**) is a graph of a data set that consists of a line extending from the minimum value to the maximum value, and a box with lines drawn at the first quartile, Q_1; the median; and the third quartile, Q_3.

Boxplots



Boxplot of Movie Budget Amounts

Boxplots - Normal Distribution



Normal Distribution:
Heights from a Simple Random Sample of Women

Boxplots - Skewed Distribution

Skewed Distribution:
Salaries (in thousands of dollars) of NCAA Football Coaches

Part 2

Outliers and Modified Boxplots

Outliers

- An **outlier** is a value that lies very far away from the vast majority of the other values in a data set.

Important Principles

- An outlier can have a dramatic effect on the mean.
- An outlier can have a dramatic effect on the standard deviation.
- An outlier can have a dramatic effect on the scale of the histogram so that the true nature of the distribution is totally obscured.

Outliers for Modified Boxplots

For purposes of constructing *modified boxplots*, we can consider outliers to be data values meeting specific criteria.

In modified boxplots, a data value is an outlier if it is . . .

above Q_3 by an amount greater than $1.5 \times$ IQR

or

below Q_1 by an amount greater than $1.5 \times$ IQR

Modified Boxplots

Boxplots described earlier are called skeletal (or regular) boxplots.

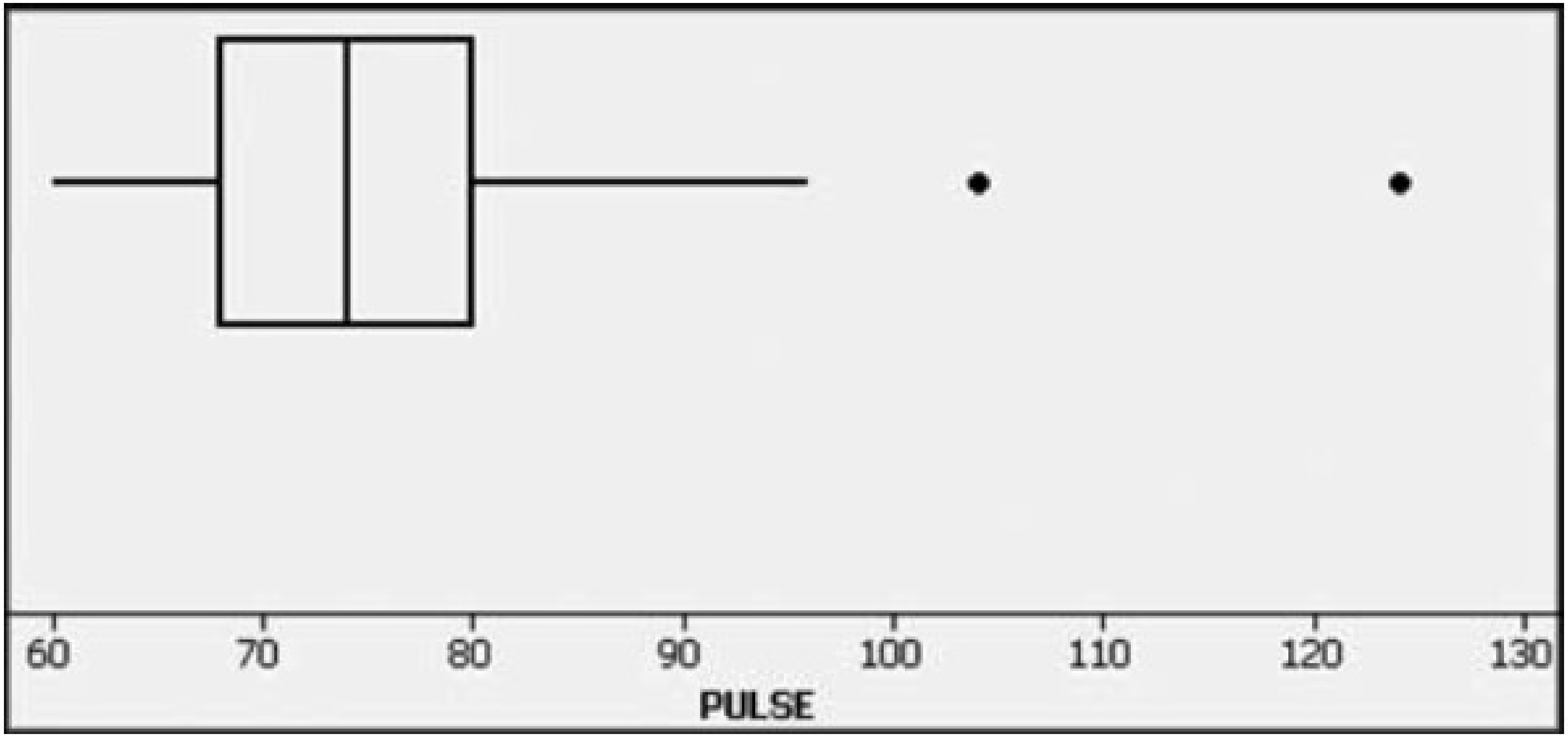
Some statistical packages provide modified boxplots which represent outliers as special points.

Modified Boxplot Construction

A modified boxplot is constructed with these specifications:

- **A special symbol (such as an asterisk) is used to identify outliers.**
- **The solid horizontal line extends only as far as the minimum data value that is not an outlier and the maximum data value that is not an outlier.**

Modified Boxplots - Example



Pulse rates of females listed in Data Set 1 in Appendix B.

Recap

In this section we have discussed:

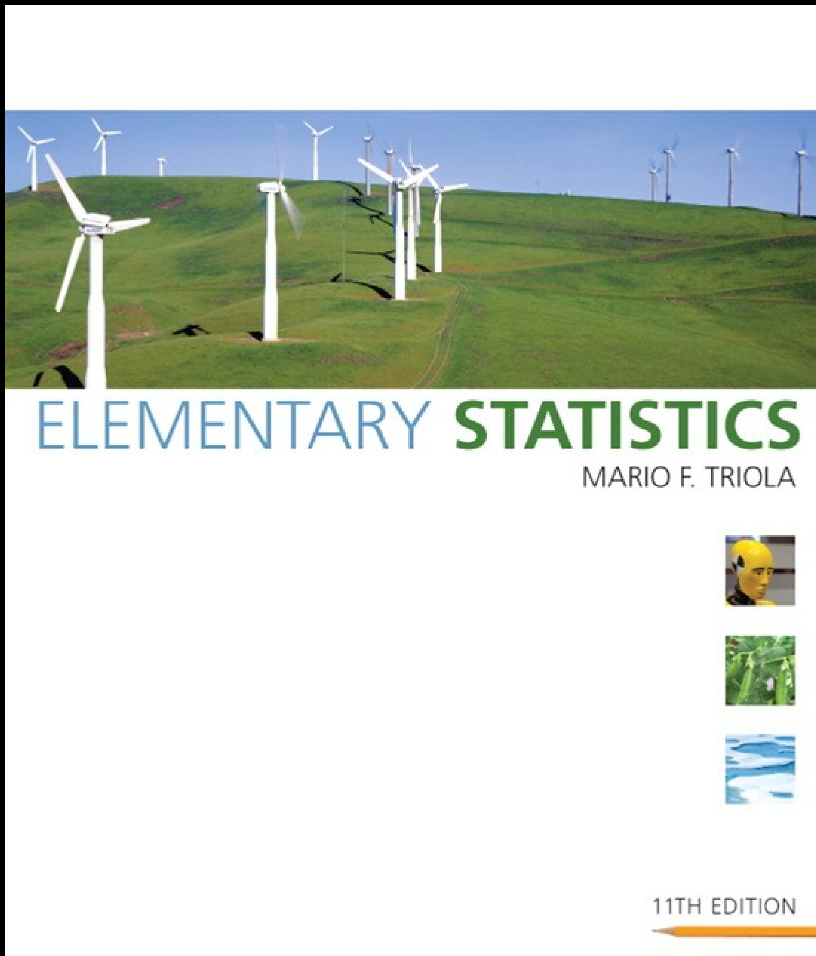
- *z* Scores
- *z* Scores and unusual values
- Percentiles
- Quartiles
- Converting a percentile to corresponding data values
- Other statistics
- 5-number summary
- Boxplots and modified boxplots
- Effects of outliers

Putting It All Together

Always consider certain key factors:

- Context of the data
- Source of the data
- Sampling Method
- Measures of Center
- Measures of Variation
- Distribution
- Outliers
- Changing patterns over time
- Conclusions
- Practical Implications

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 4
Probability

4-1 Review and Preview

4-2 Basic Concepts of Probability

4-3 Addition Rule

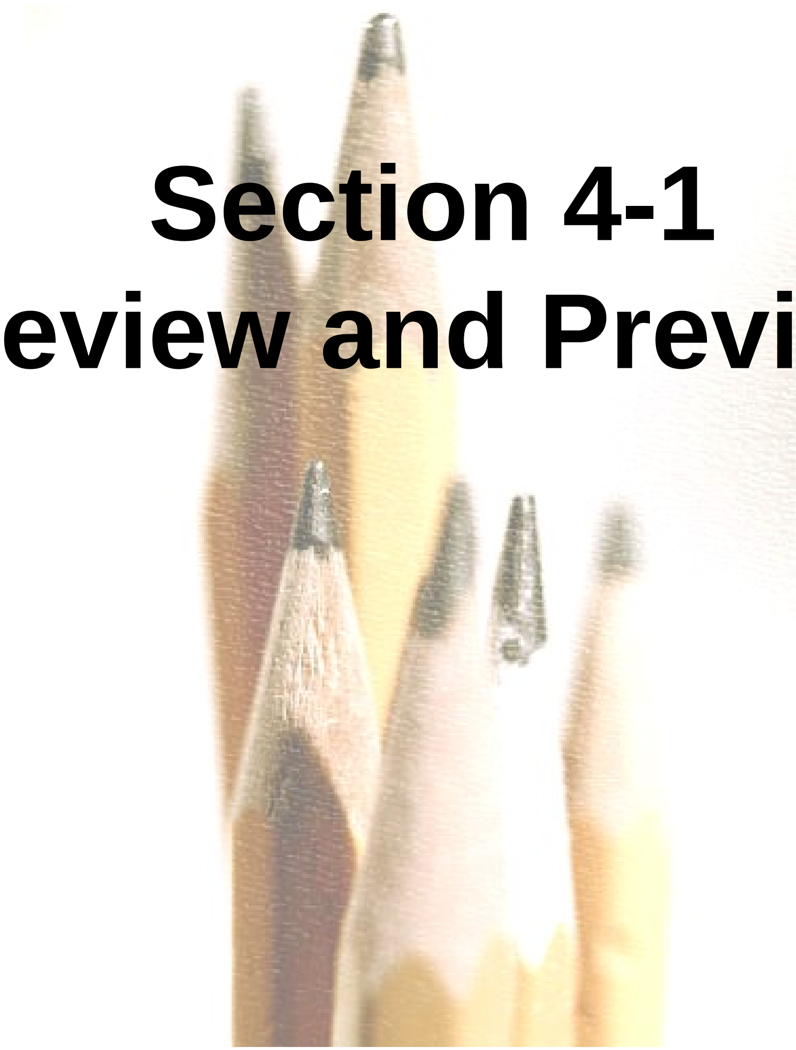
4-4 Multiplication Rule: Basics

4-5 Multiplication Rule: Complements and Conditional Probability

4-6 Probabilities Through Simulations

4-7 Counting

Section 4-1
Review and Preview



Review

Necessity of sound sampling methods.

Common measures of characteristics of data

Mean

Standard deviation

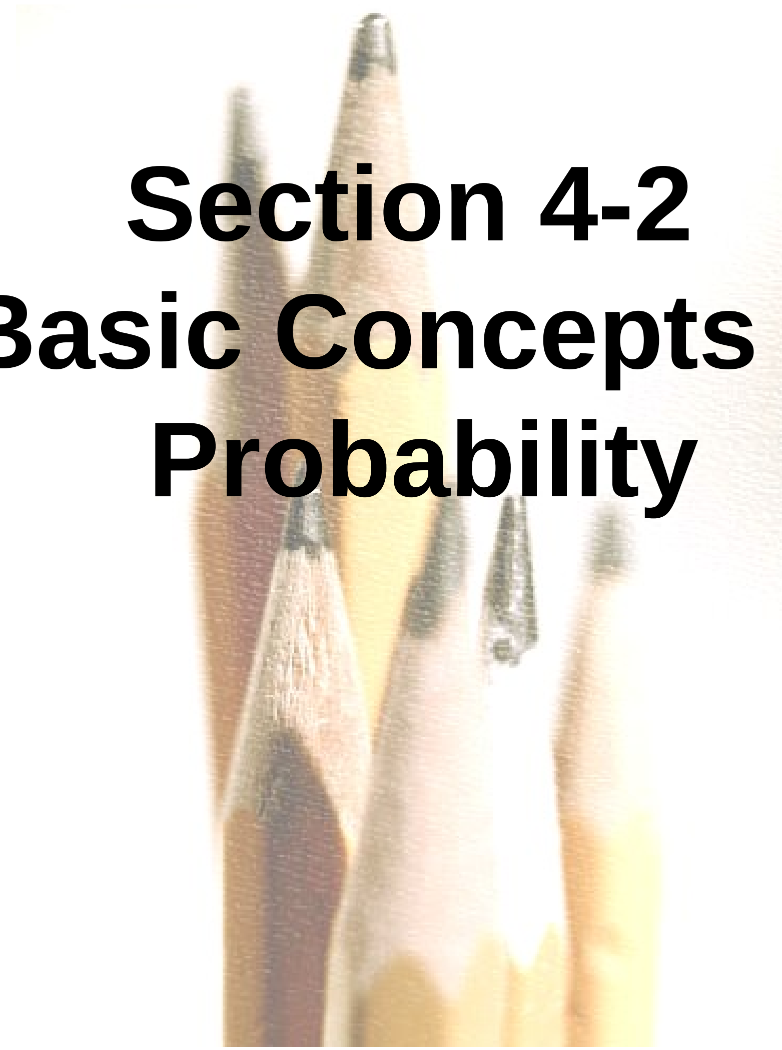
Preview

Rare Event Rule for Inferential Statistics:

If, under a given assumption, the probability of a particular observed event is extremely small, we conclude that the assumption is probably not correct.

Statisticians use the rare event rule for inferential statistics.

Section 4-2
Basic Concepts of Probability



Key Concept

This section presents three approaches to finding the probability of an event.

The most important objective of this section is to learn how to interpret probability values.

Part 1

Basics of Probability

Events and Sample Space

- **Event**

  any collection of results or outcomes of a procedure

- **Simple Event**

  an outcome or an event that cannot be further broken down into simpler components

- **Sample Space**

  for a procedure consists of all possible simple events; that is, the sample space consists of all outcomes that cannot be broken down any further

Notation for Probabilities

P - **denotes a probability.**

A*, *B*, and *C - **denotes specific events.**

***P*(*A*)** - **denotes the probability of event *A* occurring.**

Basic Rules for Computing Probability

Rule 1: Relative Frequency Approximation of Probability

Conduct (or observe) a procedure, and count the number of times event *A* actually occurs. Based on these actual results, *P*(*A*) is approximated as follows:

$$P(A) = \frac{\text{\# of times } A \text{ occurred}}{\text{\# of times procedure was repeated}}$$

Basic Rules for Computing Probability - continued

Rule 2: Classical Approach to Probability (Requires Equally Likely Outcomes)

Assume that a given procedure has n different simple events and that each of those simple events has an equal chance of occurring. If event A can occur in s of these n ways, then

$$P(A) = \frac{s}{n} = \frac{\text{number of ways } A \text{ can occur}}{\text{number of different simple events}}$$

Basic Rules for Computing Probability - continued

Rule 3: Subjective Probabilities

$P(A)$, the probability of event A, is estimated by using knowledge of the relevant circumstances.

Law of Large Numbers

As a procedure is repeated again and again, the relative frequency probability of an event tends to approach the actual probability.

Probability Limits

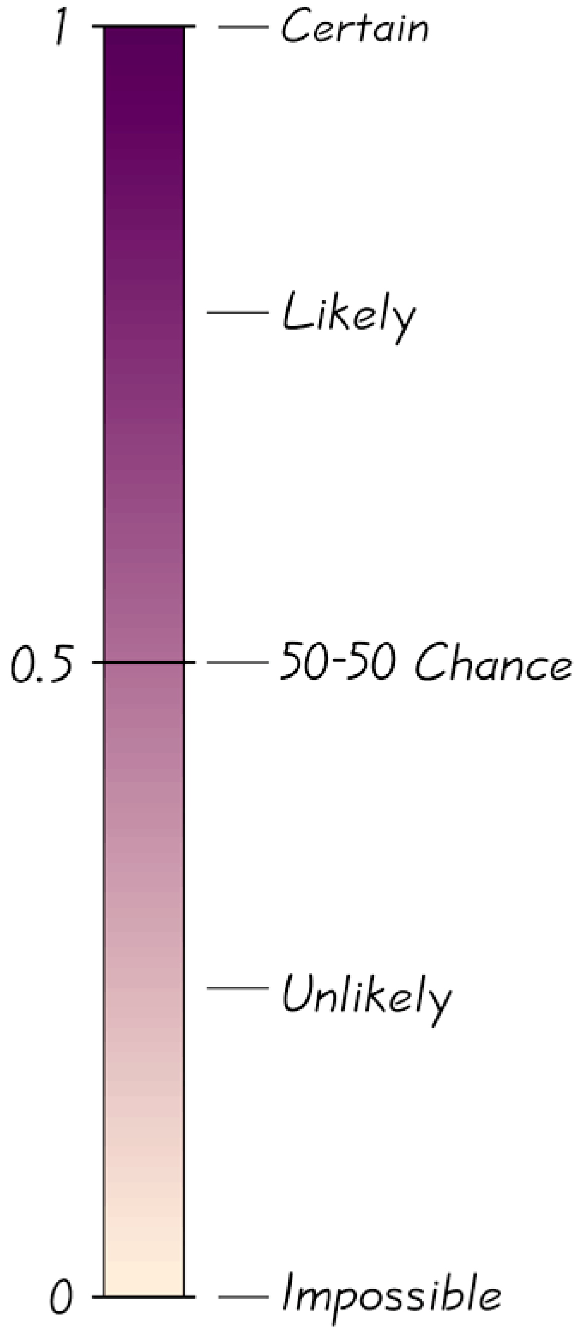
Always express a probability as a fraction or decimal number between 0 and 1.

- **The probability of an impossible event is 0.**

- **The probability of an event that is certain to occur is 1.**

- **For any event *A*, the probability of *A* is between 0 and 1 inclusive.**
  That is, $0 \leq P(A) \leq 1$.

Possible Values for Probabilities



Complementary Events

The complement of event *A*, denoted by $\overline{A}$, consists of all outcomes in which the event *A* does not occur.

Rounding Off Probabilities

When expressing the value of a probability, either give the exact fraction or decimal or round off final decimal results to three significant digits. (*Suggestion*: When a probability is not a simple fraction such as 2/3 or 5/9, express it as a decimal so that the number can be better understood.)

Part 2

Beyond the Basics of Probability: Odds

Odds

The actual odds against event A occurring are the ratio $P(A)/P(\overline{A})$, usually expressed in the form of a:b (or "a to b"), where a and b are integers having no common factors.

The actual odds in favor of event A occurring are the ratio $P(\overline{A})/P(A)$, which is the reciprocal of the actual odds against the event. If the odds against A are a:b, then the odds in favor of A are b:a.

The payoff odds against event A occurring are the ratio of the net profit (if you win) to the amount bet.

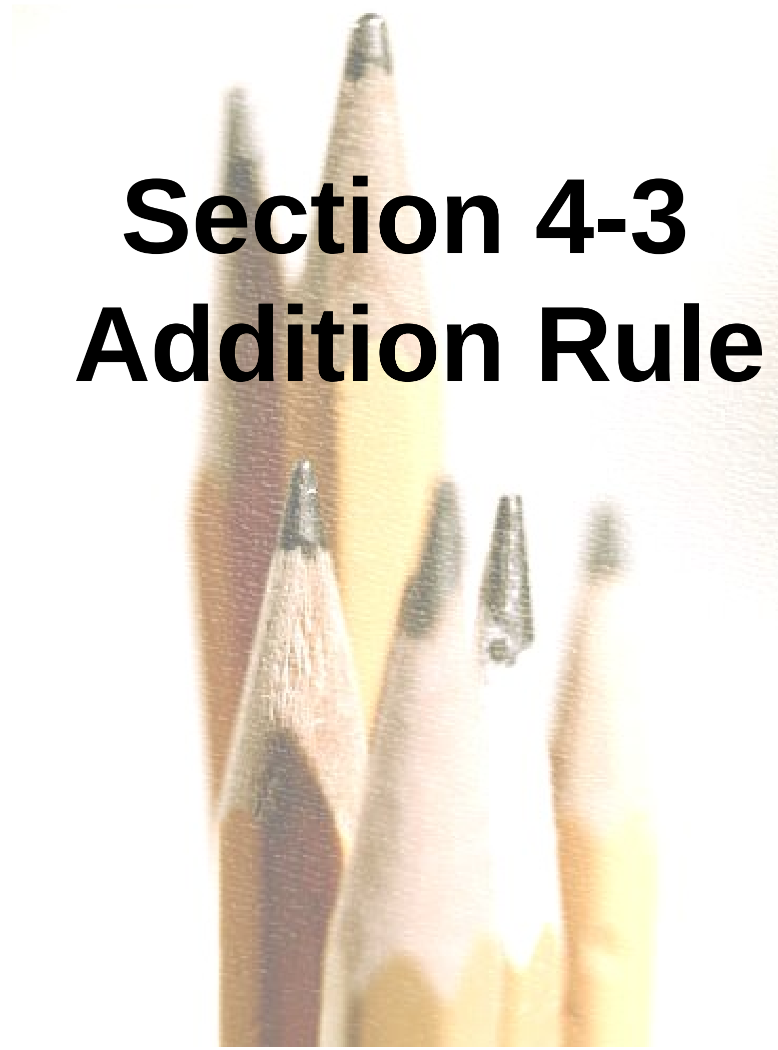
payoff odds against event A = (net profit) : (amount bet)

Recap

In this section we have discussed:

- **Rare event rule for inferential statistics.**
- **Probability rules.**
- **Law of large numbers.**
- **Complementary events.**
- **Rounding off probabilities.**
- **Odds.**

Section 4-3
Addition Rule



Key Concept

This section presents the addition rule as a device for finding probabilities that can be expressed as $P(A \text{ or } B)$, the probability that either event A occurs or event B occurs (or they both occur) as the single outcome of the procedure.

The key word in this section is “or.” It is the *inclusive or*, which means either one or the other or both.

Compound Event

Compound Event

any event combining 2 or more simple events

Notation

$P(A \text{ or } B) = P$ (in a single trial, event A occurs or event B occurs or they both occur)

General Rule for a Compound Event

When finding the probability that event *A* occurs or event *B* occurs, find the total number of ways *A* can occur and the number of ways *B* can occur, but find that total in such a way that no outcome is counted more than once.

Compound Event

Formal Addition Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

where $P(A \text{ and } B)$ denotes the probability that A and B both occur at the same time as an outcome in a trial of a procedure.

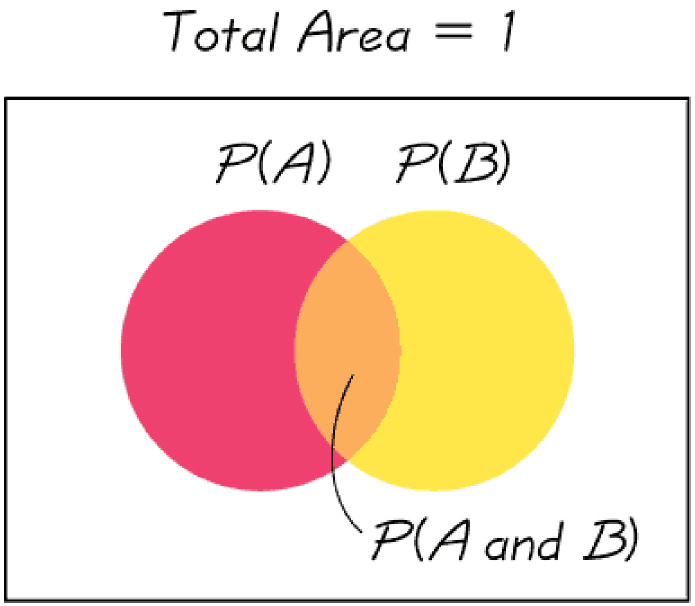
Compound Event

Intuitive Addition Rule

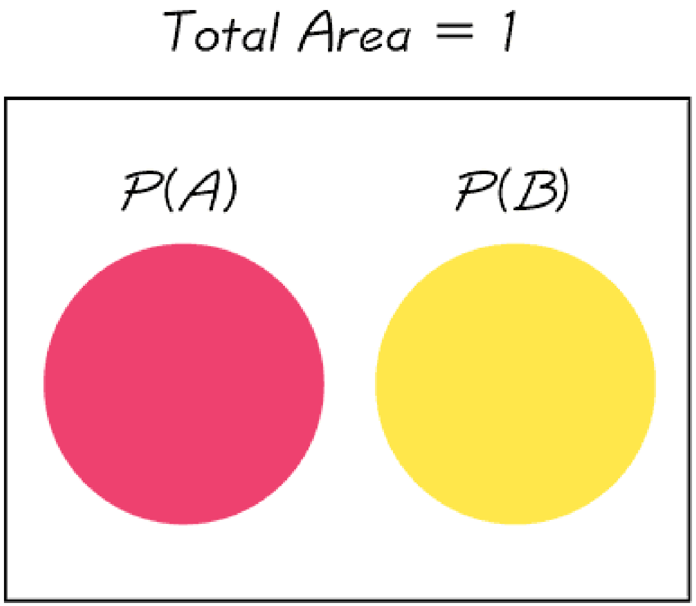
To find $P(A \text{ or } B)$, find the sum of the number of ways event A can occur and the number of ways event B can occur, adding in such a way that every outcome is counted only once. $P(A \text{ or } B)$ is equal to that sum, divided by the total number of outcomes in the sample space.

Disjoint or Mutually Exclusive

Events *A* and *B* are disjoint (or mutually exclusive) if they cannot occur at the same time. (That is, disjoint events do not overlap.)



Venn Diagram for Events That Are Not Disjoint



Venn Diagram for Disjoint Events

Complementary Events

$P(A)$ and $P(\bar{A})$ are disjoint

It is impossible for an event and its complement to occur at the same time.

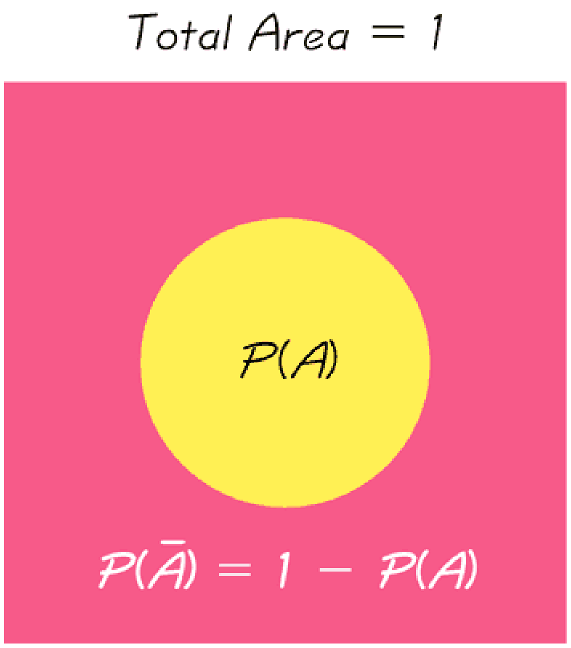
Rule of Complementary Events

$$P(A) + P(\bar{A}) = 1$$

$$P(\bar{A}) = 1 - P(A)$$

$$P(A) = 1 - P(\bar{A})$$

Venn Diagram for the Complement of Event *A*



Recap

In this section we have discussed:

- **Compound events.**
- **Formal addition rule.**
- **Intuitive addition rule.**
- **Disjoint events.**
- **Complementary events.**

Section 4-4
Multiplication Rule: Basics

Key Concept

The basic multiplication rule is used for finding *P*(*A* and *B*), the probability that event *A* occurs in a first trial and event *B* occurs in a second trial.

If the outcome of the first event *A* somehow affects the probability of the second event *B*, it is important to adjust the probability of *B* to reflect the occurrence of event *A*.

Notation

$P(A$ and $B) =$

$P($event A occurs in a first trial and event B occurs in a second trial)

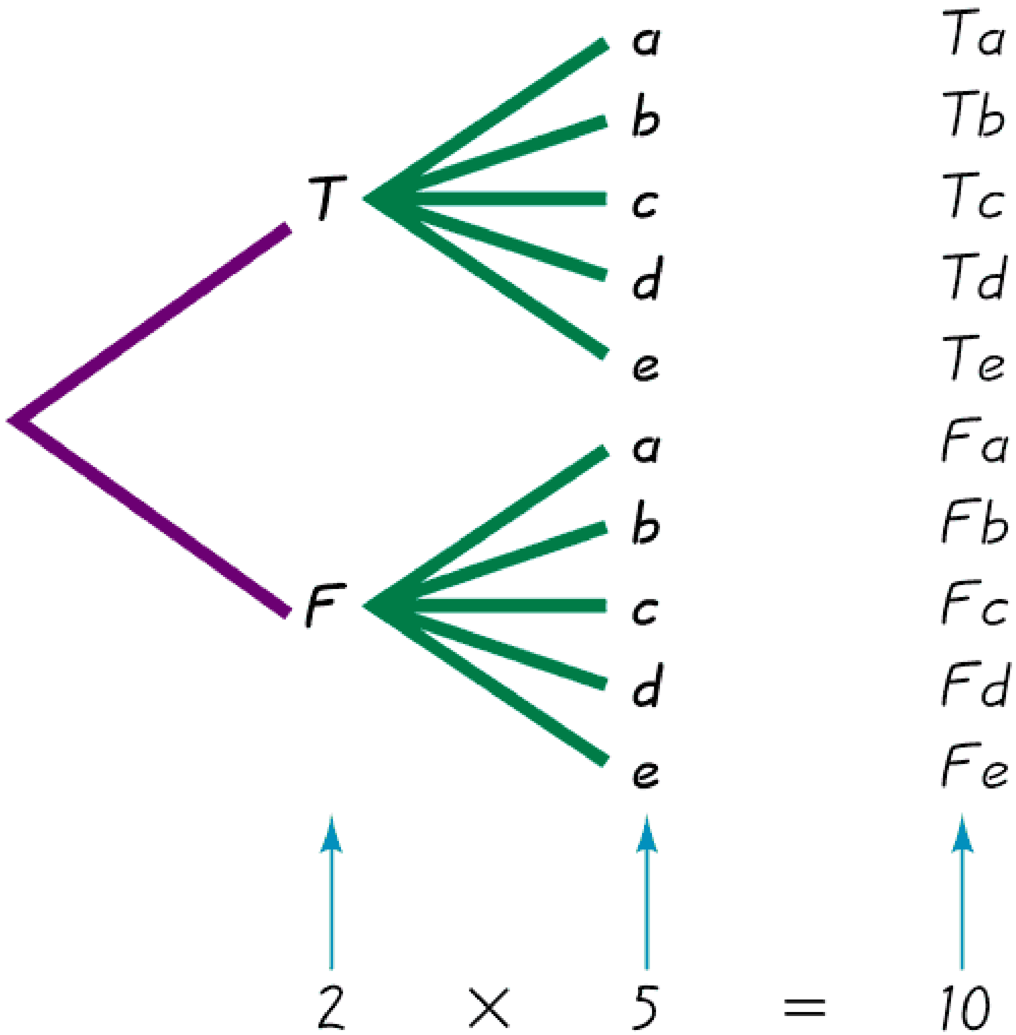
Tree Diagrams

A tree diagram is a picture of the possible outcomes of a procedure, shown as line segments emanating from one starting point. These diagrams are sometimes helpful in determining the number of possible outcomes in a sample space, if the number of possibilities is not too large.

Tree Diagrams

This figure summarizes the possible outcomes for a true/false question followed by a multiple choice question.

Note that there are 10 possible combinations.



Conditional Probability Key Point

We must adjust the probability of the second event to reflect the outcome of the first event.

Conditional Probability Important Principle

The probability for the second event *B* should take into account the fact that the first event *A* has already occurred.

Notation for Conditional Probability

$P(B|A)$ represents the probability of event B occurring after it is assumed that event A has already occurred (read $B|A$ as "B given A.")

Dependent and Independent

Two events *A* and *B* are **independent** if the occurrence of one does not affect the *probability* of the occurrence of the other. (Several events are similarly independent if the occurrence of any does not affect the probabilities of the occurrence of the others.) If *A* and *B* are not independent, they are said to be **dependent**.

Dependent Events

Two events are dependent if the occurrence of one of them affects the *probability* of the occurrence of the other, but this does not necessarily mean that one of the events is a *cause* of the other.

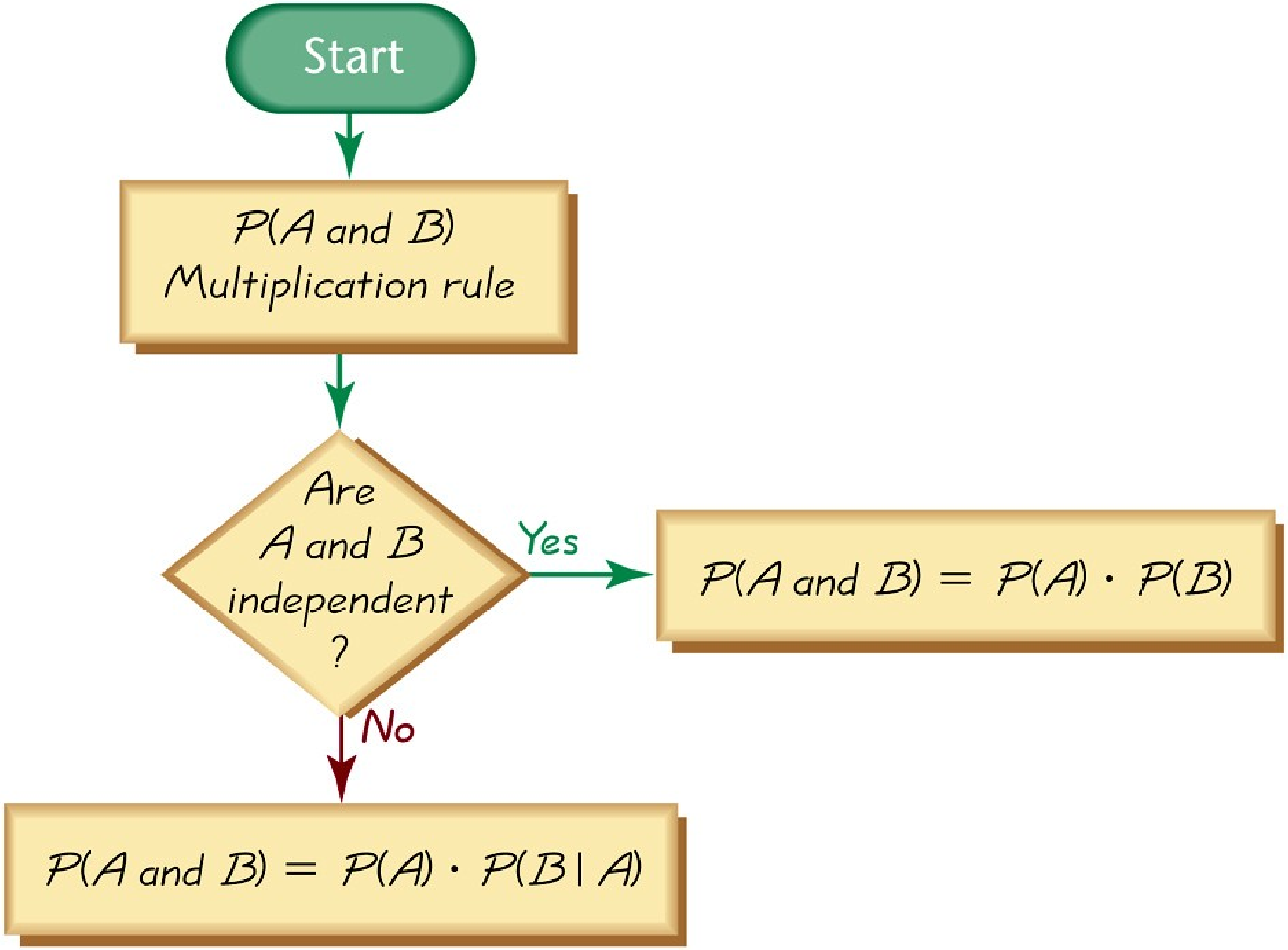
Formal Multiplication Rule

- $P(A \text{ and } B) = P(A) \cdot P(B|A)$
- Note that if A and B are independent events, P(B|A) is really the same as P(B).

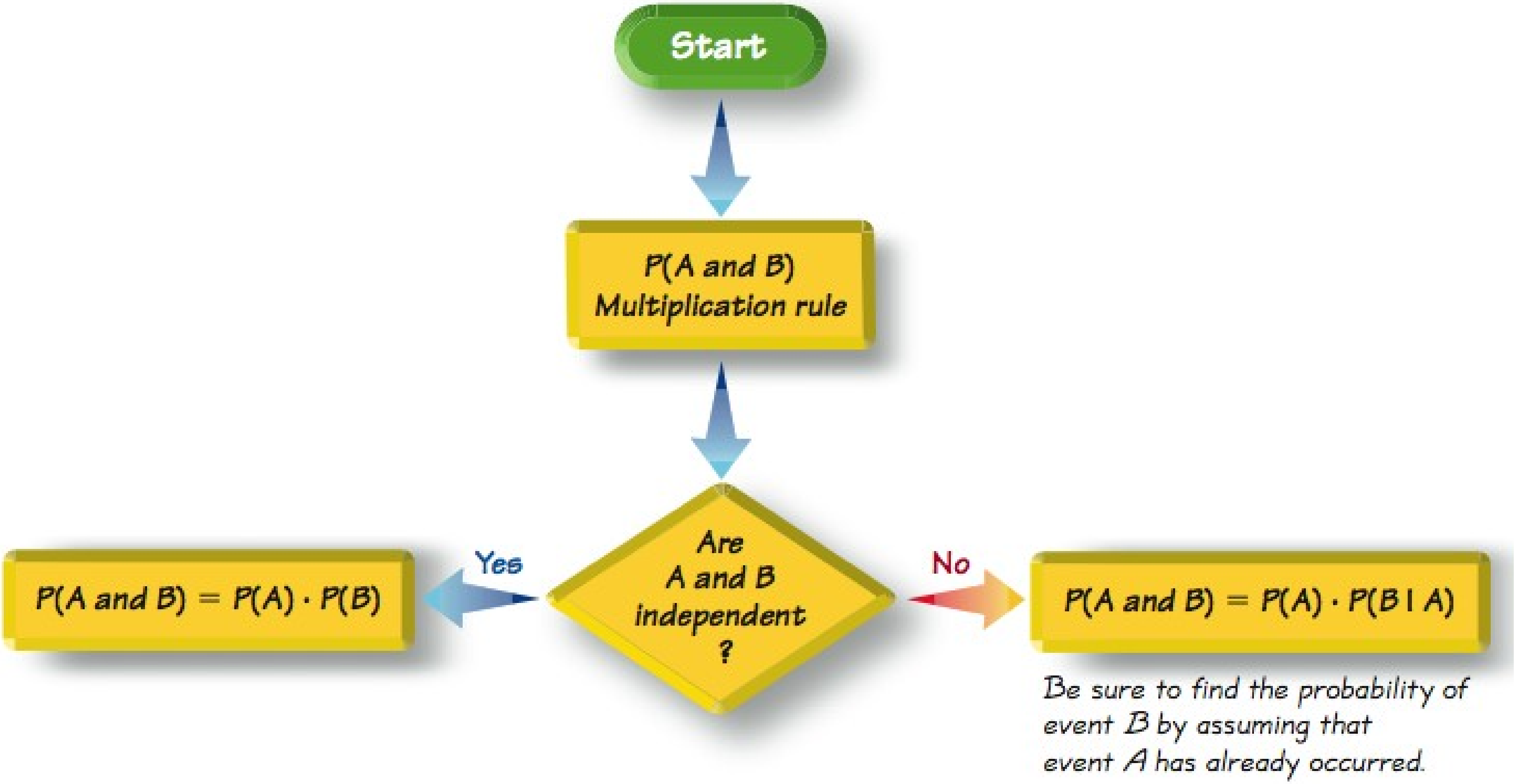
Intuitive Multiplication Rule

When finding the probability that event *A* occurs in one trial and event *B* occurs in the next trial, multiply the probability of event *A* by the probability of event *B*, but be sure that the probability of event *B* takes into account the previous occurrence of event *A*.

Applying the Multiplication Rule



Applying the Multiplication Rule



Caution

When applying the multiplication rule, always consider whether the events are independent or dependent, and adjust the calculations accordingly.

Multiplication Rule for Several Events

In general, the probability of any sequence of independent events is simply the product of their corresponding probabilities.

Treating Dependent Events as Independent

Some calculations are cumbersome, but they can be made manageable by using the common practice of treating events as independent when small samples are drawn from large populations. In such cases, it is rare to select the same item twice.

The 5% Guideline for Cumbersome Calculations

If a sample size is no more than 5% of the size of the population, treat the selections as being independent (even if the selections are made without replacement, so they are technically dependent).

Principle of Redundancy

One design feature contributing to reliability is the use of redundancy, whereby critical components are duplicated so that if one fails, the other will work. For example, single-engine aircraft now have two independent electrical systems so that if one electrical system fails, the other can continue to work so that the engine does not fail.

Summary of Fundamentals

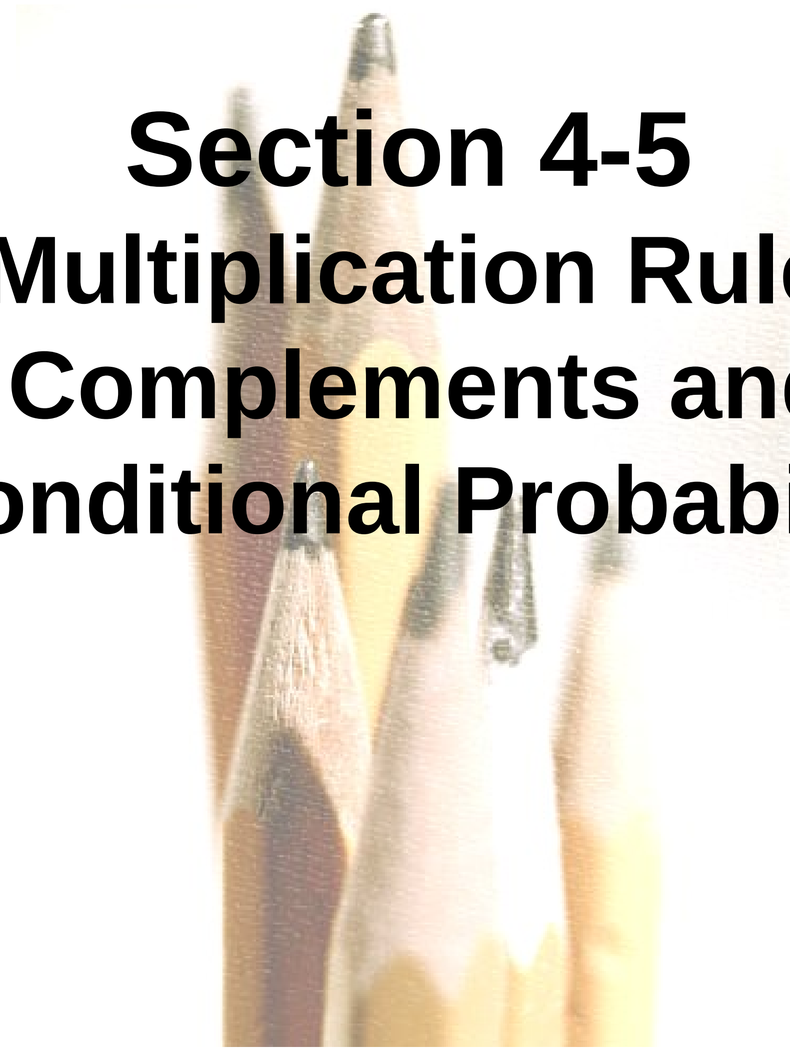
- **In the addition rule, the word "or" in *P*(*A* or *B*) suggests addition. Add *P*(*A*) and *P*(*B*), being careful to add in such a way that every outcome is counted only once.**
- **In the multiplication rule, the word "and" in *P*(*A* and *B*) suggests multiplication. Multiply *P*(*A*) and *P*(*B*), but be sure that the probability of event *B* takes into account the previous occurrence of event *A*.**

Recap

In this section we have discussed:

- **Notation for *P*(*A* and *B*).**
- **Tree diagrams.**
- **Notation for conditional probability.**
- **Independent events.**
- **Formal and intuitive multiplication rules.**

Section 4-5
Multiplication Rule: Complements and Conditional Probability



Key Concepts

Probability of "at least one":
Find the probability that among several trials, we get at least one of some specified event.

Conditional probability:
Find the probability of an event when we have additional information that some other event has already occurred.

Complements: The Probability of "At Least One"

- **"At least one" is equivalent to "one or more."**

- **The complement of getting at least one item of a particular type is that you get no items of that type.**

Finding the Probability of "At Least One"

To find the probability of at least one of something, calculate the probability of none, then subtract that result from 1. That is,

$$P(\text{at least one}) = 1 - P(\text{none}).$$

Conditional Probability

A conditional probability of an event is a probability obtained with the additional information that some other event has already occurred. $P(B|A)$ denotes the conditional probability of event B occurring, given that event A has already occurred, and it can be found by dividing the probability of events A and B both occurring by the probability of event A:

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$$

Intuitive Approach to Conditional Probability

The conditional probability of *B* given *A* can be found by assuming that event *A* has occurred, and then calculating the probability that event *B* will occur.

Confusion of the Inverse

To incorrectly believe that $P(A|B)$ and $P(B|A)$ are the same, or to incorrectly use one value for the other, is often called confusion of the inverse.

Recap

In this section we have discussed:

- **Concept of “at least one.”**
- **Conditional probability.**
- **Intuitive approach to conditional probability.**

Section 4-6
Probabilities Through Simulations

Key Concept

In this section we use simulations as an alternative approach to finding probabilities. The advantage to using simulations is that we can overcome much of the difficulty encountered when using the formal rules discussed in the preceding sections.

Simulation

A simulation of a procedure is a process that behaves the same way as the procedure, so that similar results are produced.

Simulation Example

Gender Selection In a test of the MicroSort method of gender selection developed by the Genetics & IVF Institute, 127 boys were born among 152 babies born to parents who used the YSORT method for trying to have a baby boy. In order to properly evaluate these results, we need to know the probability of getting at least 127 boys among 152 births, assuming that boys and girls are equally likely. Assuming that male and female births are equally likely, describe a simulation that results in the genders of 152 newborn babies.

Solution

One approach is simply to flip a fair coin 152 times, with heads representing females and tails representing males. Another approach is to use a calculator or computer to randomly generate 152 numbers that are 0s and 1s, with 0 representing a male and 1 representing a female. The numbers must be generated in such a way that they are equally likely. Here are typical results:

Simulation Examples

Solution 1:

- Flipping a fair coin 100 times where heads = female
  tails = male

H H T H T T H H H H

female female male female male male male female female female

- Generating 0's and 1's with a computer or calculator where

Solution 2:

0 = male
1 = female

0 0 1 0 1 1 1 0 0 0

male male female male female female female male male male

Random Numbers

In many experiments, random numbers are used in the simulation of naturally occurring events. Below are some ways to generate random numbers.

- **A table of random of digits**
- **STATDISK**
- **Minitab**
- **Excel**
- **TI-83/84 Plus calculator**

Random Numbers

STATDISK

Row	1 Ran...
1	7
2	8
3	16
4	38
5	42
6	46
7	68
8	68
9	104
10	117
11	140
12	195
13	204
14	244
15	271
16	274

Minitab

↓	C1	C2
1	38	
2	48	
3	59	
4	71	
5	101	
6	107	
7	122	
8	129	
9	153	
10	153	
11	163	

Random Numbers

Excel

	A
1	15
2	3
3	15
4	362
5	164
6	184
7	158
8	59
9	143
10	85
11	134

TI-83/84 Plus calculator

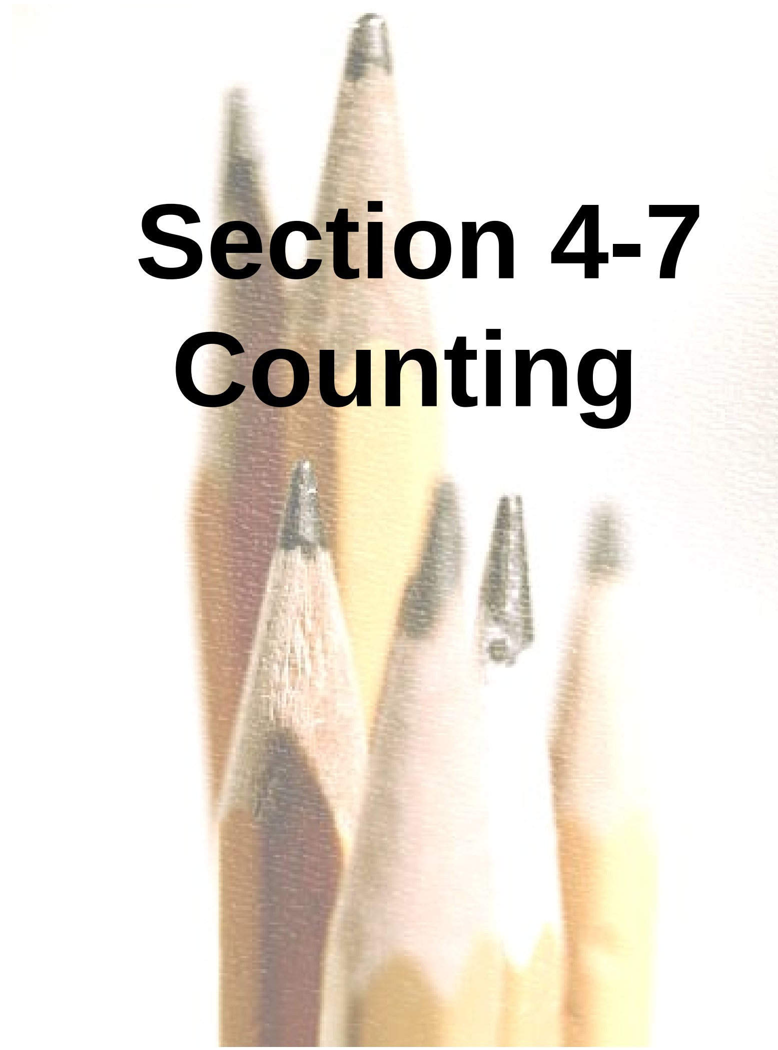
```
randInt(1,365,25
→L1
{79 206 340 133…
SortA(L1)
                Done
L1
{17 34 46 70 79…
```

Recap

In this section we have discussed:

- **The definition of a simulation.**
- **How to create a simulation.**
- **Ways to generate random numbers.**

Section 4-7
Counting



Key Concept

In many probability problems, the big obstacle is finding the total number of outcomes, and this section presents several methods for finding such numbers without directly listing and counting the possibilities.

Fundamental Counting Rule

For a sequence of two events in which the first event can occur *m* ways and the second event can occur *n* ways, the events together can occur a total of $m \cdot n$ ways.

Notation

The factorial symbol ! denotes the product of decreasing positive whole numbers.

For example,

$$4! = 4 \bullet 3 \bullet 2 \bullet 1 = 24.$$

By special definition, 0! = 1.

Factorial Rule

A collection of n different items can be arranged in order $n!$ different ways. (This **factorial rule** reflects the fact that the first item may be selected in n different ways, the second item may be selected in $n - 1$ ways, and so on.)

Permutations Rule
(when items are all different)

Requirements:

1. There are n different items available. (This rule does not apply if some of the items are identical to others.)
2. We select r of the n items (without replacement).
3. We consider rearrangements of the same items to be different sequences. (The permutation of ABC is different from CBA and is counted separately.)

If the preceding requirements are satisfied, the number of permutations (or sequences) of r items selected from n available items (without replacement) is

$$_nP_r = \frac{n!}{(n-r)!}$$

Permutations Rule
(when some items are identical to others)

Requirements:

1. **There are n items available, and some items are identical to others.**
2. **We select all of the n items (without replacement).**
3. **We consider rearrangements of distinct items to be different sequences.**

If the preceding requirements are satisfied, and if there are n_1 alike, n_2 alike, . . . n_k alike, the number of permutations (or sequences) of all items selected without replacement is

$$\frac{n!}{n_1! \cdot n_2! \cdots \cdots n_k!}$$

Combinations Rule

Requirements:

1. **There are *n* different items available.**
2. **We select *r* of the *n* items (without replacement).**
3. **We consider rearrangements of the same items to be the same. (The combination of ABC is the same as CBA.)**

If the preceding requirements are satisfied, the number of combinations of *r* items selected from *n* different items is

$$_{n}C_{r} = \frac{n!}{(n-r)!\,r!}$$

Permutations versus Combinations

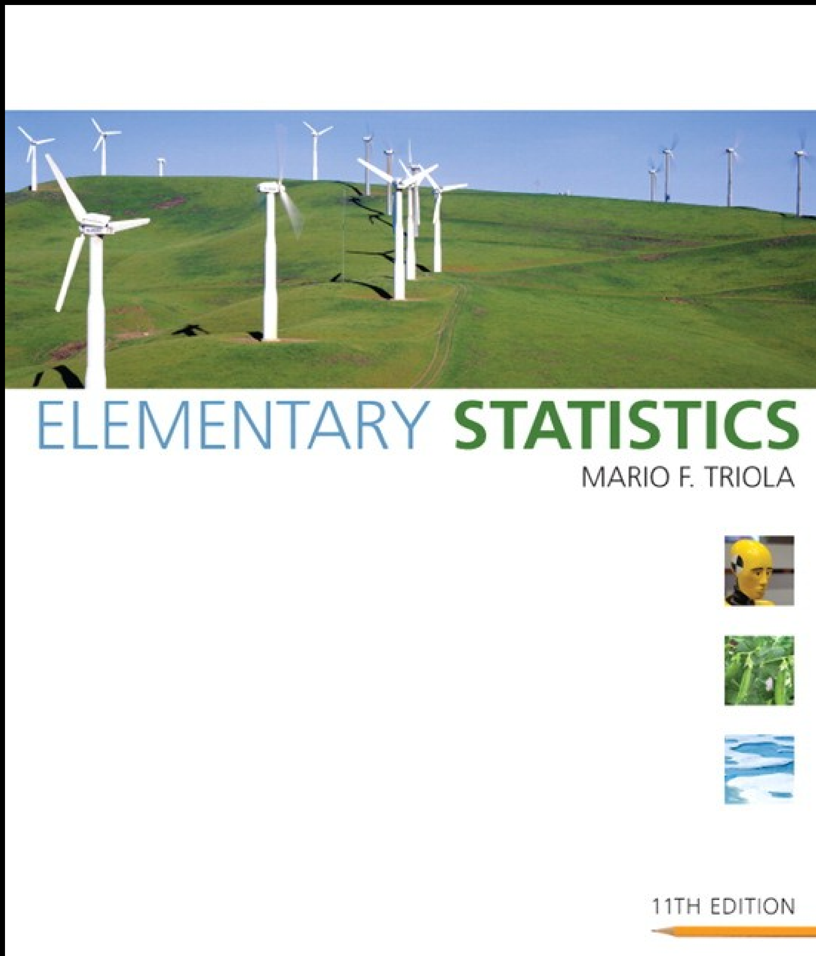
When different orderings of the same items are to be counted separately, we have a permutation problem, but when different orderings are not to be counted separately, we have a combination problem.

Recap

In this section we have discussed:

- **The fundamental counting rule.**
- **The factorial rule.**
- **The permutations rule (when items are all different).**
- **The permutations rule (when some items are identical to others).**
- **The combinations rule.**

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 5
Probability Distributions

5-1 Review and Preview

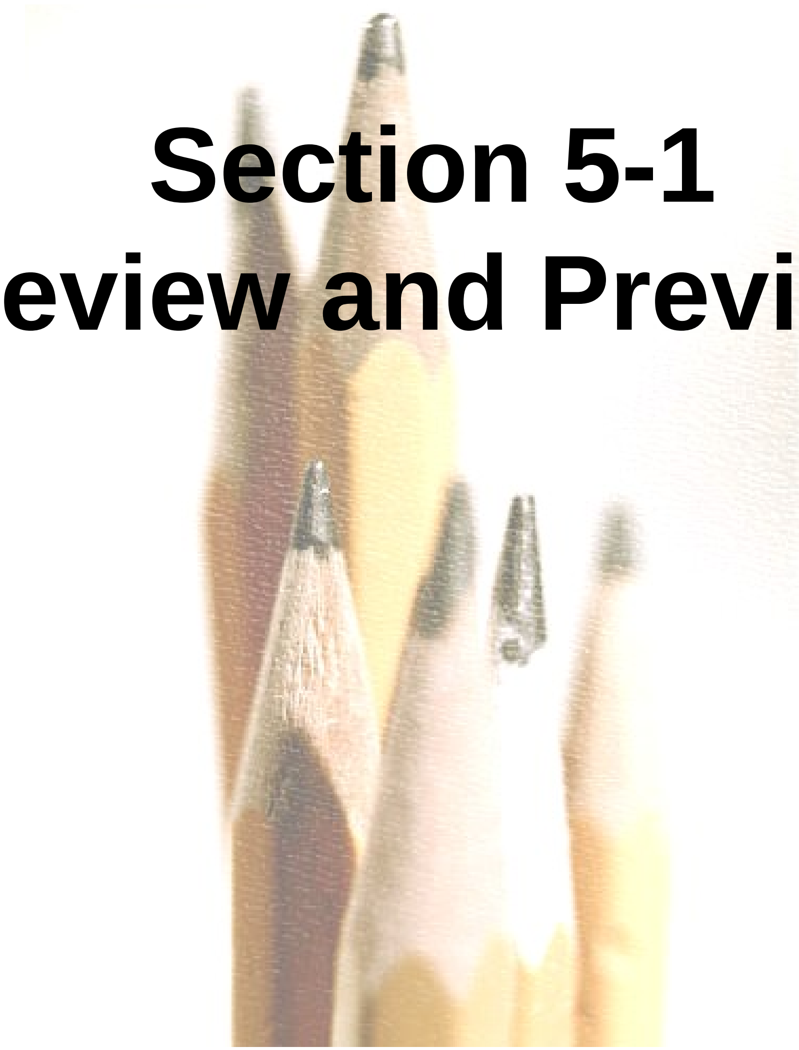
5-2 Random Variables

5-3 Binomial Probability Distributions

5-4 Mean, Variance and Standard Deviation for the Binomial Distribution

5-5 Poisson Probability Distributions

Section 5-1
Review and Preview



Review and Preview

This chapter combines the methods of **descriptive statistics** presented in Chapter 2 and 3 and those of **probability** presented in Chapter 4 to describe and analyze

probability distributions.

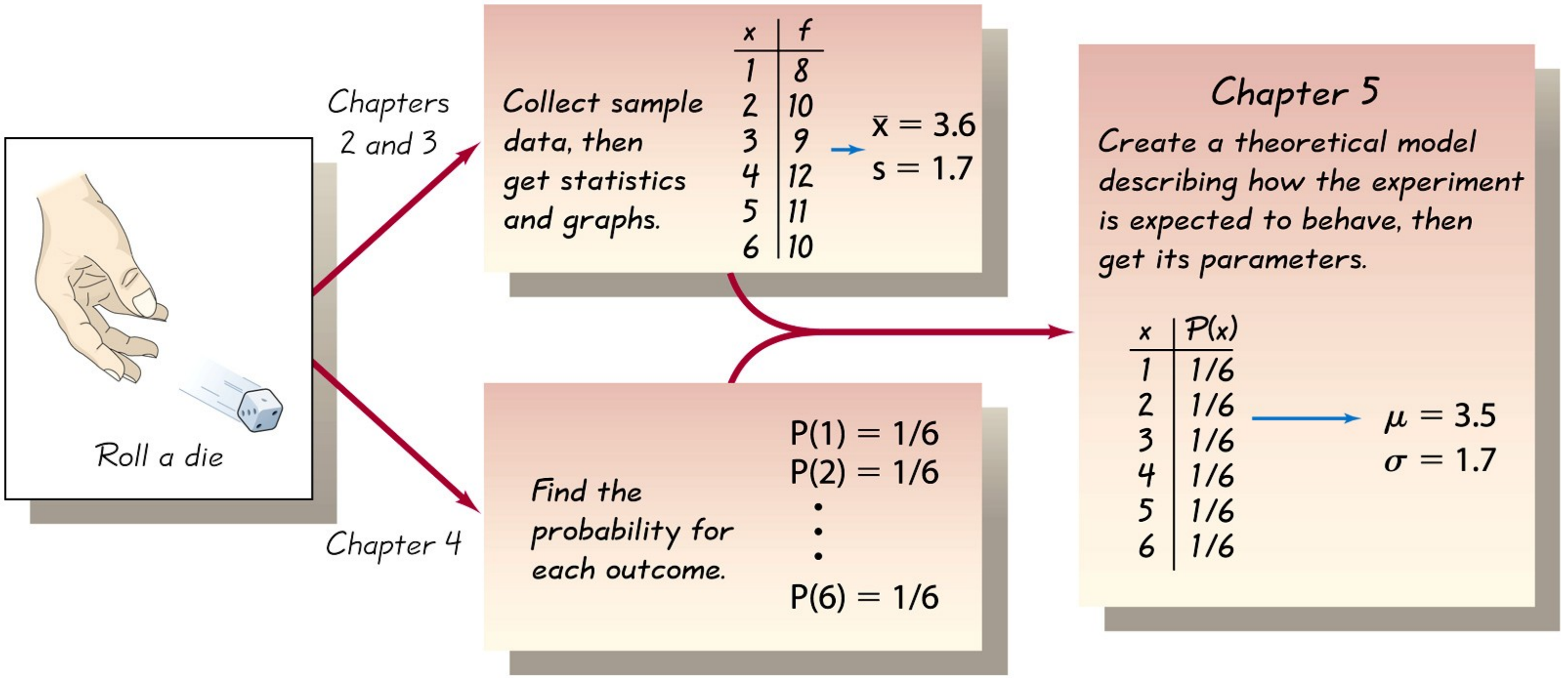
Probability Distributions describe what will **probably** happen instead of what actually **did** happen, and they are often given in the format of a graph, table, or formula.

Preview

In order to fully understand probability distributions, we must first understand the concept of a random variable, and be able to distinguish between discrete and continuous random variables. In this chapter we focus on discrete probability distributions. In particular, we discuss binomial and Poisson probability distributions.

Combining Descriptive Methods and Probabilities

In this chapter we will construct probability distributions by presenting possible outcomes along with the relative frequencies we expect.



Section 5-2
Random Variables

Key Concept

This section introduces the important concept of a probability distribution, which gives the probability for each value of a variable that is determined by chance.

Give consideration to distinguishing between outcomes that are likely to occur by chance and outcomes that are "unusual" in the sense they are not likely to occur by chance.

Key Concept

- The concept of random variables and how they relate to probability distributions
- Distinguish between discrete random variables and continuous random variables
- Develop formulas for finding the mean, variance, and standard deviation for a probability distribution
- Determine whether outcomes are likely to occur by chance or they are unusual (in the sense that they are not likely to occur by chance)

Random Variable Probability Distribution

- **Random variable**
  a variable (typically represented by x) that has a single numerical value, determined by chance, for each outcome of a procedure
- **Probability distribution**
  a description that gives the probability for each value of the random variable; often expressed in the format of a graph, table, or formula

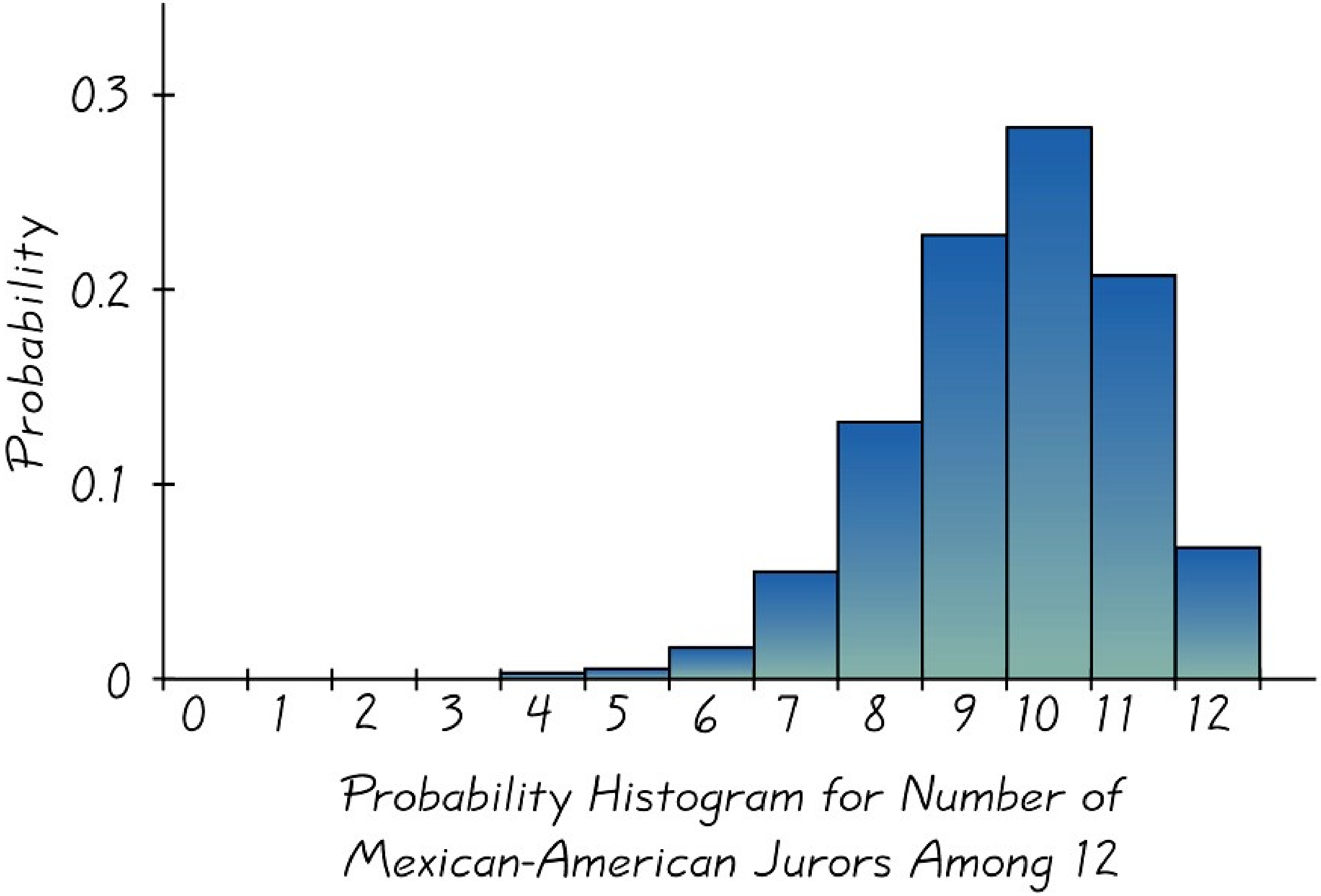
Discrete and Continuous Random Variables

- **Discrete random variable**
  either a finite number of values or countable number of values, where “countable” refers to the fact that there might be infinitely many values, but they result from a counting process

- **Continuous random variable**
  infinitely many values, and those values can be associated with measurements on a continuous scale without gaps or interruptions

Graphs

The probability histogram is very similar to a relative frequency histogram, but the vertical scale shows probabilities.



Requirements for Probability Distribution

$$\sum P(x) = 1$$

where x assumes all possible values.

$$0 \le P(x) \le 1$$

for every individual value of x.

Mean, Variance and Standard Deviation of a Probability Distribution

$$\mu = \Sigma [x \cdot P(x)]$$ **Mean**

$$\sigma^2 = \Sigma [(x - \mu)^2 \cdot P(x)]$$ **Variance**

$$\sigma^2 = \Sigma [x^2 \cdot P(x)] - \mu^2$$ **Variance (shortcut)**

$$\sigma = \sqrt{\Sigma [x^2 \cdot P(x)] - \mu^2}$$ **Standard Deviation**

Roundoff Rule for μ, σ, and σ^2

Round results by carrying one more decimal place than the number of decimal places used for the random variable x.

If the values of x are integers, round μ, σ, and σ^2 to one decimal place.

Identifying *Unusual* Results
Range Rule of Thumb

According to the range rule of thumb, most values should lie within 2 standard deviations of the mean.

We can therefore identify "unusual" values by determining if they lie outside these limits:

Maximum usual value = $\mu + 2\sigma$

Minimum usual value = $\mu - 2\sigma$

Identifying Unusual Results Probabilities

Rare Event Rule for Inferential Statistics

If, under a given assumption (such as the assumption that a coin is fair), the probability of a particular observed event (such as 992 heads in 1000 tosses of a coin) is extremely small, we conclude that the assumption is probably not correct.

Identifying Unusual Results Probabilities

Using Probabilities to Determine When Results Are Unusual

- **Unusually high**: x successes among n trials is an **unusually high** number of successes if $P(x \text{ or more}) \leq 0.05$.
- **Unusually low**: x successes among n trials is an **unusually low** number of successes if $P(x \text{ or fewer}) \leq 0.05$.

Expected Value

The **expected value** of a discrete random variable is denoted by E, and it represents the mean value of the outcomes. It is obtained by finding the value of $\Sigma [x \cdot P(x)]$.

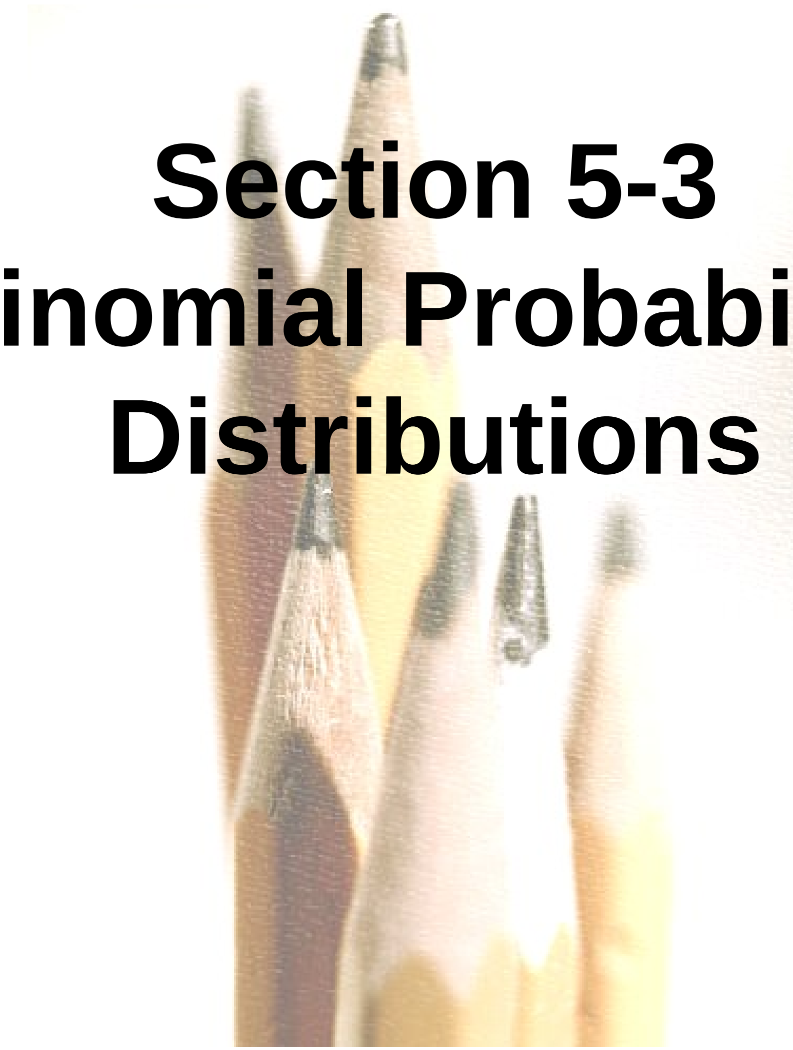
$$E = \Sigma [x \cdot P(x)]$$

Recap

In this section we have discussed:

- **Combining methods of descriptive statistics with probability.**
- **Random variables and probability distributions.**
- **Probability histograms.**
- **Requirements for a probability distribution.**
- **Mean, variance and standard deviation of a probability distribution.**
- **Identifying unusual results.**
- **Expected value.**

Section 5-3
Binomial Probability Distributions



Key Concept

This section presents a basic definition of a binomial distribution along with notation, and methods for finding probability values.

Binomial probability distributions allow us to deal with circumstances in which the outcomes belong to **two** relevant categories such as acceptable/defective or survived/died.

Binomial Probability Distribution

A **binomial probability distribution** results from a procedure that meets all the following requirements:

1. The procedure has a **fixed number of trials**.
2. The trials must be **independent**. (The outcome of any individual trial doesn't affect the probabilities in the other trials.)
3. Each trial must have all outcomes classified into **two categories** (commonly referred to as **success** and **failure**).
4. The probability of a success remains the same in all trials.

Notation for Binomial Probability Distributions

S and **F** (success and failure) denote the two possible categories of all outcomes; p and q will denote the probabilities of **S** and **F**, respectively, so

$P(S) = p$ (p = probability of success)

$P(F) = 1 - p = q$ (q = probability of failure)

Notation (continued)

n denotes the fixed number of trials.

x denotes a specific number of successes in ***n*** trials, so ***x*** can be any whole number between **0** and ***n***, inclusive.

p denotes the probability of **success** in *one* of the *n* trials.

q denotes the probability of **failure** in **one** of the ***n*** trials.

P*(*x*)** denotes the probability of getting exactly ***x successes among the ***n*** trials.

Important Hints

- Be sure that x and p both refer to the **same** category being called a success.

- When sampling without replacement, consider events to be independent if $n < 0.05N$.

Methods for Finding Probabilities

We will now discuss three methods for finding the probabilities corresponding to the random variable *x* in a binomial distribution.

Method 1: Using the Binomial Probability Formula

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x}$$

for $x = 0, 1, 2, \ldots, n$

where

n = number of trials

x = number of successes among n trials

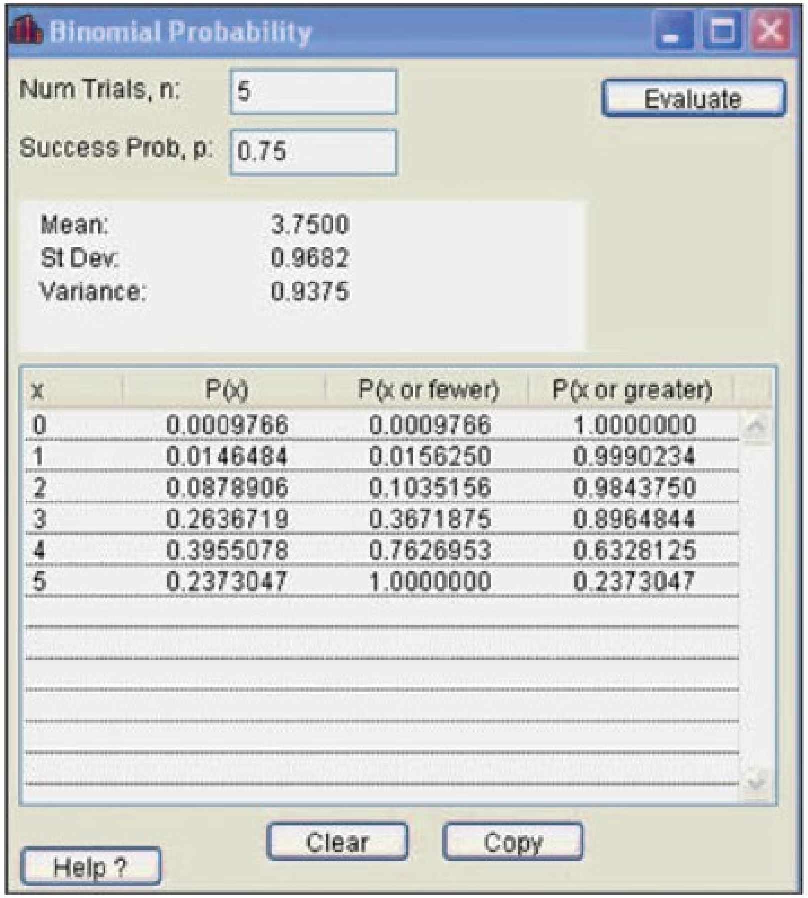
p = probability of success in any one trial

q = probability of failure in any one trial ($q = 1 - p$)

Method 2: Using Technology

STATDISK, Minitab, Excel, SPSS, SAS and the TI-83/84 Plus calculator can be used to find binomial probabilities.

STATDISK



x	P(x)	P(x or fewer)	P(x or greater)
0	0.0009766	0.0009766	1.0000000
1	0.0146484	0.0156250	0.9990234
2	0.0878906	0.1035156	0.9843750
3	0.2636719	0.3671875	0.8964844
4	0.3955078	0.7626953	0.6328125
5	0.2373047	1.0000000	0.2373047

MINITAB

x	P(x)
0	0.000977
1	0.014648
2	0.087891
3	0.263672
4	0.395508
5	0.237305

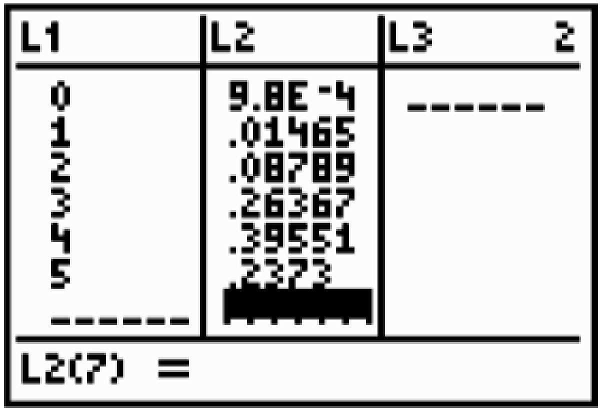
Method 2: Using Technology

STATDISK, Minitab, Excel and the TI-83 Plus calculator can all be used to find binomial probabilities.

EXCEL

	A	B
1	0	0.000977
2	1	0.014648
3	2	0.087891
4	3	0.263672
5	4	0.395508
6	5	0.237305

TI-83 PLUS Calculator



Method 3: Using Table A-1 in Appendix A

Part of Table A-1 is shown below. With *n* = 12 and *p* = 0.80 in the binomial distribution, the probabilities of 4, 5, 6, and 7 successes are 0.001, 0.003, 0.016, and 0.053 respectively.

n	x	p 0.80
	4	0.001
	5	0.003
	6	0.016
	7	0.053

→

x	p
4	0.001
5	0.003
6	0.016
7	0.053

Strategy for Finding Binomial Probabilities

- Use computer software or a TI-83 Plus calculator if available.
- If neither software nor the TI-83 Plus calculator is available, use Table A-1, if possible.
- If neither software nor the TI-83 Plus calculator is available and the probabilities can’t be found using Table A-1, use the binomial probability formula.

Rationale for the Binomial Probability Formula

$$P(x) = \underbrace{\frac{n!}{(n-x)!\,x!}}_{\substack{\text{The number of outcomes with} \\ \text{exactly } x \text{ successes} \\ \text{among } n \text{ trials}}} \cdot \; p^x \cdot q^{n-x}$$

Binomial Probability Formula

$$P(x) = \underbrace{\frac{n\,!}{(n-x)!\,x!}}_{\substack{\text{Number of outcomes with}\\ \text{exactly } x \text{ successes}\\ \text{among } n \text{ trials}}} \cdot \underbrace{p^x \cdot q^{n-x}}_{\substack{\text{The probability of } x \text{ successes}\\ \text{among } n \text{ trials for any one}\\ \text{particular order}}}$$

Number of outcomes with exactly *x* successes among *n* trials

The probability of *x* successes among *n* trials for any one particular order

Recap

In this section we have discussed:

- **The definition of the binomial probability distribution.**
- **Notation.**
- **Important hints.**
- **Three computational methods.**
- **Rationale for the formula.**

Section 5-4
Mean, Variance, and Standard Deviation for the Binomial Distribution

Key Concept

In this section we consider important characteristics of a binomial distribution including center, variation and distribution. That is, given a particular binomial probability distribution we can find its mean, variance and standard deviation.

A strong emphasis is placed on interpreting and understanding those values.

For Any Discrete Probability Distribution: Formulas

Mean $\mu = \Sigma[x \cdot P(x)]$

Variance $\sigma^2 = [\Sigma x^2 \cdot P(x)] - \mu^2$

Std. Dev $\sigma = \sqrt{[\Sigma x^2 \cdot P(x)] - \mu^2}$

Binomial Distribution: Formulas

Mean $\mu = n \cdot p$

Variance $\sigma^2 = n \cdot p \cdot q$

Std. Dev. $\sigma = \sqrt{n \cdot p \cdot q}$

Where

n **= number of fixed trials**

p **= probability of success in one of the *n* trials**

q **= probability of failure in one of the *n* trials**

Interpretation of Results

It is especially important to interpret results. The range rule of thumb suggests that values are unusual if they lie outside of these limits:

$$\text{Maximum usual values} = \mu + 2\sigma$$

$$\text{Minimum usual values} = \mu - 2\sigma$$

Recap

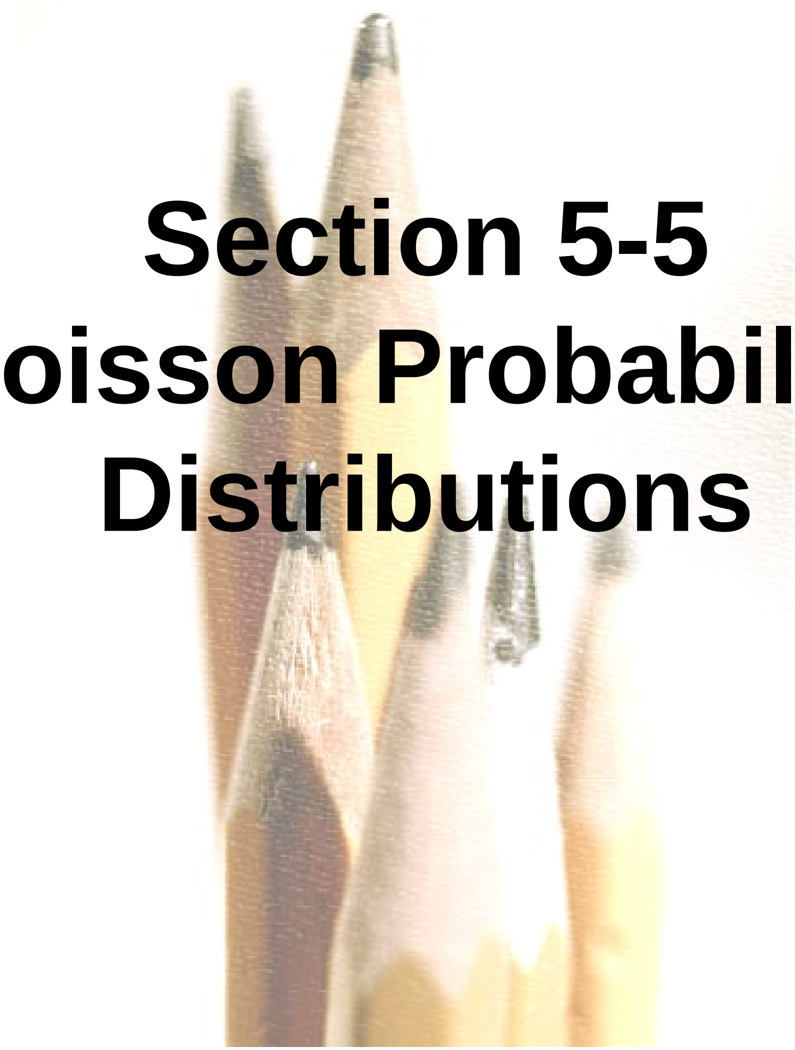
In this section we have discussed:

- **Mean, variance and standard deviation formulas for any discrete probability distribution.**

- **Mean, variance and standard deviation formulas for the binomial probability distribution.**

- **Interpreting results.**

Section 5-5
Poisson Probability Distributions



Key Concept

The Poisson distribution is another discrete probability distribution which is important because it is often used for describing the behavior of rare events (with small probabilities).

Poisson Distribution

The Poisson distribution is a discrete probability distribution that applies to occurrences of some event over a specified interval. The random variable *x* is the number of occurrences of the event in an interval. The interval can be time, distance, area, volume, or some similar unit.

Formula

$$P(x) = \frac{\mu^{x} \cdot e^{-\mu}}{x!} \quad \text{where } e \approx 2.71828$$

Requirements of the Poisson Distribution

- The random variable x is the number of occurrences of an event over some interval.
- The occurrences must be random.
- The occurrences must be independent of each other.
- The occurrences must be uniformly distributed over the interval being used.

Parameters

- The mean is μ.
- The standard deviation is $\sigma = \sqrt{\mu}$.

Parameters of the Poisson Distribution

- **The mean is μ.**
- **The standard deviation is $\sigma = \sqrt{\mu}$.**

Difference from a Binomial Distribution

The Poisson distribution differs from the binomial distribution in these fundamental ways:

- **The binomial distribution is affected by the sample size n and the probability p, whereas the Poisson distribution is affected only by the mean μ.**
- **In a binomial distribution the possible values of the random variable x are 0, 1, . . . n, but a Poisson distribution has possible x values of 0, 1, 2, . . . , with no upper limit.**

Poisson as an Approximation to the Binomial Distribution

The Poisson distribution is sometimes used to approximate the binomial distribution when n is large and p is small.

Rule of Thumb

- $n \geq 100$
- $np \leq 10$

Poisson as an Approximation to the Binomial Distribution - μ

If both of the following requirements are met,

- $n \geq 100$
- $np \leq 10$

then use the following formula to calculate μ,

Value for μ

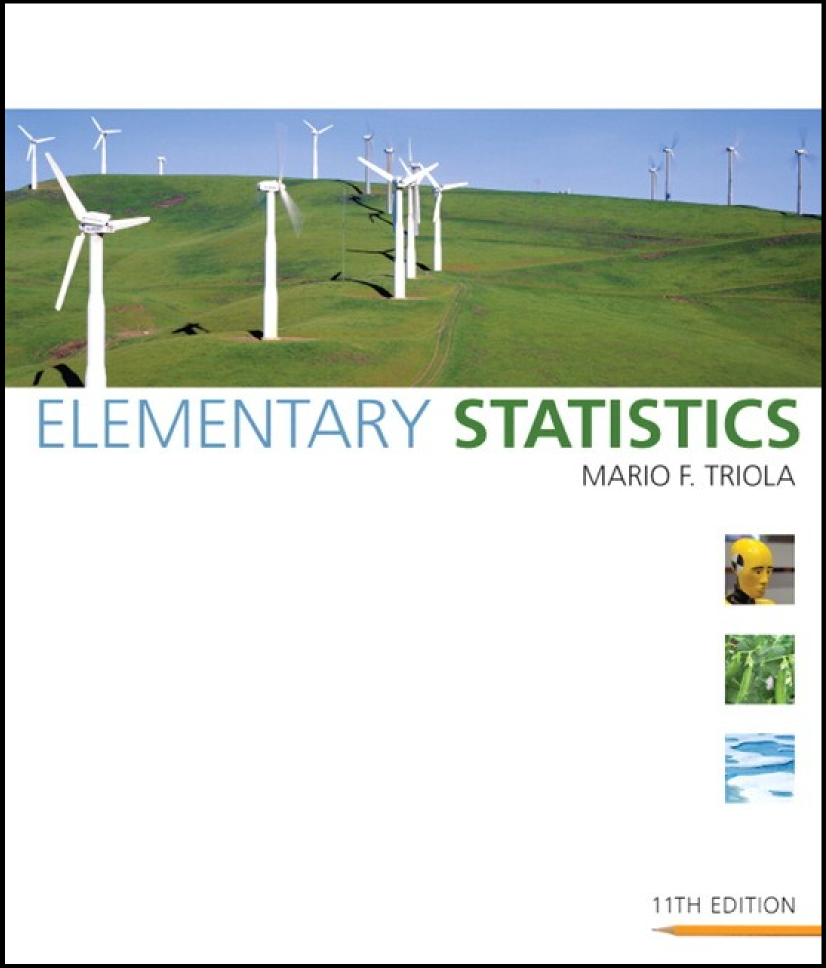
$$\mu = n \cdot p$$

Recap

In this section we have discussed:

- **Definition of the Poisson distribution.**
- **Requirements for the Poisson distribution.**
- **Difference between a Poisson distribution and a binomial distribution.**
- **Poisson approximation to the binomial.**

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

Chapter 6
Normal Probability Distributions

6-1 Review and Preview

6-2 The Standard Normal Distribution

6-3 Applications of Normal Distributions

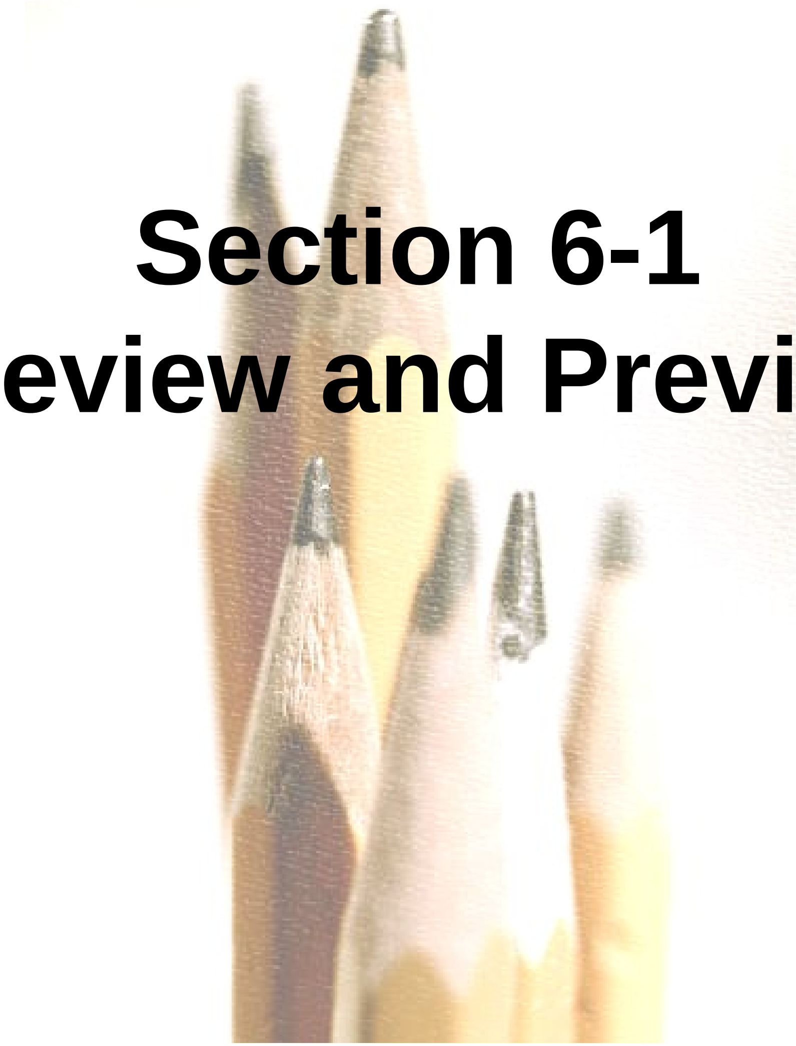
6-4 Sampling Distributions and Estimators

6-5 The Central Limit Theorem

6-6 Normal as Approximation to Binomial

6-7 Assessing Normality

Section 6-1
Review and Preview



Review

- **Chapter 2: Distribution of data**
- **Chapter 3: Measures of data sets, including measures of center and variation**
- **Chapter 4: Principles of probability**
- **Chapter 5: Discrete probability distributions**

Preview

Chapter focus is on:

- **Continuous random variables**
- **Normal distributions**

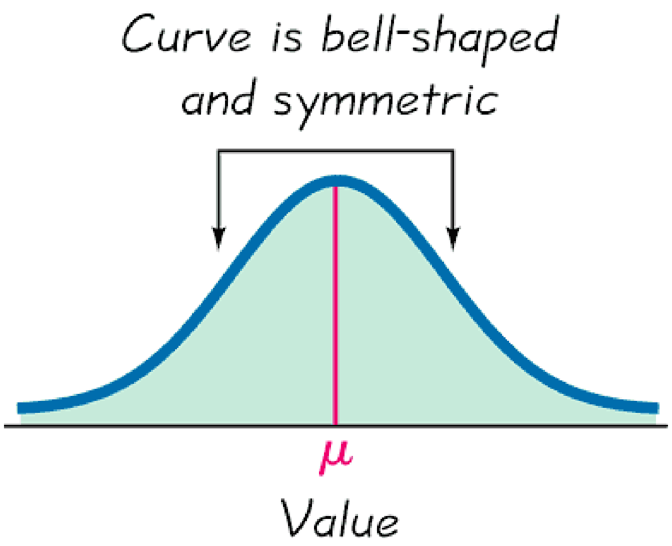


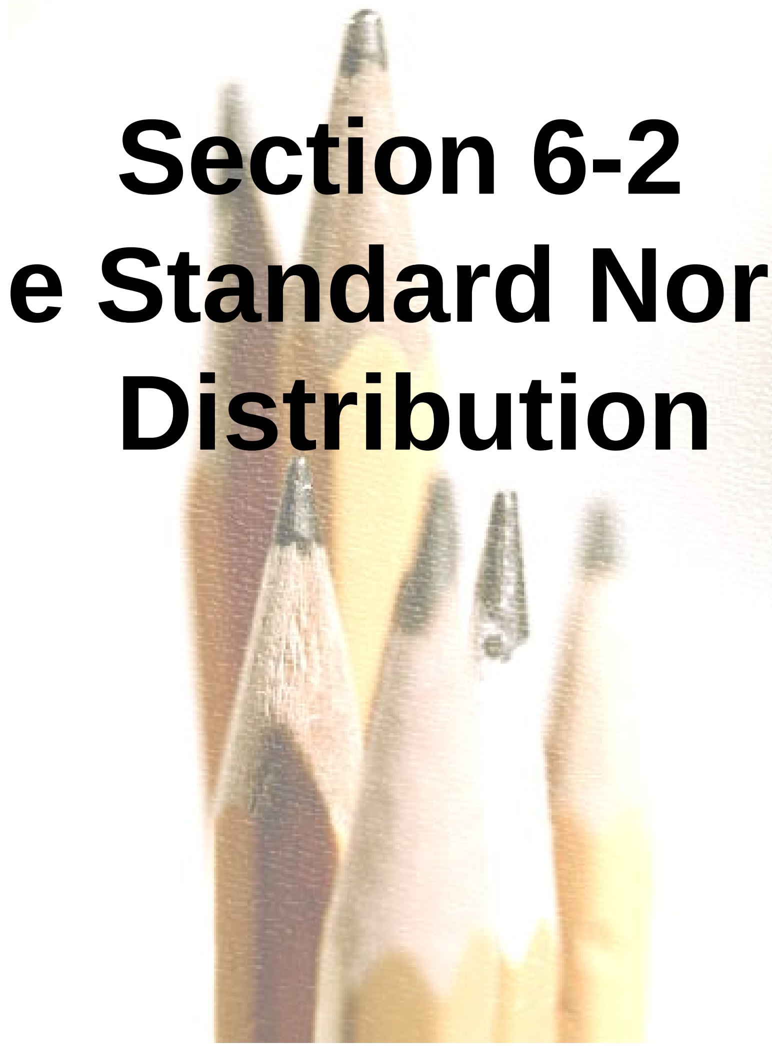
Figure 6-1

$$f(x)=\frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Formula 6-1

Distribution determined by fixed values of mean and standard deviation

Section 6-2
The Standard Normal Distribution



Key Concept

This section presents the *standard normal distribution* which has three properties:

1. It's graph is bell-shaped.

2. It's mean is equal to 0 ($\mu = 0$).

3. It's standard deviation is equal to 1 ($\sigma = 1$).

Develop the skill to find areas (or probabilities or relative frequencies) corresponding to various regions under the graph of the standard normal distribution. Find *z*-scores that correspond to area under the graph.

Uniform Distribution

A continuous random variable has a uniform distribution if its values are spread evenly over the range of probabilities. The graph of a uniform distribution results in a rectangular shape.

Density Curve

A density curve is the graph of a continuous probability distribution. It must satisfy the following properties:

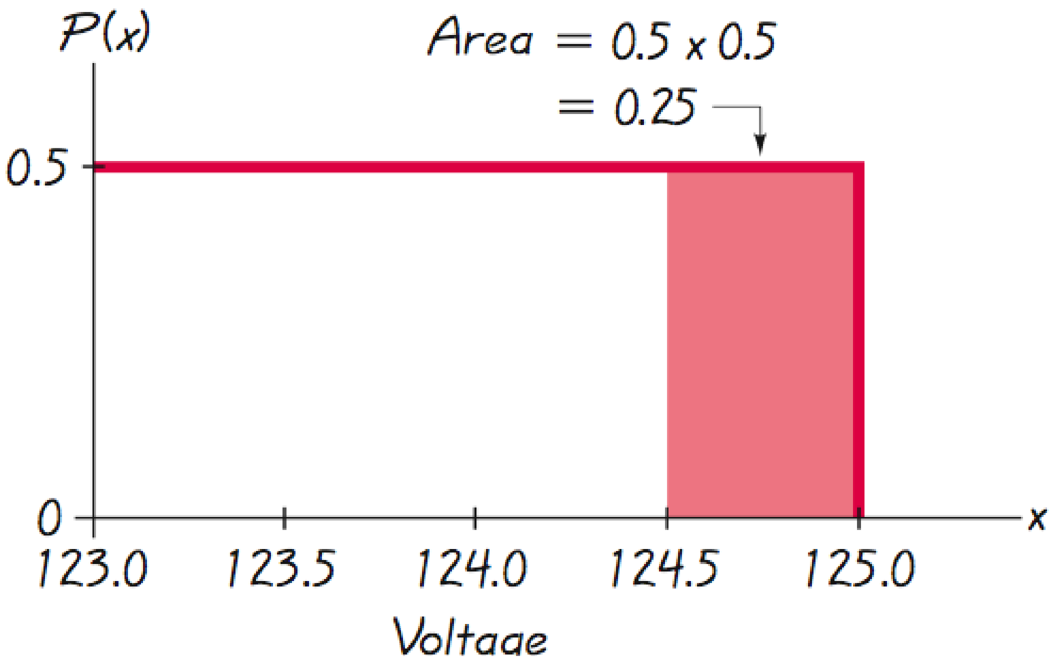
1. **The total area under the curve must equal 1.**
2. **Every point on the curve must have a vertical height that is 0 or greater. (That is, the curve cannot fall below the *x*-axis.)**

Area and Probability

Because the total area under the density curve is equal to 1, there is a correspondence between *area* and *probability*.

Using Area to Find Probability

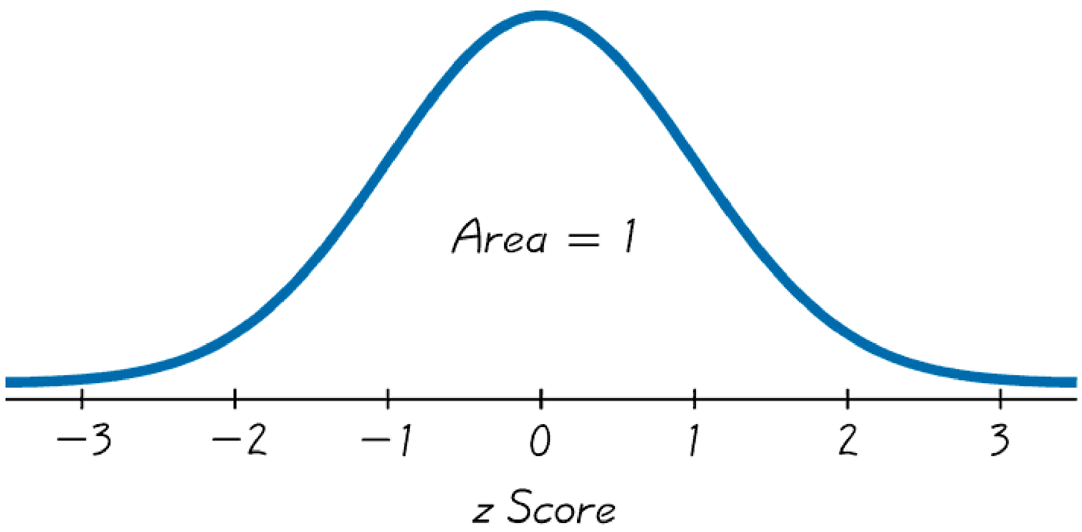
Given the uniform distribution illustrated, find the probability that a randomly selected voltage level is greater than 124.5 volts.



Shaded area represents voltage levels greater than 124.5 volts. Correspondence between area and probability: 0.25.

Standard Normal Distribution

The standard normal distribution is a normal probability distribution with $\mu = 0$ and $\sigma = 1$. The total area under its density curve is equal to 1.



Finding Probabilities When Given *z*-scores

- **Table A-2 (in Appendix A)**
- **Formulas and Tables insert card**
- **Find areas for many different regions**

Finding Probabilities – Other Methods

- STATDISK
- Minitab
- Excel
- TI-83/84 Plus

Methods for Finding Normal Distribution Areas

Table A-2, STATDISK, Minitab, Excel

Gives the cumulative area from the left up to a vertical line above a specific value of *z*.

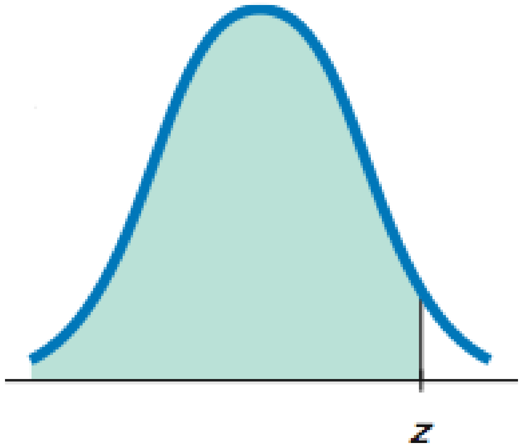


Table A-2 The procedure for using Table A-2 is described in the text.

STATDISK Select **Analysis, Probability Distributions, Normal Distribution.** Enter the *z* value, then click on **Evaluate.**

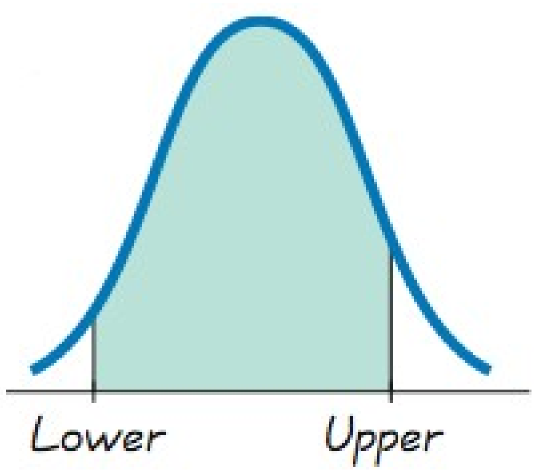
MINITAB Select **Calc, Probability Distributions, Normal.** In the dialog box, select **Cumulative Probability, Input Constant.**

EXCEL Select **fx, Statistical, NORMDIST.** In the dialog box, enter the value and mean, the standard deviation, and "true."

Methods for Finding Normal Distribution Areas

TI-83/84 Plus Calculator

Gives area bounded on the left and bounded on the right by vertical lines above any specific values.



TI-83/84 Press 2ND VARS **[2: normal cdf ( ],** then enter the two *z* scores separated by a comma, as in (left *z* score, right *z* score).

Table A-2

TABLE A-2 Standard Normal (z) Distribution: Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.50 and lower	.0001									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	* .0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	* .0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559

Using Table A-2

1. It is designed only for the *standard* normal distribution, which has a mean of 0 and a standard deviation of 1.

2. It is on two pages, with one page for *negative* *z*-scores and the other page for *positive* *z*-scores.

3. Each value in the body of the table is a *cumulative area from the left* up to a vertical boundary above a specific *z*-score.

Using Table A-2

4. **When working with a graph, avoid confusion between *z*-scores and areas.**

   ***z* Score**

   Distance along horizontal scale of the standard normal distribution; refer to the <u>leftmost column and top row</u> of Table A-2.

   Area

   Region under the curve; refer to the values in the <u>body</u> of Table A-2.

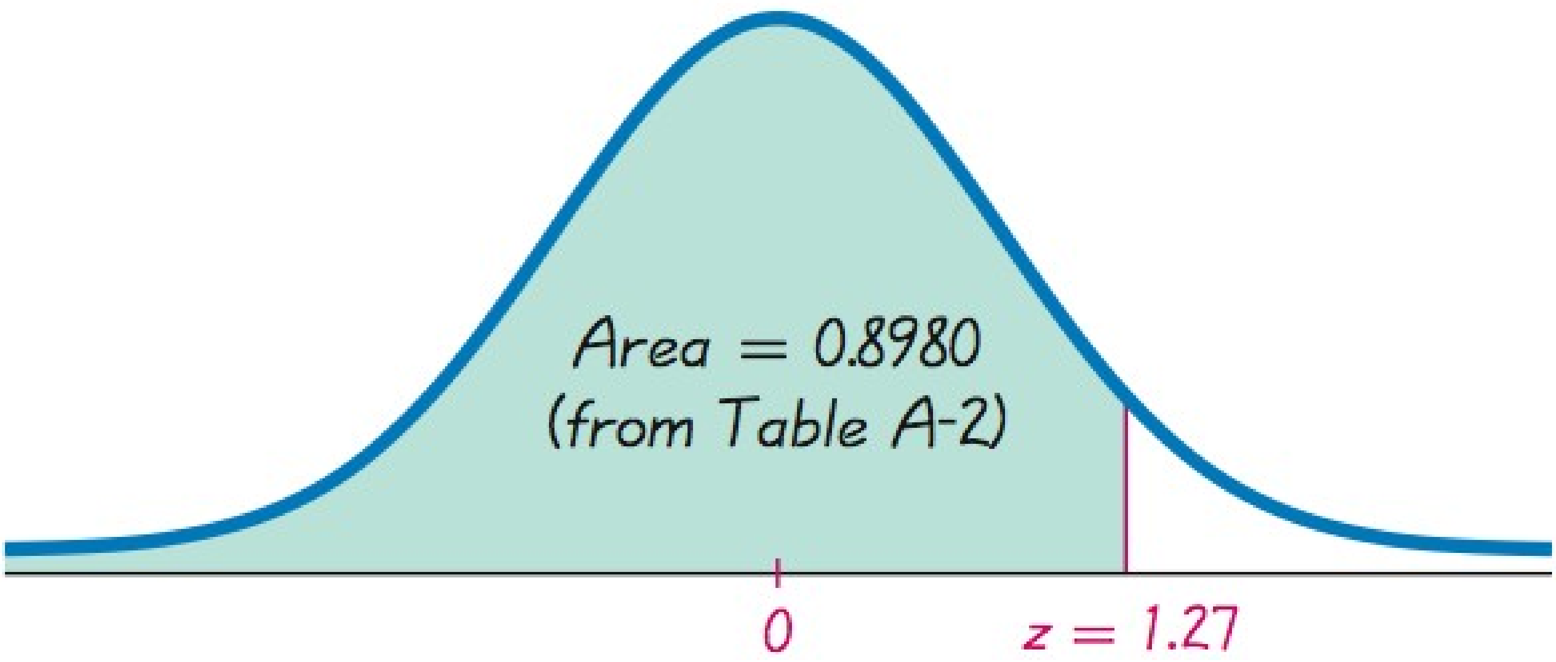
5. **The part of the z-score denoting hundredths is found across the top.**

Example - Thermometers

The Precision Scientific Instrument Company manufactures thermometers that are supposed to give readings of 0ºC at the freezing point of water. Tests on a large sample of these instruments reveal that at the freezing point of water, some thermometers give readings below 0º (denoted by negative numbers) and some give readings above 0º (denoted by positive numbers). Assume that the mean reading is 0ºC and the standard deviation of the readings is 1.00ºC. Also assume that the readings are normally distributed. If one thermometer is randomly selected, find the probability that, at the freezing point of water, the reading is less than 1.27º.

Example - (Continued)

$$P(z < 1.27) =$$

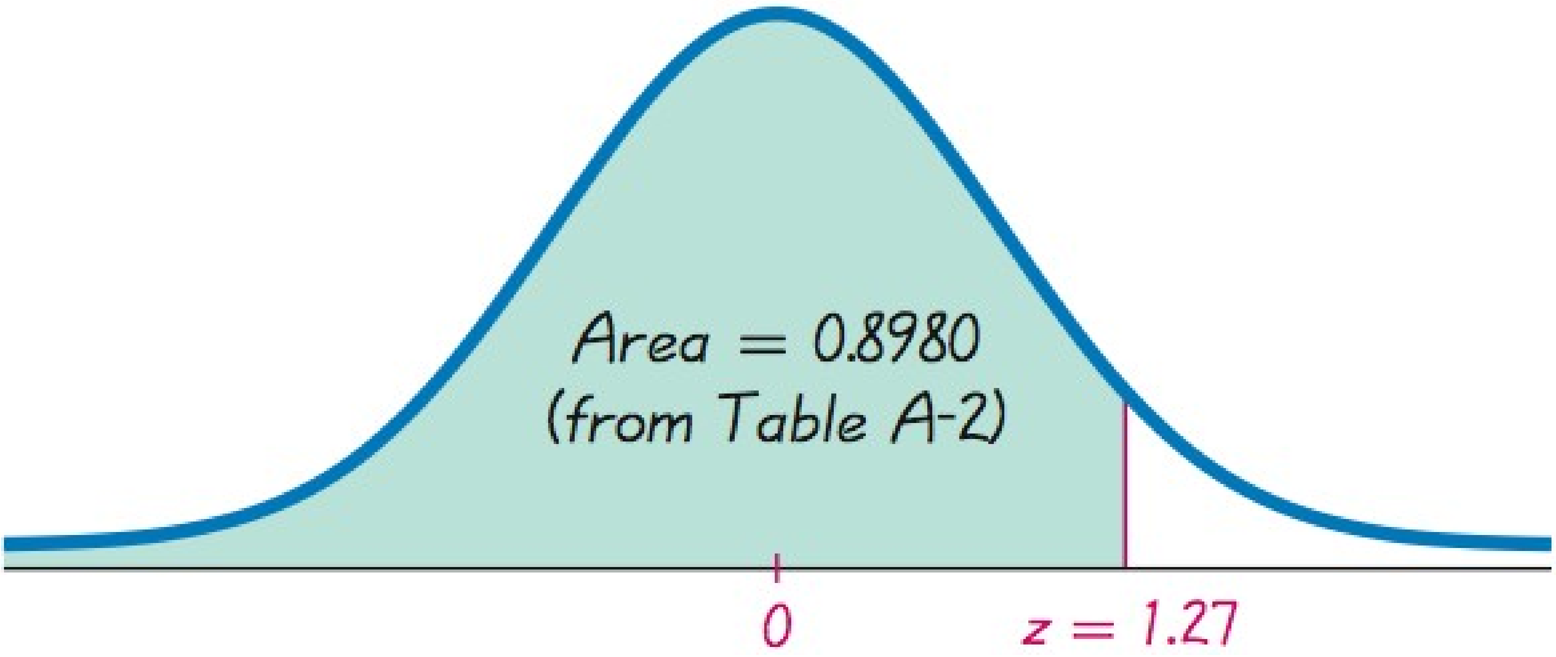


Look at Table A-2

TABLE A-2	(*continued*) Cumulative Area from the LEFT							
z	.00	.01	.02	.03	.04	.05	.06	.07
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292

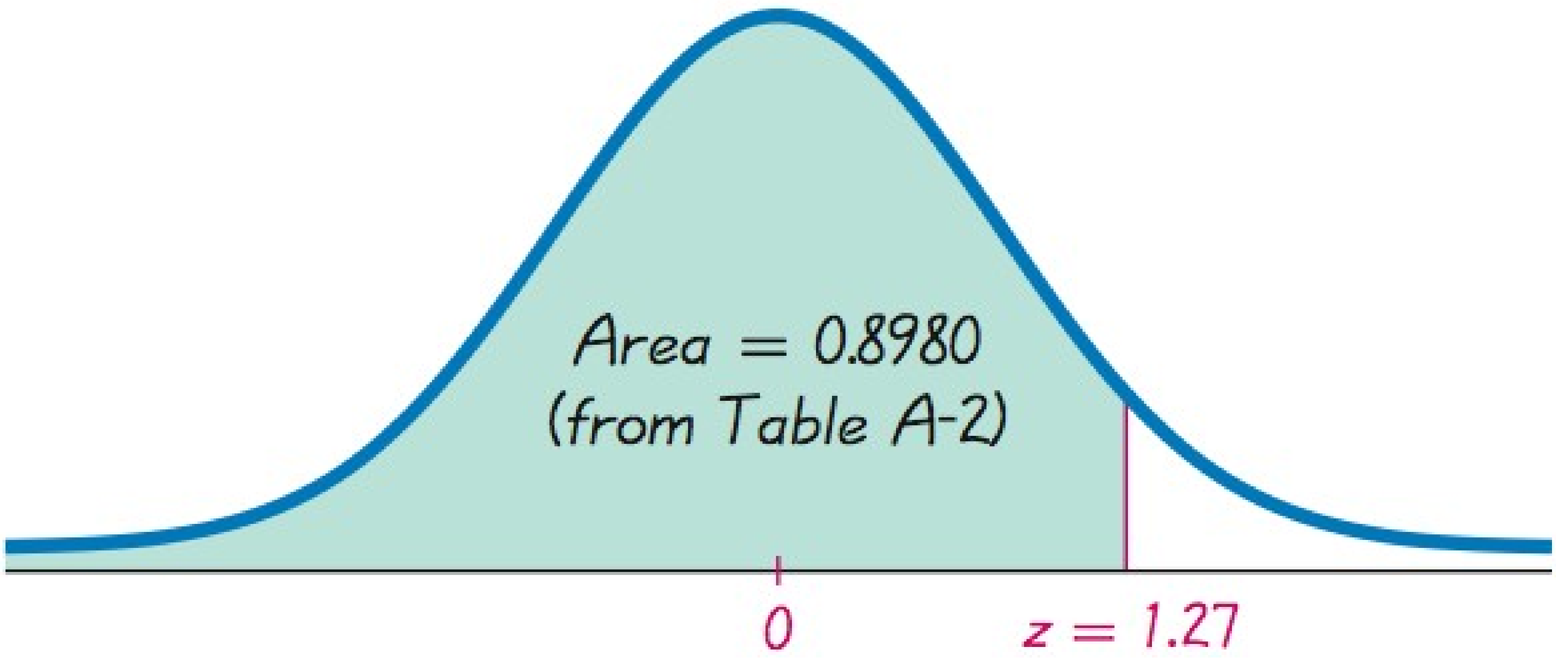
Example - cont

P (z < 1.27) = 0.8980



Example - cont

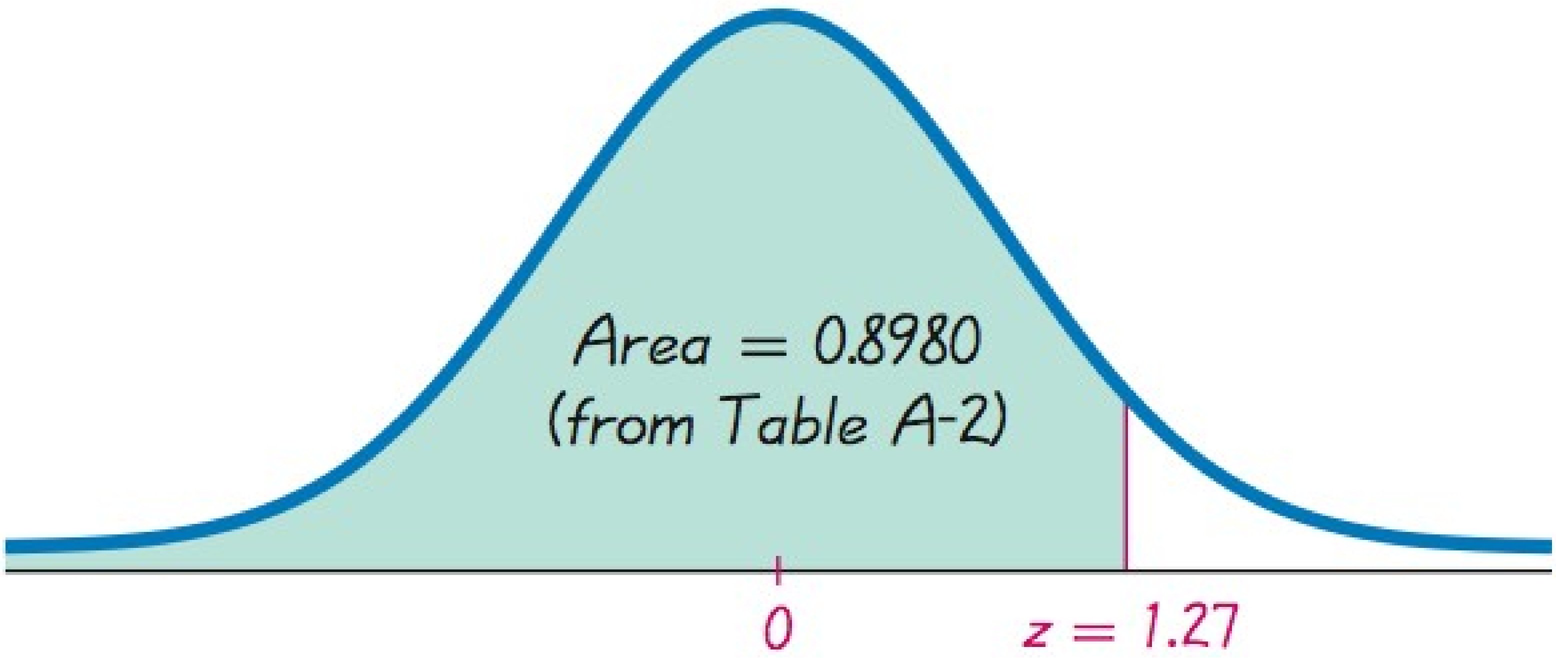
P (z < 1.27) = 0.8980



The *probability* of randomly selecting a thermometer with a reading less than 1.27º is 0.8980.

Example - cont

P (z < 1.27) = 0.8980

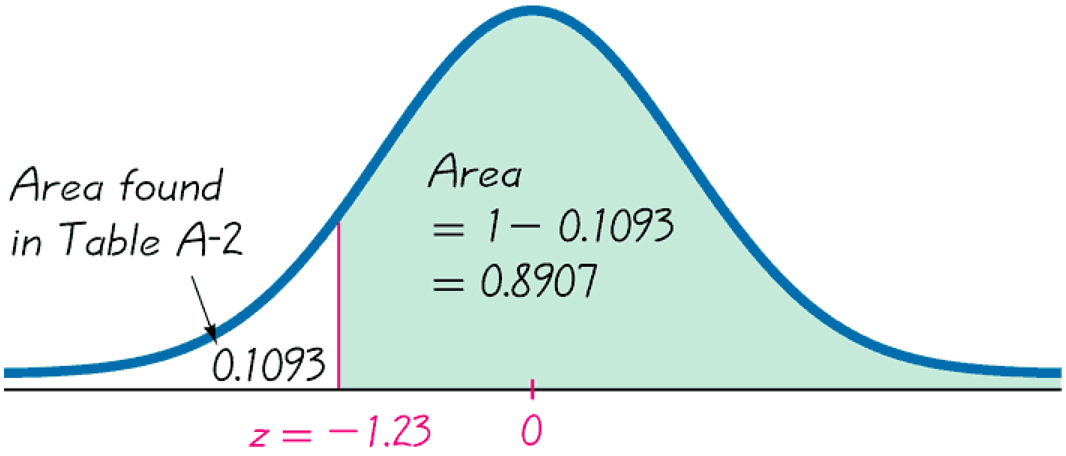


Or 89.80% will have readings below 1.27º.

Example - Thermometers Again

If thermometers have an average (mean) reading of 0 degrees and a standard deviation of 1 degree for freezing water, and if one thermometer is randomly selected, find the probability that it reads (at the freezing point of water) above –1.23 degrees.

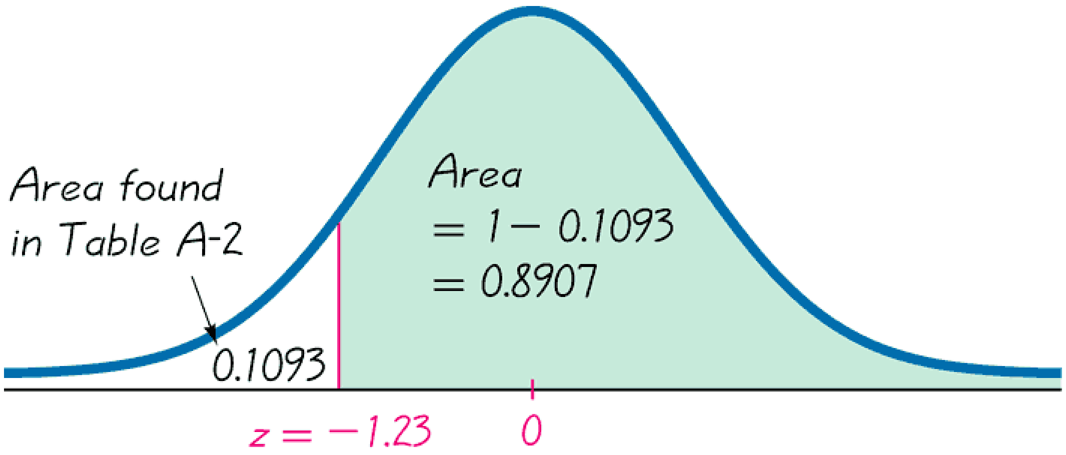
$P(z > -1.23) = 0.8907$



Probability of randomly selecting a thermometer with a reading above –1.23º is 0.8907.

Example - cont

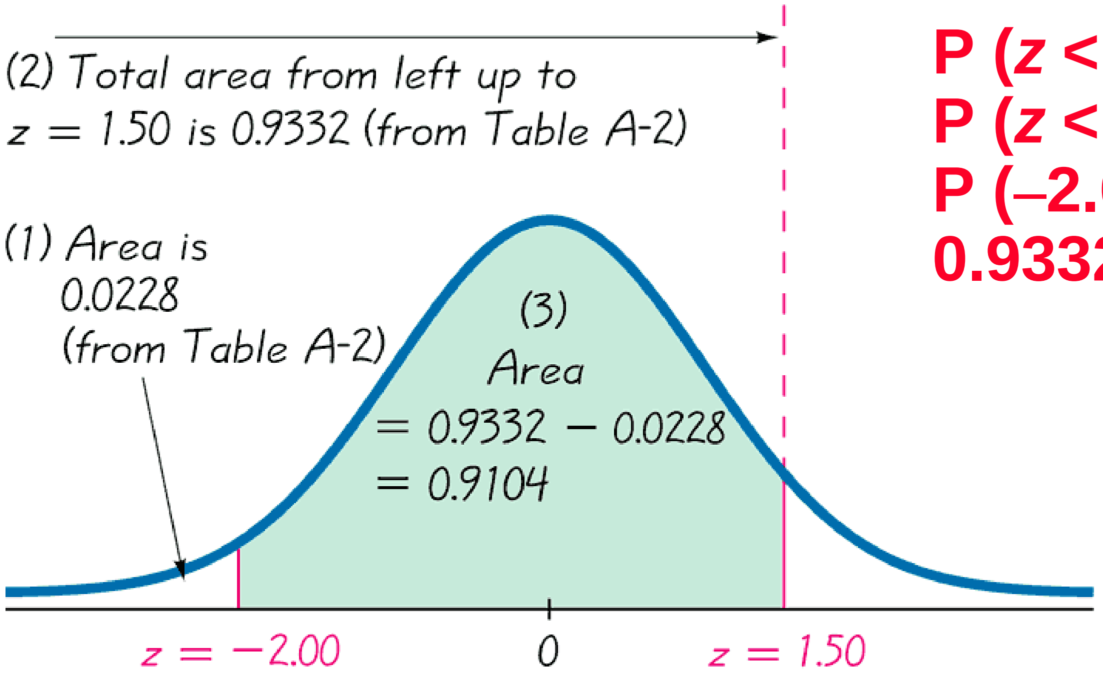
P (z > –1.23) = 0.8907



89.07% of the thermometers have readings above –1.23 degrees.

Example - Thermometers III

A thermometer is randomly selected. Find the probability that it reads (at the freezing point of water) between –2.00 and 1.50 degrees.

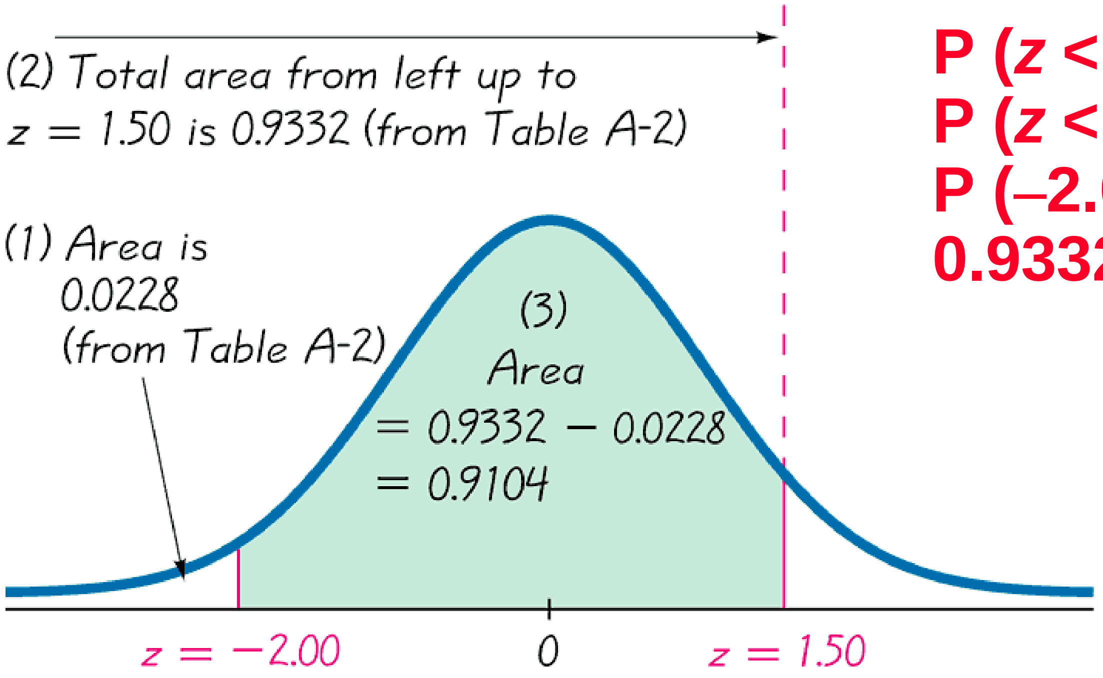


$P(z < -2.00) = 0.0228$
$P(z < 1.50) = 0.9332$
$P(-2.00 < z < 1.50) =$
$0.9332 - 0.0228 = 0.9104$

The probability that the chosen thermometer has a reading between – 2.00 and 1.50 degrees is 0.9104.

Example - cont

A thermometer is randomly selected. Find the probability that it reads (at the freezing point of water) between –2.00 and 1.50 degrees.



P (z < –2.00) = 0.0228
P (z < 1.50) = 0.9332
P (–2.00 < z < 1.50) =
0.9332 – 0.0228 = 0.9104

If many thermometers are selected and tested at the freezing point of water, then 91.04% of them will read between –2.00 and 1.50 degrees.

Notation

$$P(a < z < b)$$

denotes the probability that the *z* score is between *a* and *b.*

$$P(z > a)$$

denotes the probability that the *z* score is greater than *a.*

$$P(z < a)$$

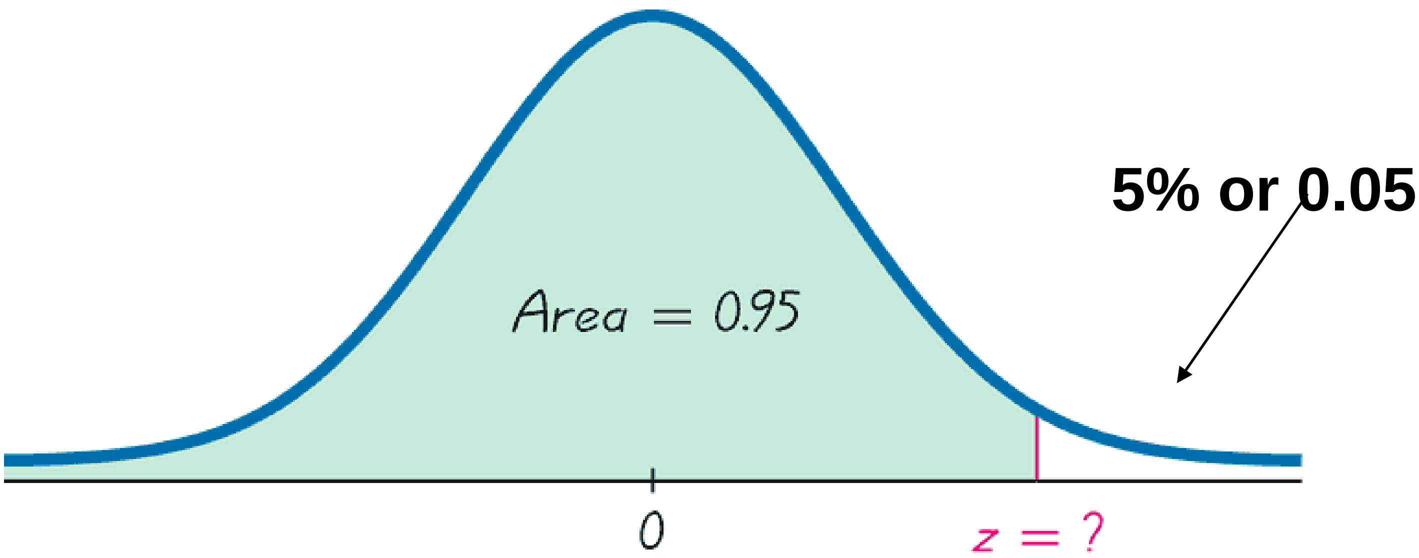
denotes the probability that the *z* score is less than *a.*

Finding a *z* Score When Given a Probability Using Table A-2

1. **Draw a bell-shaped curve and identify the region under the curve that corresponds to the given probability. If that region is not a cumulative region from the left, work instead with a known region that is a cumulative region from the left.**

2. **Using the cumulative area from the left, locate the closest probability in the body of Table A-2 and identify the corresponding *z* score.**

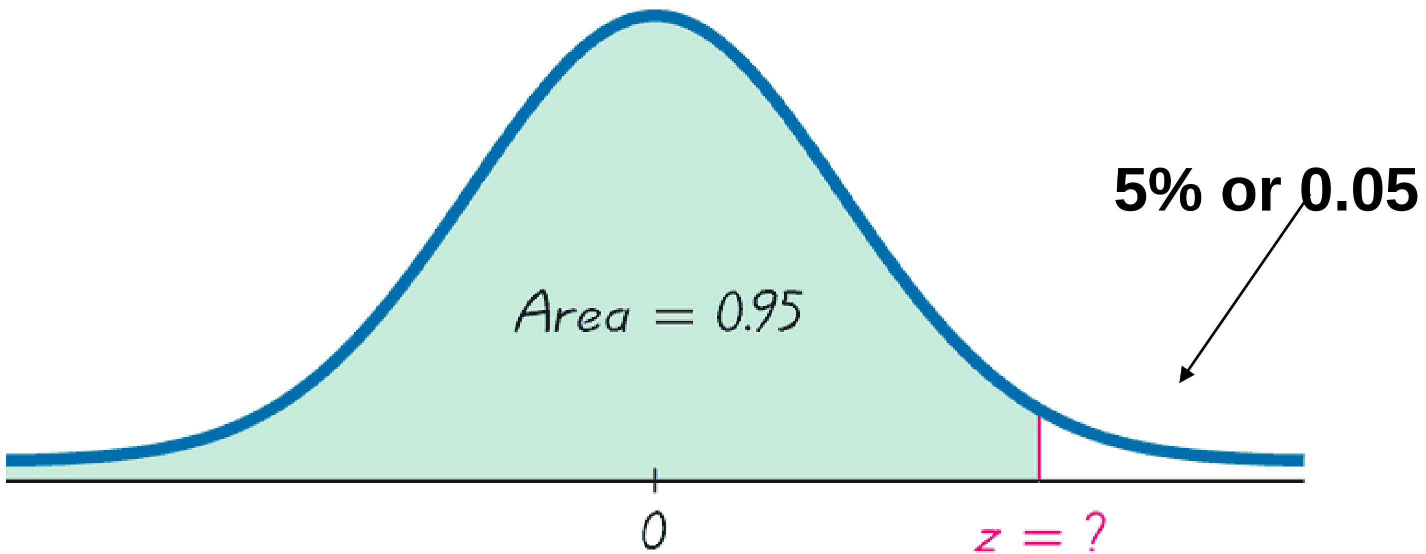
Finding *z* Scores When Given Probabilities



(*z* score will be positive)

Finding the 95th Percentile

Finding z Scores When Given Probabilities - cont

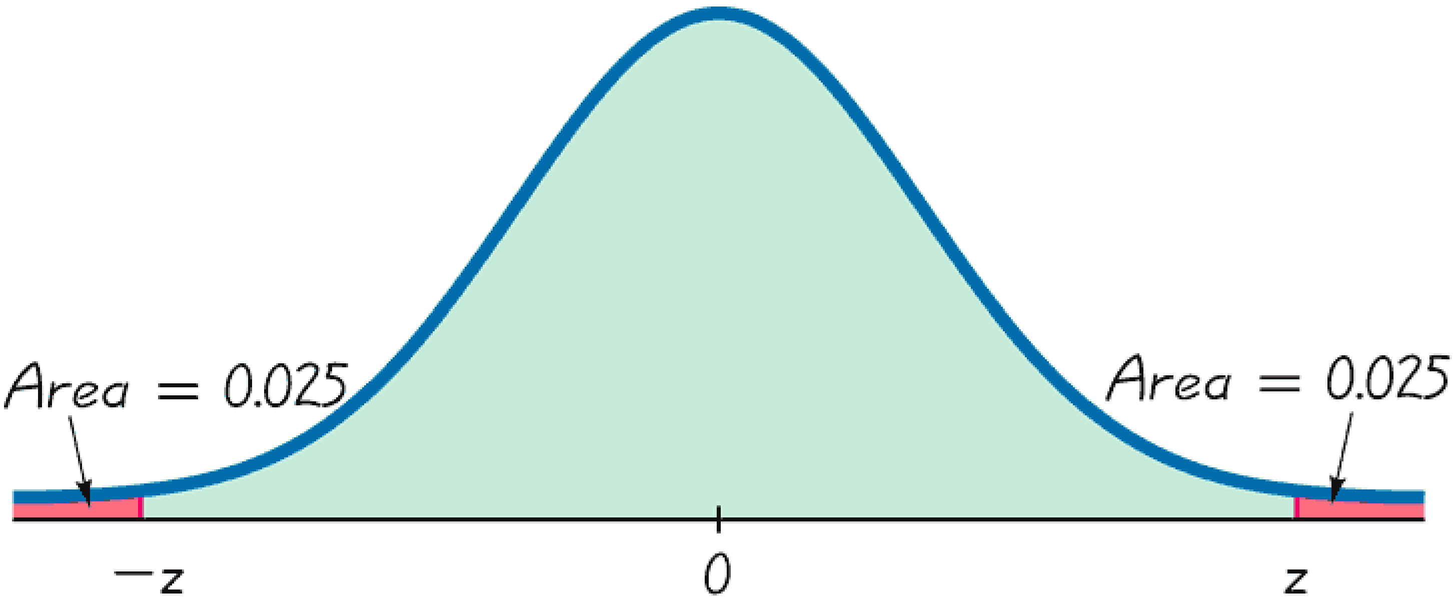


(z score will be positive)

1.645

Finding the 95th Percentile

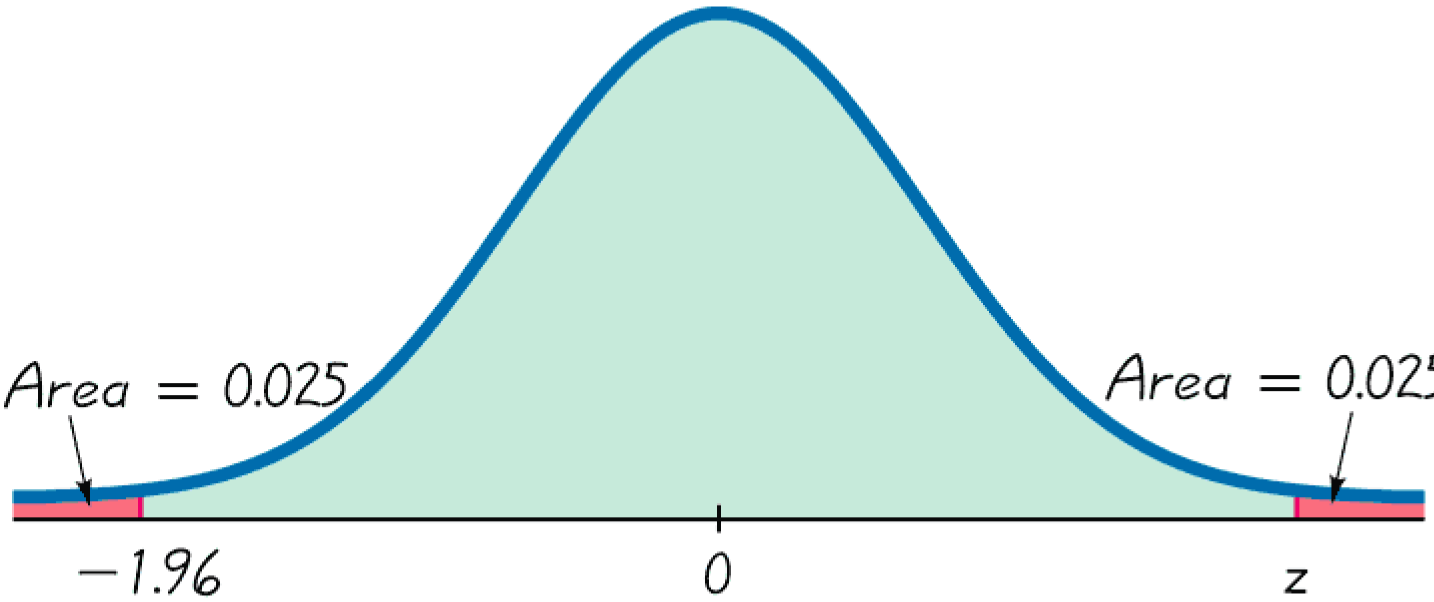
Finding *z* Scores When Given Probabilities - cont



(One *z* score will be negative and the other positive)

Finding the Bottom 2.5% and Upper 2.5%

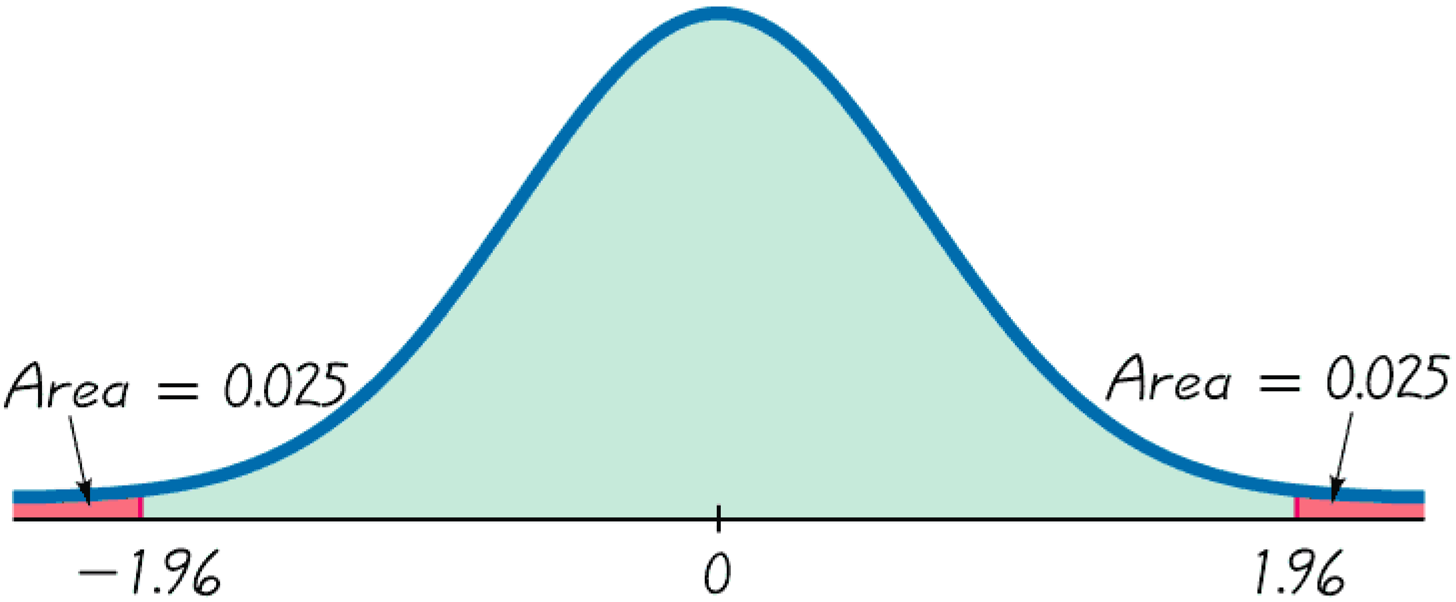
Finding *z* Scores When Given Probabilities - cont



(One *z* score will be negative and the other positive)

Finding the Bottom 2.5% and Upper 2.5%

Finding *z* Scores When Given Probabilities - cont



(One *z* score will be negative and the other positive)

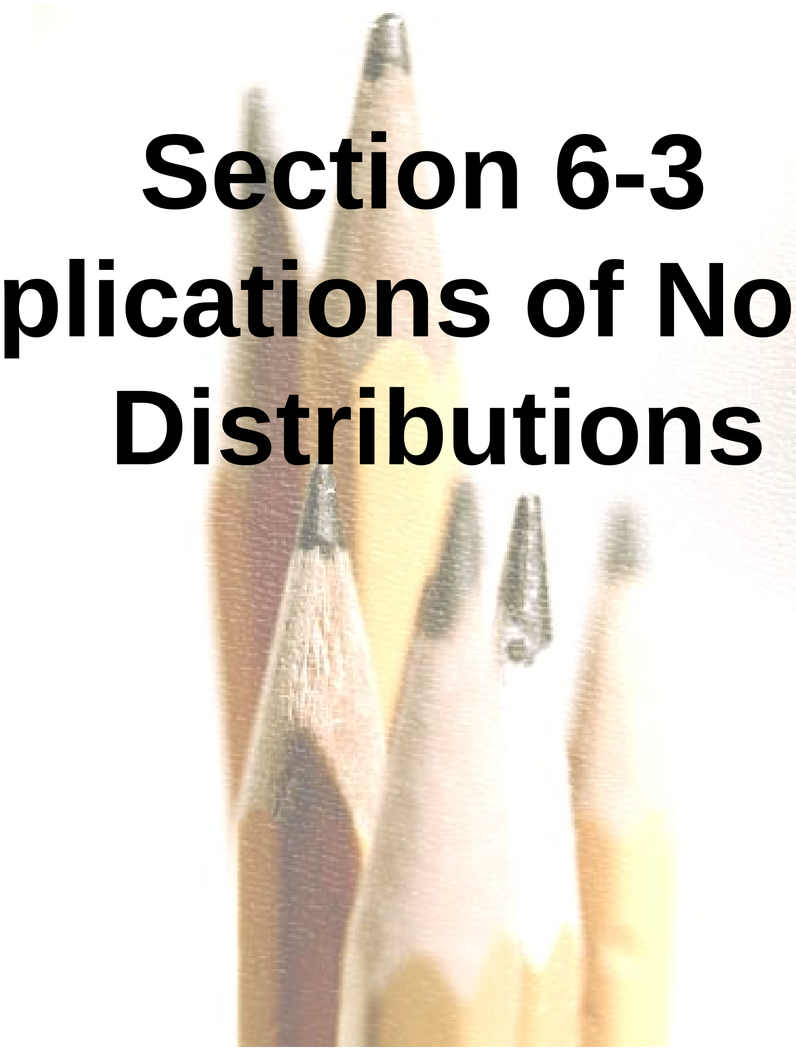
Finding the Bottom 2.5% and Upper 2.5%

Recap

In this section we have discussed:

- **Density curves.**
- **Relationship between area and probability.**
- **Standard normal distribution.**
- **Using Table A-2.**

Section 6-3
Applications of Normal Distributions



Key Concept

This section presents methods for working with normal distributions that are not standard. That is, the mean is not 0 or the standard deviation is not 1, or both.

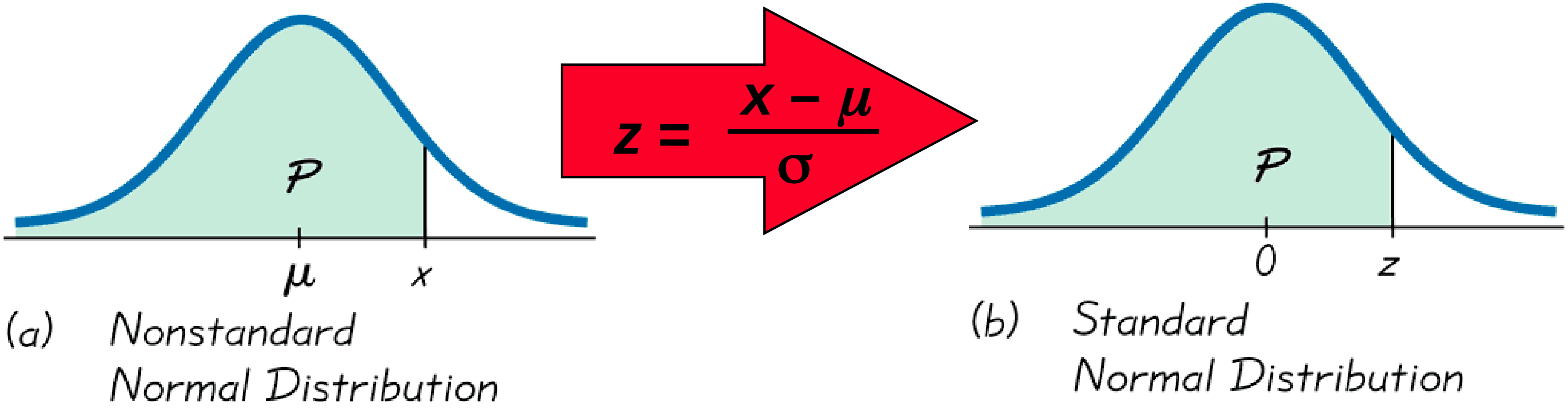
The key concept is that we can use a simple conversion that allows us to standardize any normal distribution so that the same methods of the previous section can be used.

Conversion Formula

$$z = \frac{x - \mu}{\sigma}$$

Round *z* scores to 2 decimal places

Converting to a Standard Normal Distribution



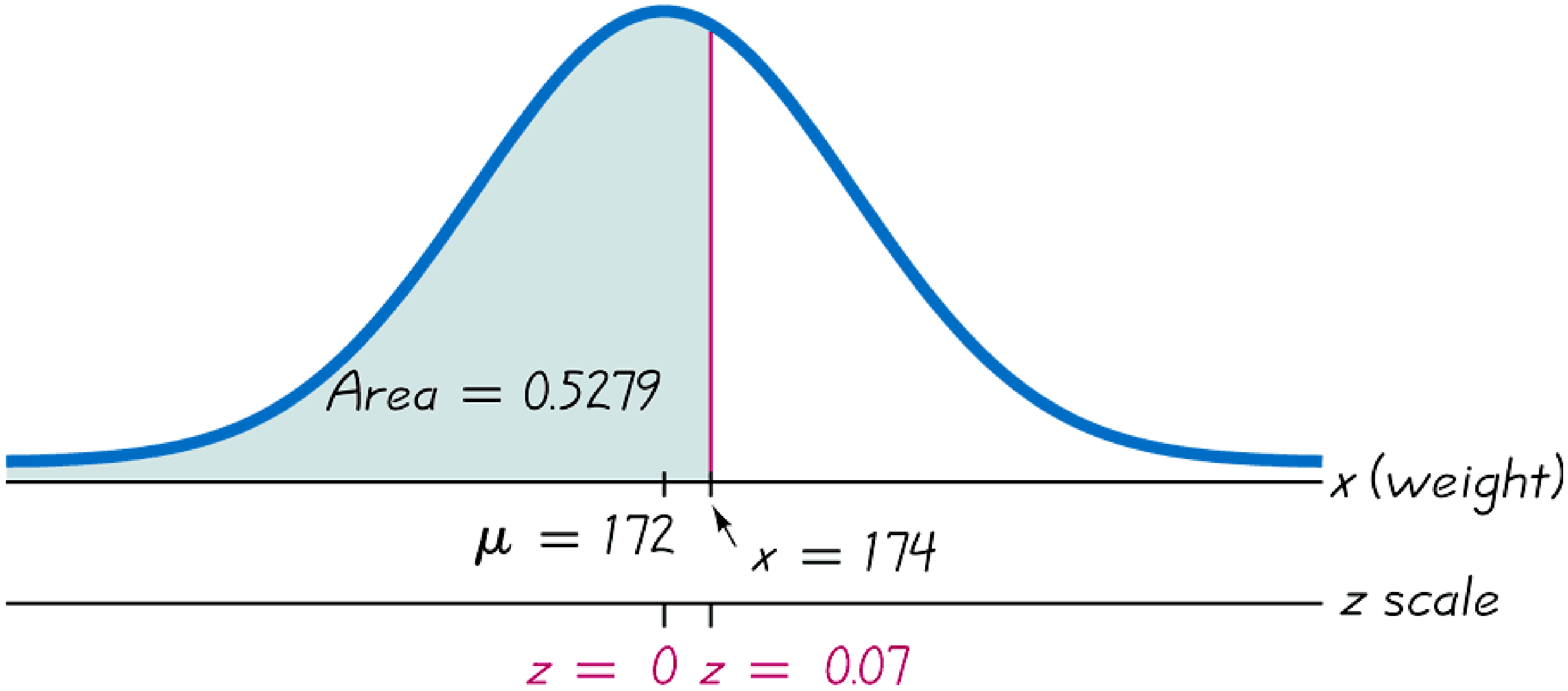
Example – Weights of Water Taxi Passengers

In the Chapter Problem, we noted that the safe load for a water taxi was found to be 3500 pounds. We also noted that the mean weight of a passenger was assumed to be 140 pounds. Assume the worst case that all passengers are men. Assume also that the weights of the men are normally distributed with a mean of 172 pounds and standard deviation of 29 pounds. If one man is randomly selected, what is the probability he weighs less than 174 pounds?

Example - cont

$\mu = 172$
$\sigma = 29$

$$z = \frac{174 - 172}{29} = 0.07$$

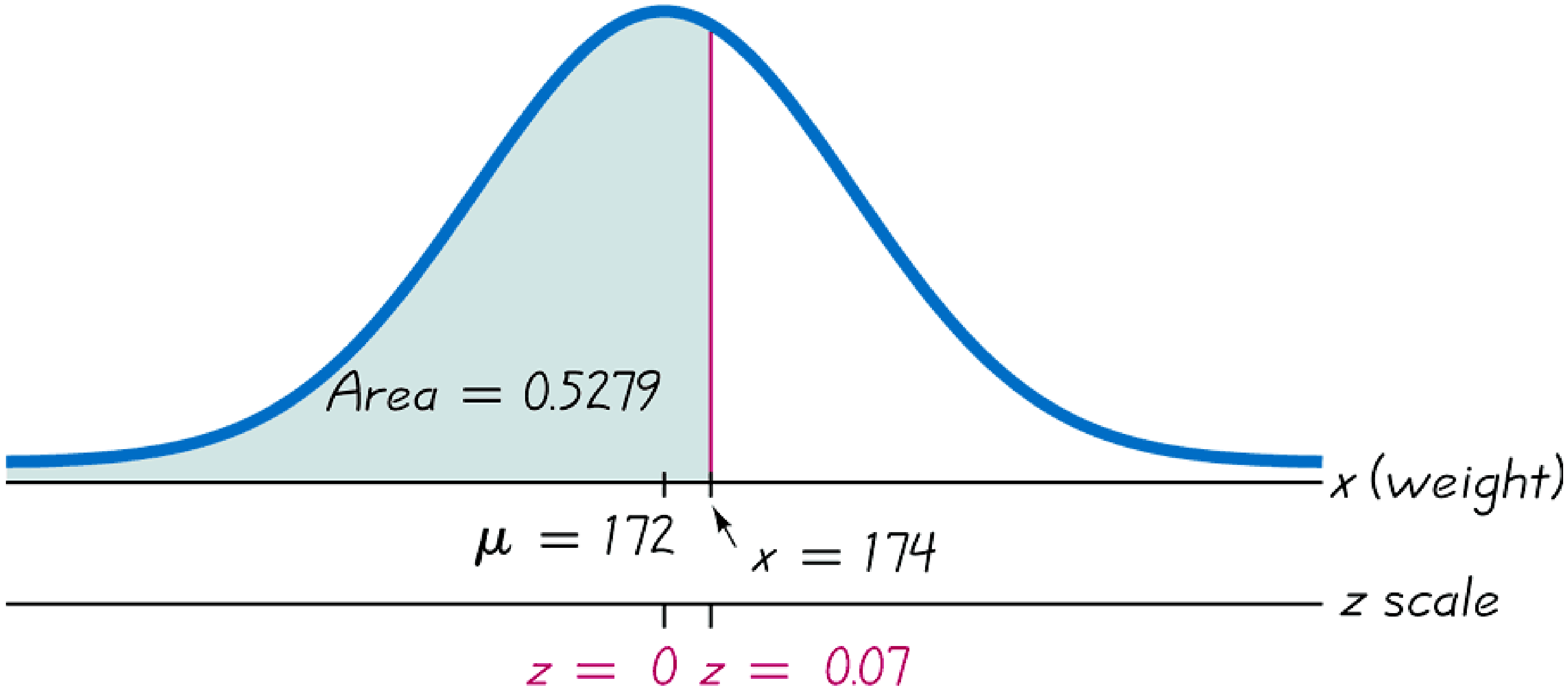


Example - cont

$\mu =$ 172

$\sigma =$ 29

P (x < 174 lb.) = P(z < 0.07)



Helpful Hints

1. Don’t confuse *z* scores and areas. *z* scores are distances along the horizontal scale, but areas are regions under the normal curve. Table A-2 lists *z* scores in the left column and across the top row, but areas are found in the body of the table.

2. Choose the correct (right/left) side of the graph.

3. A z score must be negative whenever it is located in the left half of the normal distribution.

4. Areas (or probabilities) are positive or zero values, but they are never negative.

Procedure for Finding Values Using Table A-2 and Formula 6-2

1. **Sketch a normal distribution curve, enter the given probability or percentage in the appropriate region of the graph, and identify the *x* value(s) being sought.**

2. **Use Table A-2 to find the *z* score corresponding to the cumulative left area bounded by *x*. Refer to the body of Table A-2 to find the closest area, then identify the corresponding *z* score.**

3. **Using Formula 6-2, enter the values for μ, σ, and the *z* score found in step 2, then solve for *x*.**

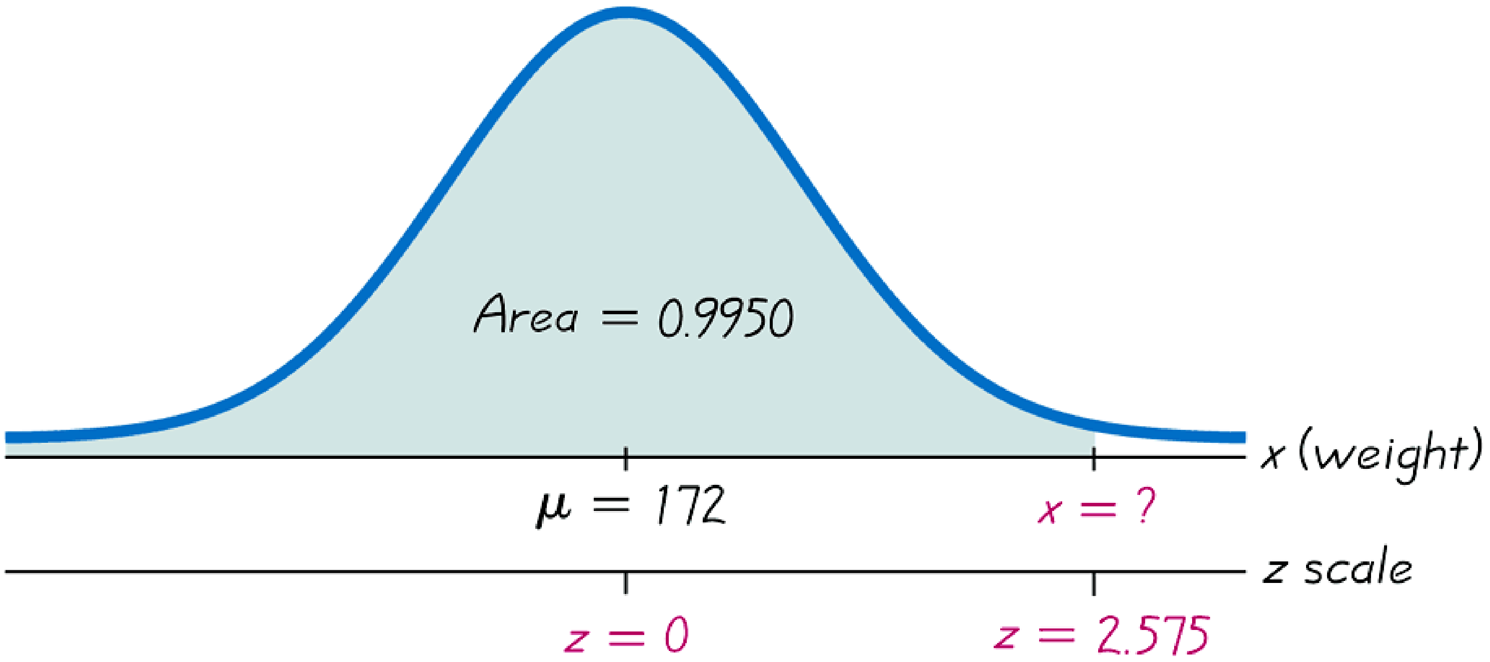
   $$x = \mu + (z \cdot \sigma)$$ **(Another form of Formula 6-2)**

   (If *z* is located to the left of the mean, be sure that it is a negative number.)

4. **Refer to the sketch of the curve to verify that the solution makes sense in the context of the graph and the context of the problem.**

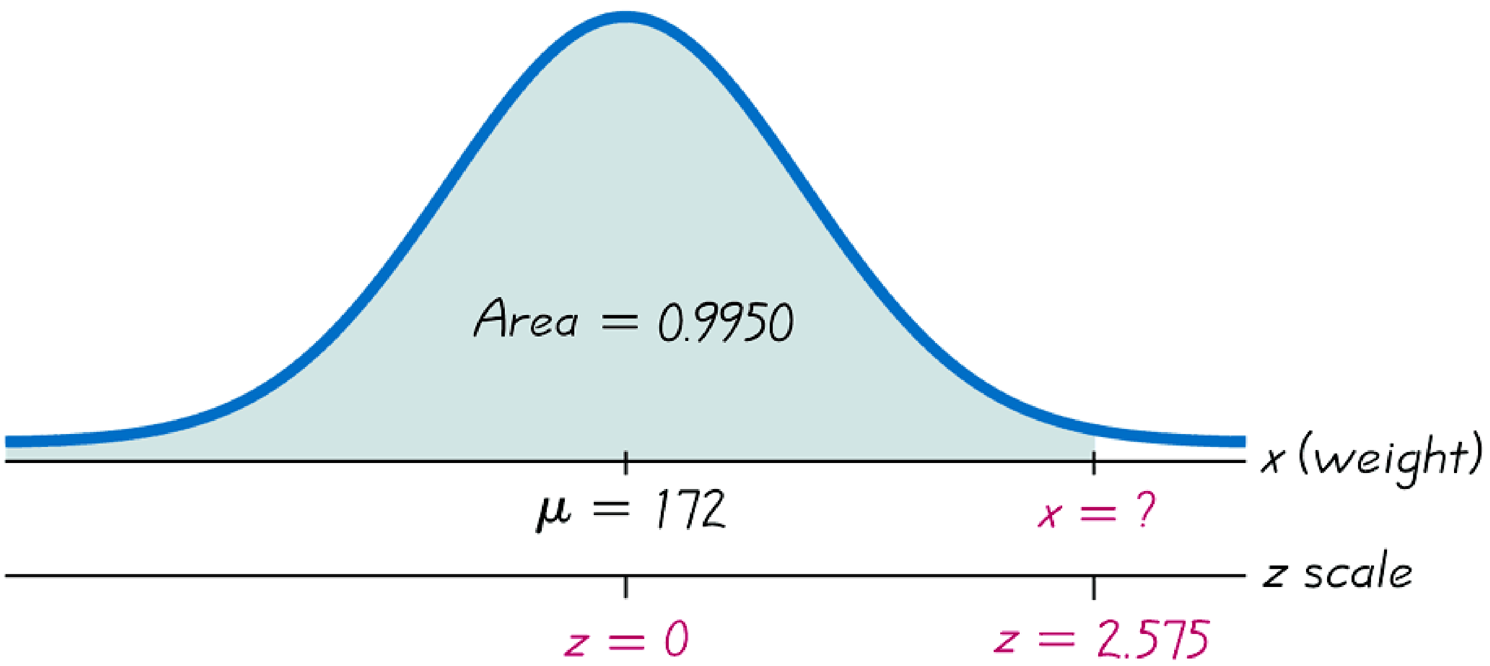
Example – Lightest and Heaviest

Use the data from the previous example to determine what weight separates the lightest 99.5% from the heaviest 0.5%?



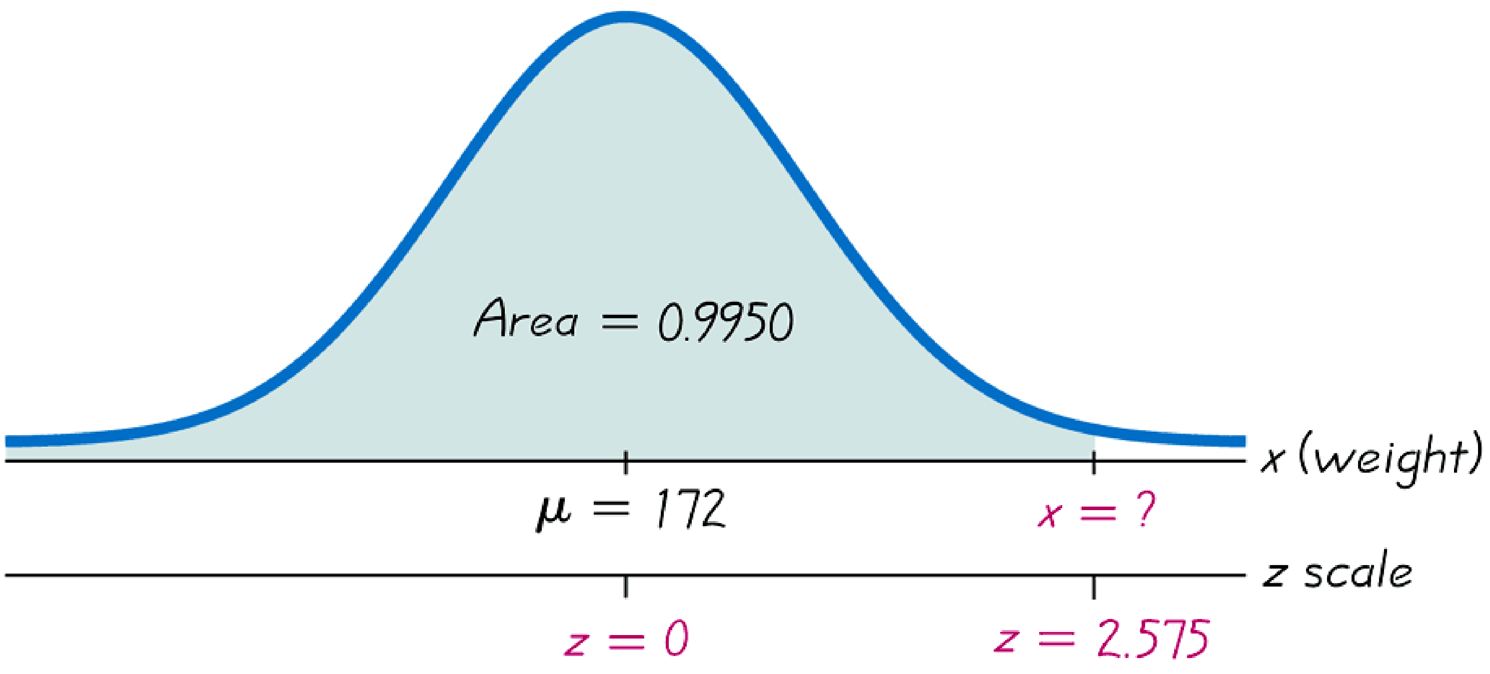
Example – Lightest and Heaviest - cont

$$x = \mu + (z \bullet \sigma)$$
$$x = 172 + (2.575 \bullet 29)$$
$$x = 246.675 \text{ (247 rounded)}$$

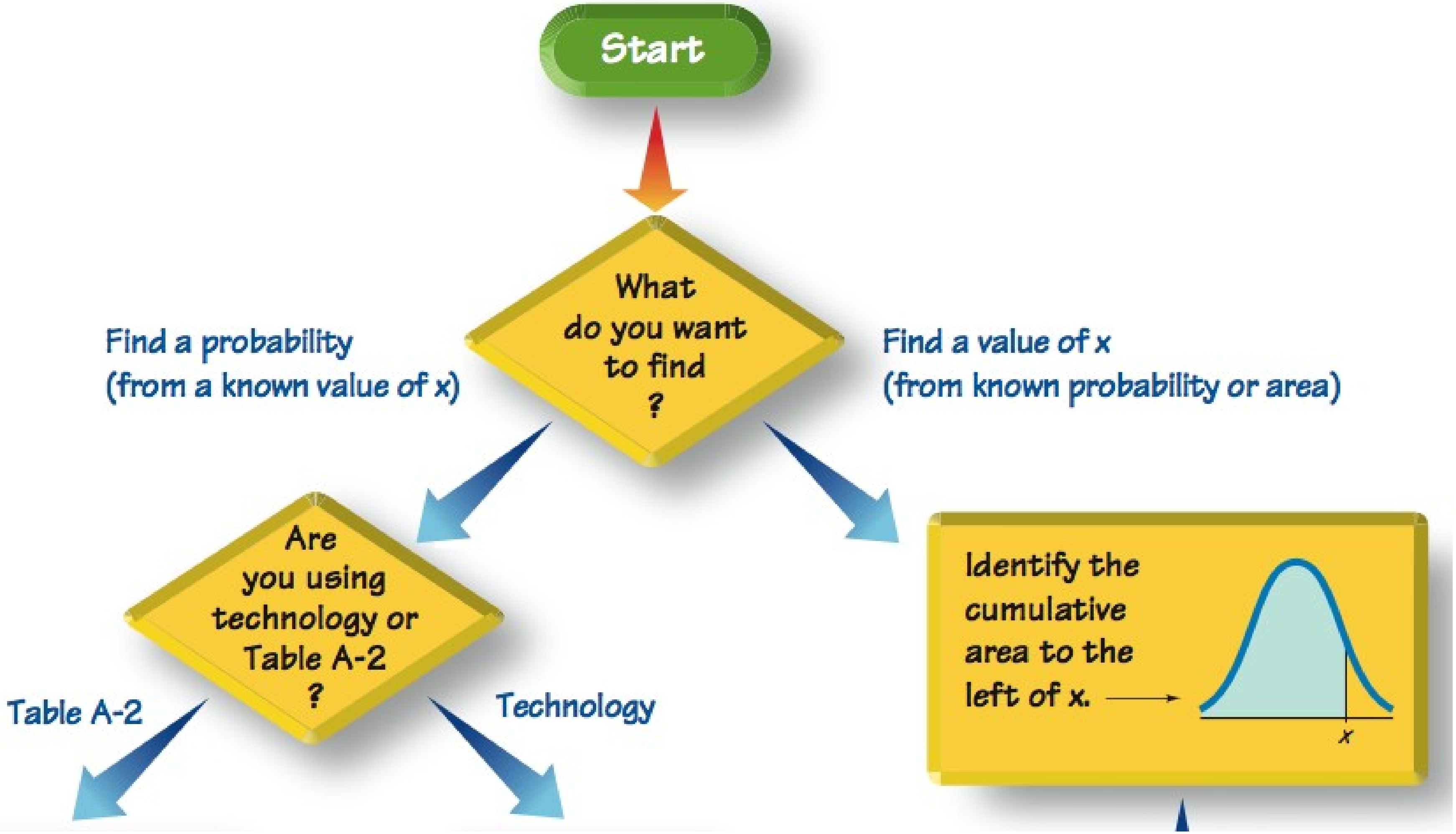


Example – Lightest and Heaviest - cont

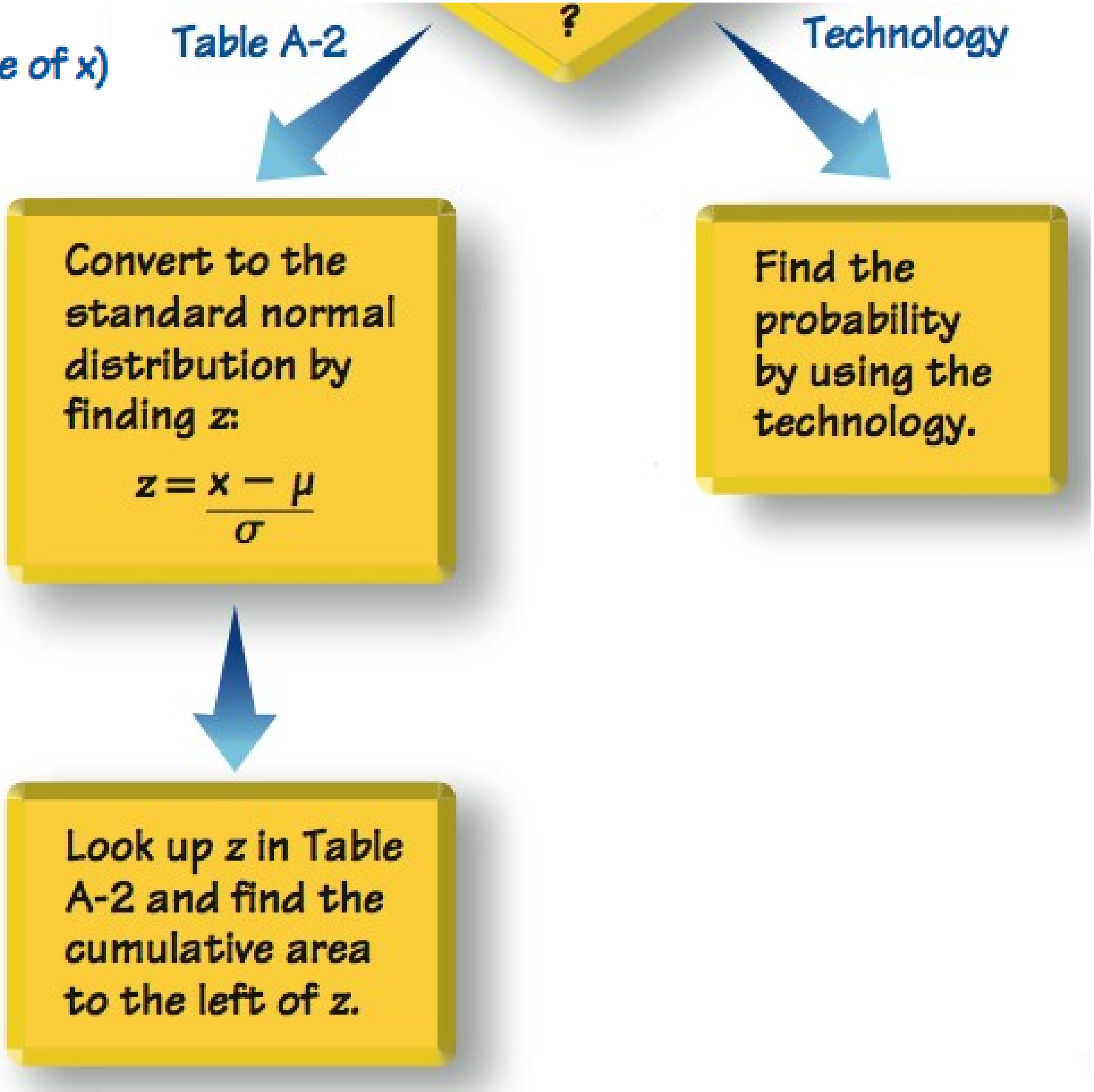
The weight of 247 pounds separates the lightest 99.5% from the heaviest 0.5%



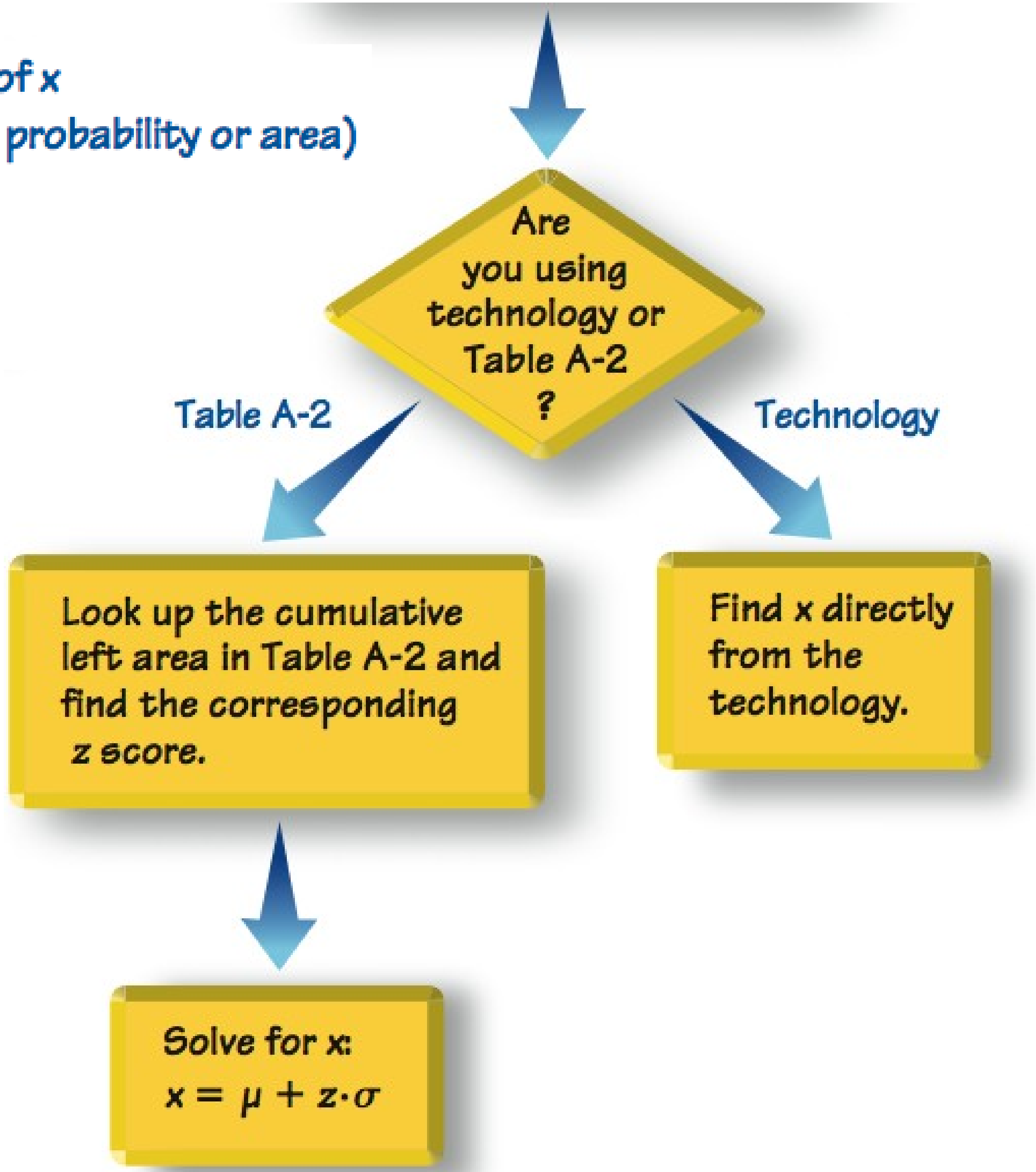
Applications with Normal Distributions



Find a probability (from a known value of x)



Find a value of x (from known probability or area)

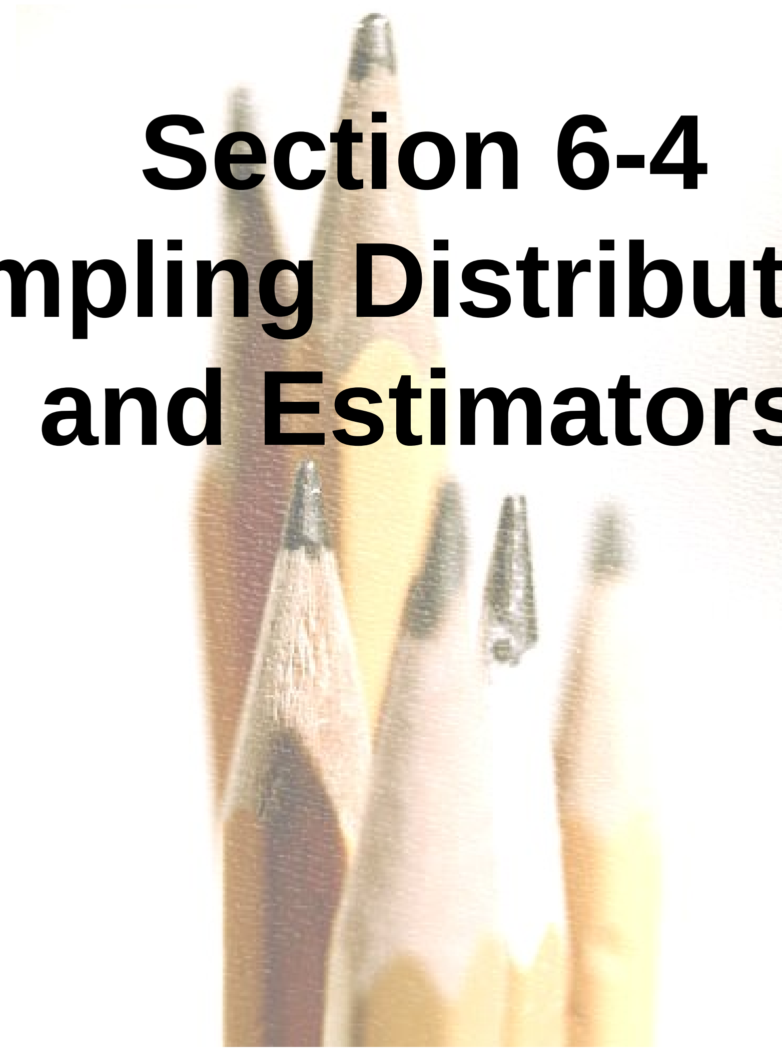


Recap

In this section we have discussed:

- **Non-standard normal distribution.**
- **Converting to a standard normal distribution.**
- **Procedures for finding values using Table A-2 and Formula 6-2.**

Section 6-4
Sampling Distributions and Estimators



Key Concept

The main objective of this section is to understand the concept of a sampling distribution of a statistic, which is the distribution of all values of that statistic when all possible samples of the same size are taken from the same population.

We will also see that some statistics are better than others for estimating population parameters.

Definition

The **sampling distribution of a statistic** (such as the sample mean or sample proportion) is the distribution of all values of the statistic when all possible samples of the same size n are taken from the same population. (The sampling distribution of a statistic is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

Definition

The **sampling distribution of the mean** is the distribution of sample means, with all samples having the same sample size n taken from the same population. (The sampling distribution of the mean is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

Properties

- **Sample means target the value of the population mean. (That is, the mean of the sample means is the population mean. The expected value of the sample mean is equal to the population mean.)**
- **The distribution of the sample means tends to be a normal distribution.**

Definition

The **sampling distribution of the variance** is the distribution of sample variances, with all samples having the same sample size *n* taken from the same population. (The sampling distribution of the variance is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

Properties

- Sample variances target the value of the population variance. (That is, the mean of the sample variances is the population variance. The expected value of the sample variance is equal to the population variance.)
- The distribution of the sample variances tends to be a distribution skewed to the right.

Definition

The sampling distribution of the proportion is the distribution of sample proportions, with all samples having the same sample size *n* taken from the same population.

Definition

We need to distinguish between a population proportion *p* and some sample proportion:

p **= population proportion**

$\hat{p}$ **= sample proportion**

Properties

- Sample proportions target the value of the population proportion. (That is, the mean of the sample proportions is the population proportion. The expected value of the sample proportion is equal to the population proportion.)
- The distribution of the sample proportion tends to be a normal distribution.

Unbiased Estimators

Sample means, variances and proportions are unbiased estimators.

That is they target the population parameter.

These statistics are better in estimating the population parameter.

Biased Estimators

Sample medians, ranges and standard deviations are biased estimators.

That is they do NOT target the population parameter.

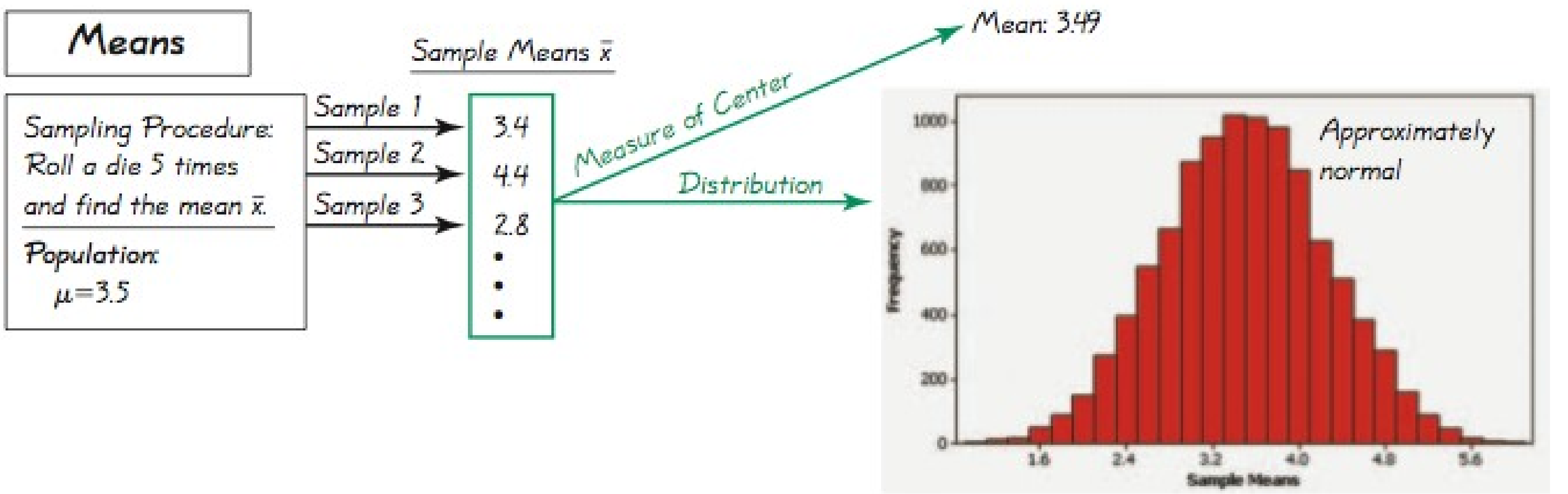
Note: the bias with the standard deviation is relatively small in large samples so *s* is often used to estimate.

Example - Sampling Distributions

Consider repeating this process: Roll a die 5 times, find the mean $\bar{x}$, variance s^2, and the proportion of *odd* numbers of the results. What do we know about the behavior of all sample means that are generated as this process continues indefinitely?

Example - Sampling Distributions

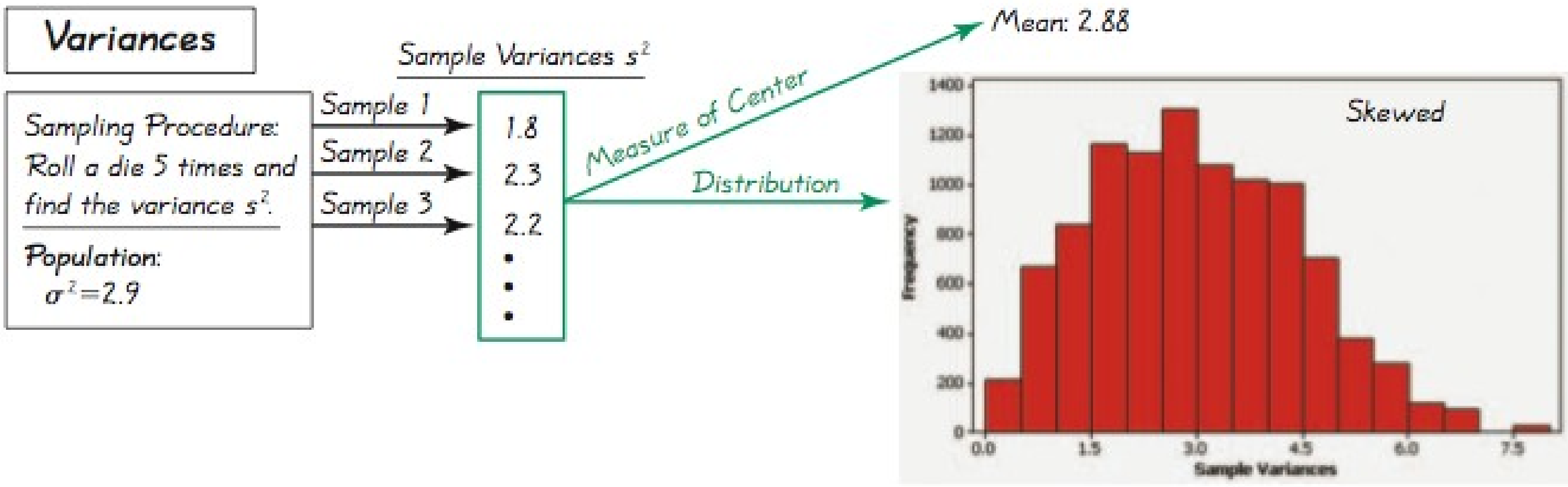
Specific results from 10,000 trials



All outcomes are equally likely so the population mean is 3.5; the mean of the 10,000 trials is 3.49. If continued indefinitely, the sample mean will be 3.5. Also, notice the distribution is "normal."

Example - Sampling Distributions

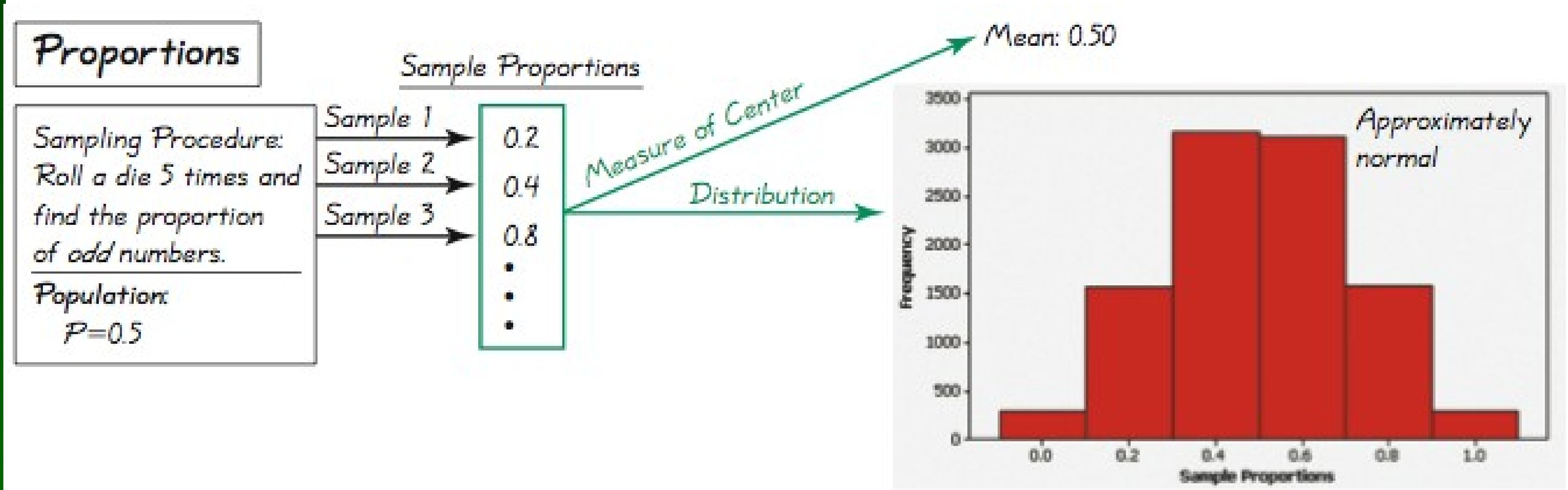
Specific results from 10,000 trials



All outcomes are equally likely so the population variance is 2.9; the mean of the 10,000 trials is 2.88. If continued indefinitely, the sample variance will be 2.9. Also, notice the distribution is "skewed to the right."

Example - Sampling Distributions

Specific results from 10,000 trials



All outcomes are equally likely so the population proportion of odd numbers is 0.50; the proportion of the 10,000 trials is 0.50. If continued indefinitely, the mean of sample proportions will be 0.50. Also, notice the distribution is "approximately normal."

Why Sample with Replacement?

Sampling *without replacement* would have the very practical advantage of avoiding wasteful duplication whenever the same item is selected more than once. However, we are interested in sampling *with replacement* for these two reasons:

1. **When selecting a relatively small sample form a large population, it makes no significant difference whether we sample with replacement or without replacement.**
2. **Sampling with replacement results in independent events that are unaffected by previous outcomes, and independent events are easier to analyze and result in simpler calculations and formulas.**

Caution

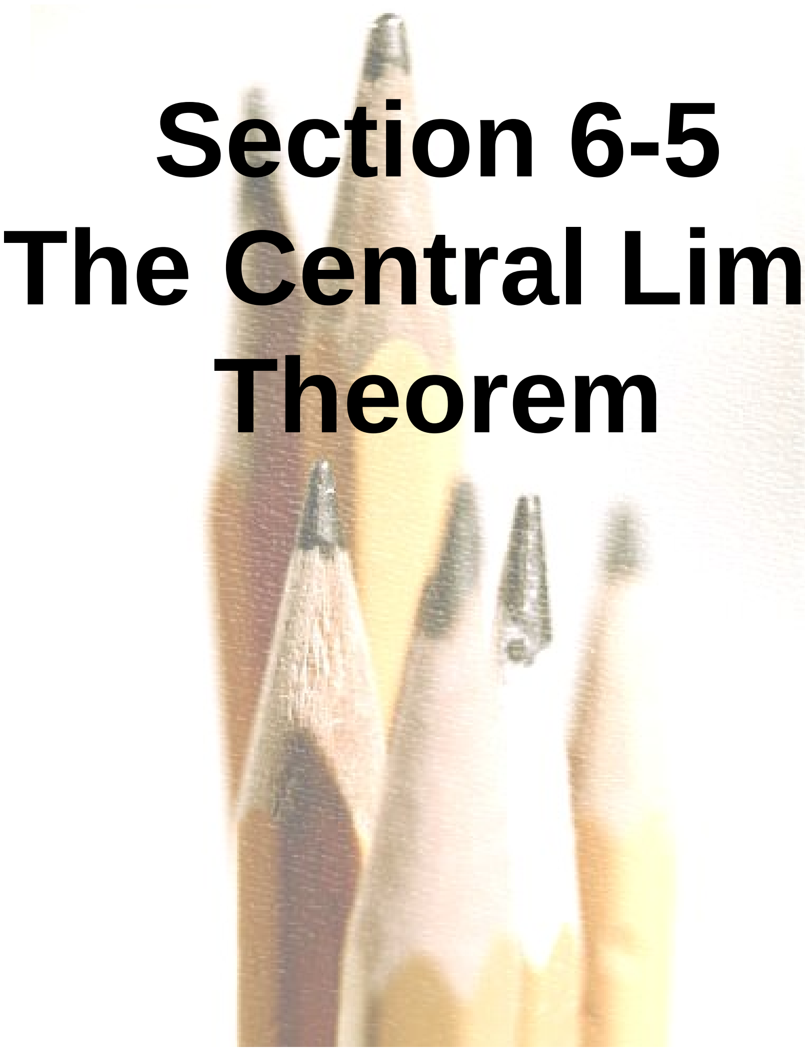
Many methods of statistics require a *simple random sample*. Some samples, such as voluntary response samples or convenience samples, could easily result in very wrong results.

Recap

In this section we have discussed:

- **Sampling distribution of a statistic.**
- **Sampling distribution of the mean.**
- **Sampling distribution of the variance.**
- **Sampling distribution of the proportion.**
- **Estimators.**

Section 6-5
The Central Limit Theorem



Key Concept

The *Central Limit Theorem* tells us that for a population with *any* distribution, the distribution of the sample means approaches a normal distribution as the sample size increases.

The procedure in this section form the foundation for estimating population parameters and hypothesis testing.

Central Limit Theorem

Given:

1. The random variable x has a distribution (which may or may not be normal) with mean μ and standard deviation σ.
2. Simple random samples all of size n are selected from the population. (The samples are selected so that all possible samples of the same size n have the same chance of being selected.)

Central Limit Theorem – cont.

Conclusions:

1. The distribution of sample $\bar{x}$ will, as the sample size increases, approach a normal distribution.
2. The mean of the sample means is the population mean μ.
3. The standard deviation of all sample means is $\sigma/\sqrt{n}$.

Practical Rules Commonly Used

1. **For samples of size *n* larger than 30, the distribution of the sample means can be approximated reasonably well by a normal distribution. The approximation gets closer to a normal distribution as the sample size *n* becomes larger.**

2. **If the original population is *normally distributed*, then for any sample size *n*, the sample means will be normally distributed (not just the values of *n* larger than 30).**

Notation

the mean of the sample means

$$\mu_{\bar{x}} = \mu$$

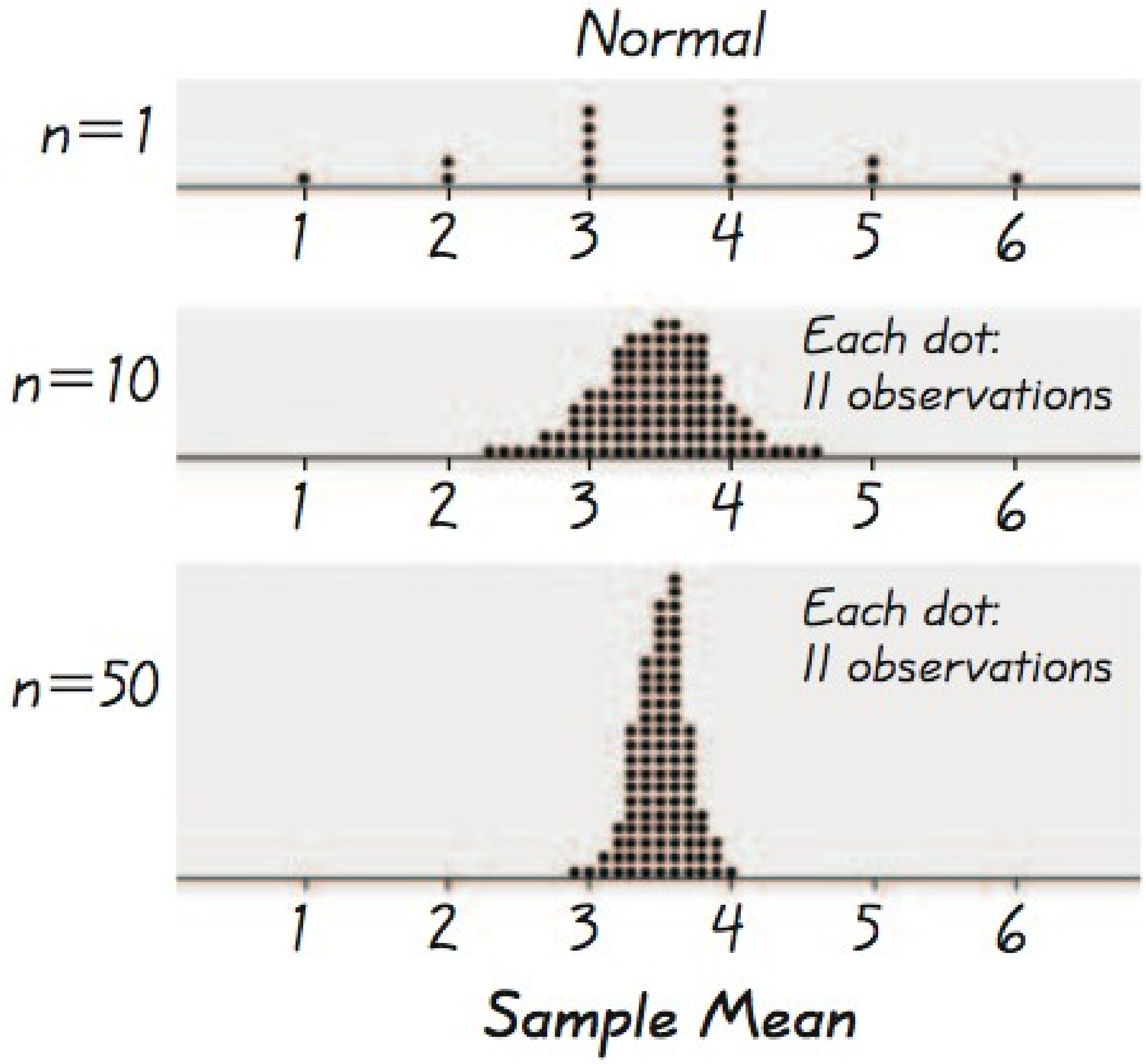
the standard deviation of sample mean

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

(often called the standard error of the mean)

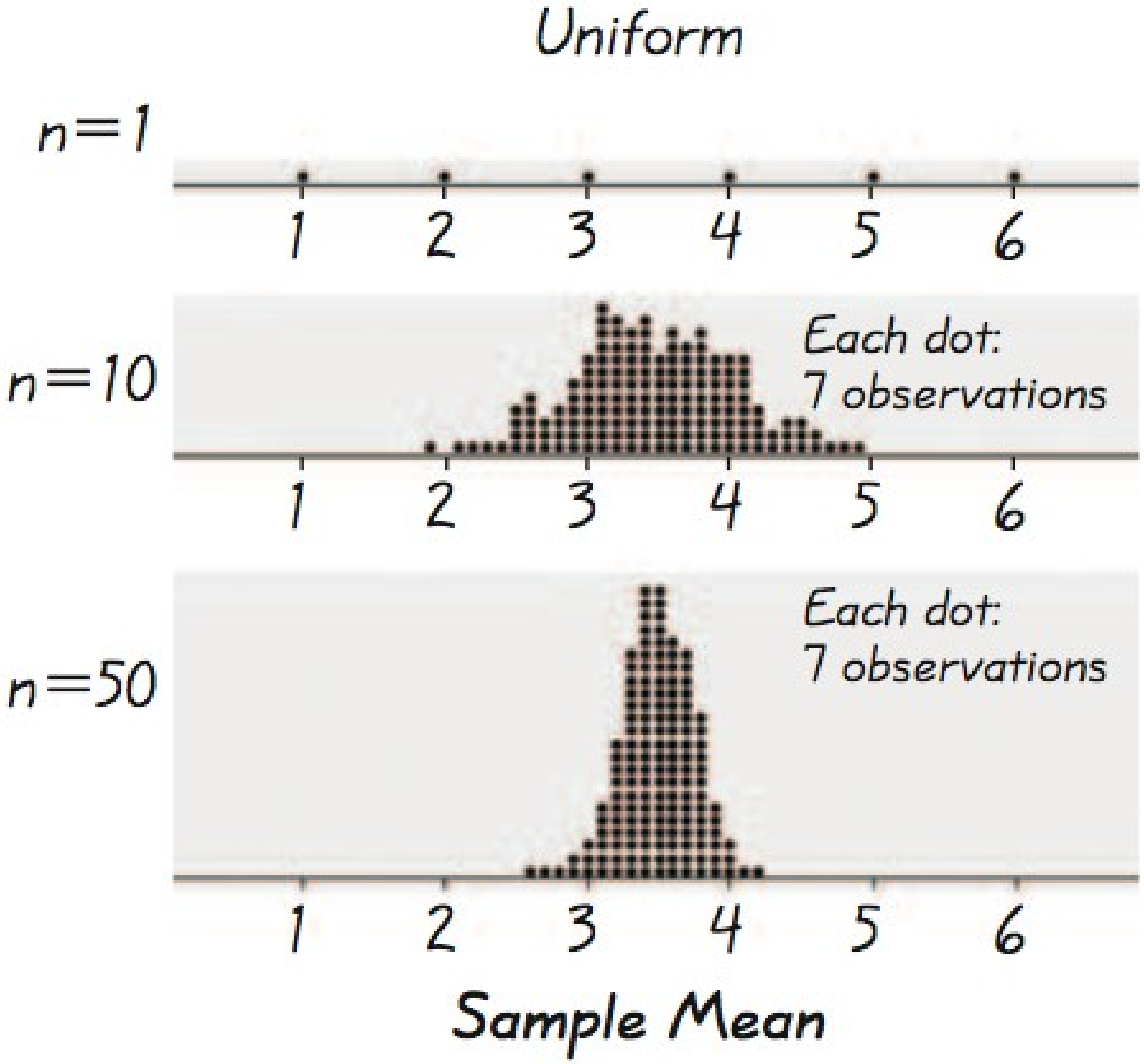
Example - Normal Distribution

As we proceed from *n* = 1 to *n* = 50, we see that the distribution of sample means is approaching the shape of a normal distribution.



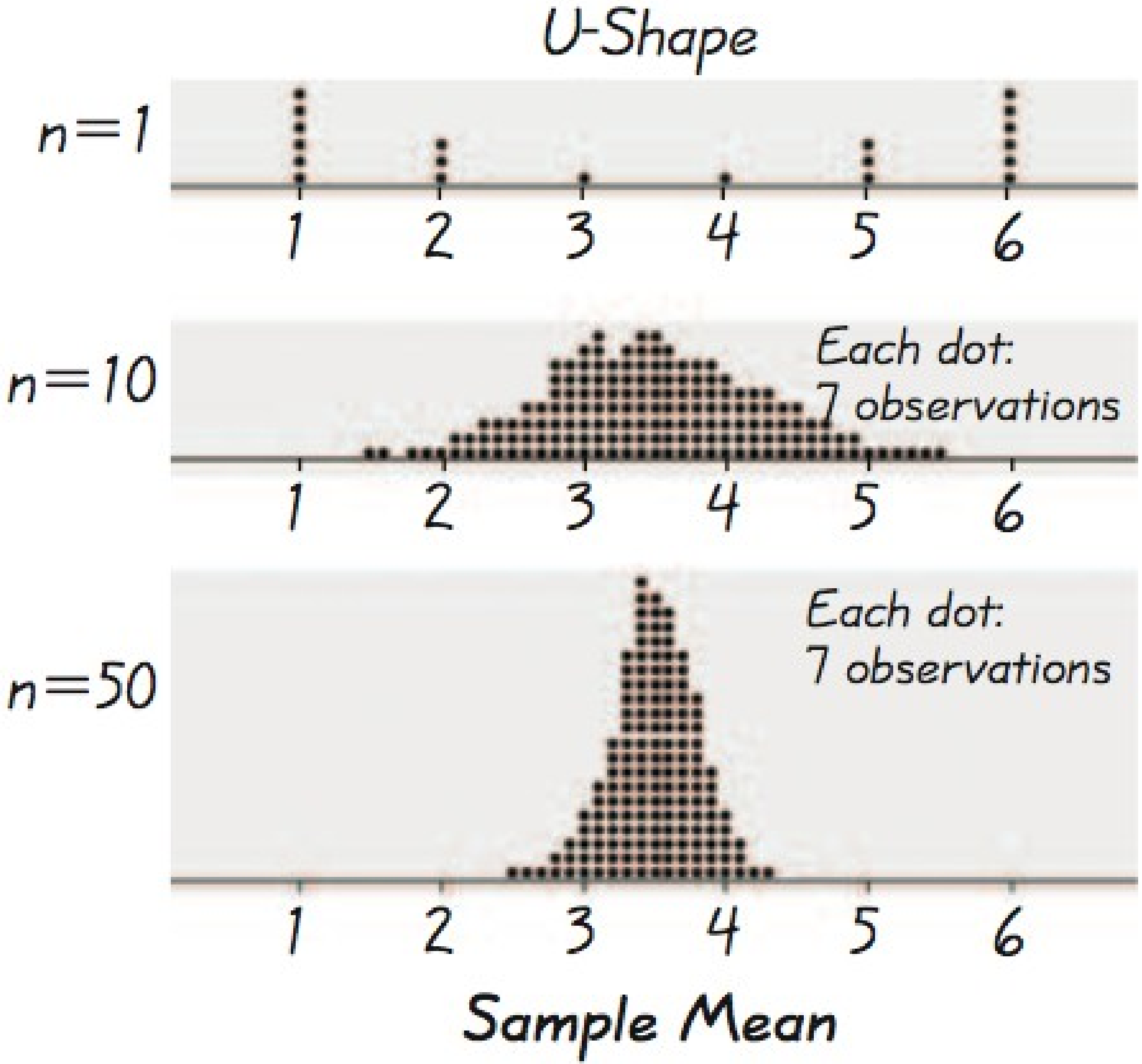
Example - Uniform Distribution

As we proceed from *n* = 1 to *n* = 50, we see that the distribution of sample means is approaching the shape of a normal distribution.



Example - U-Shaped Distribution

As we proceed from *n* = 1 to *n* = 50, we see that the distribution of sample means is approaching the shape of a normal distribution.



Important Point

As the sample size increases, the sampling distribution of sample means approaches a normal distribution.

Example – Water Taxi Safety

Use the Chapter Problem. Assume the population of weights of men is normally distributed with a mean of 172 lb and a standard deviation of 29 lb.

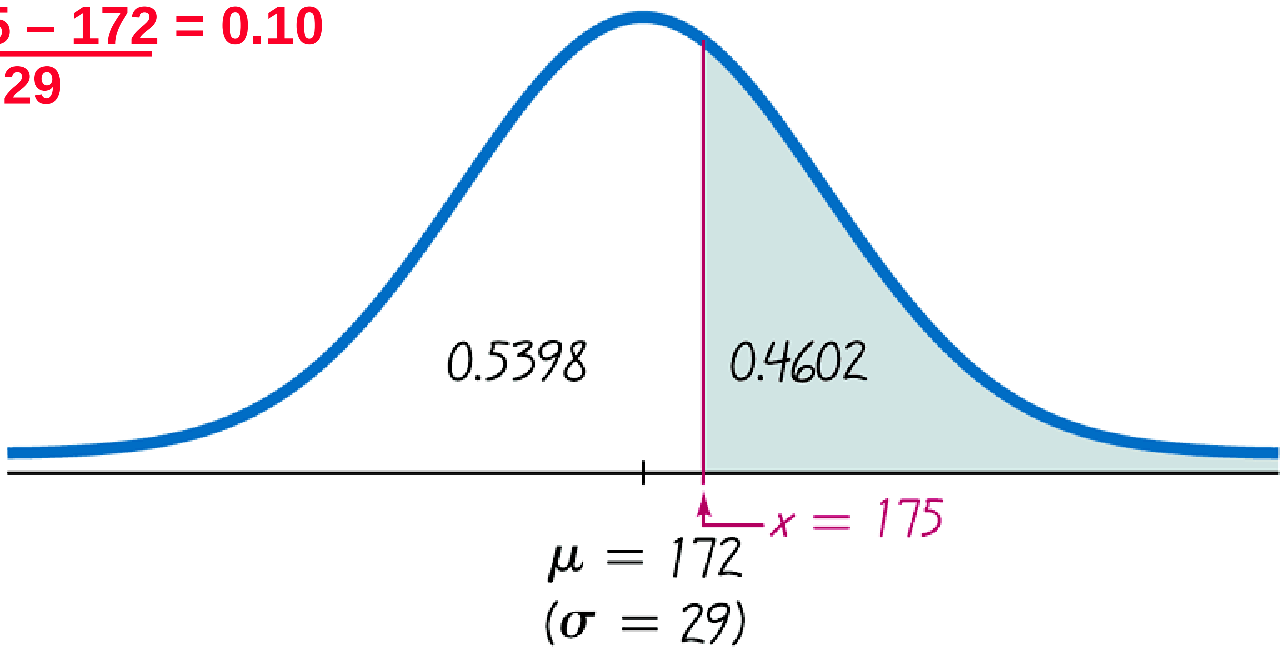
a) Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.

b) b) Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).

Example – cont

a) Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.

$$z = \frac{175 - 172}{29} = 0.10$$

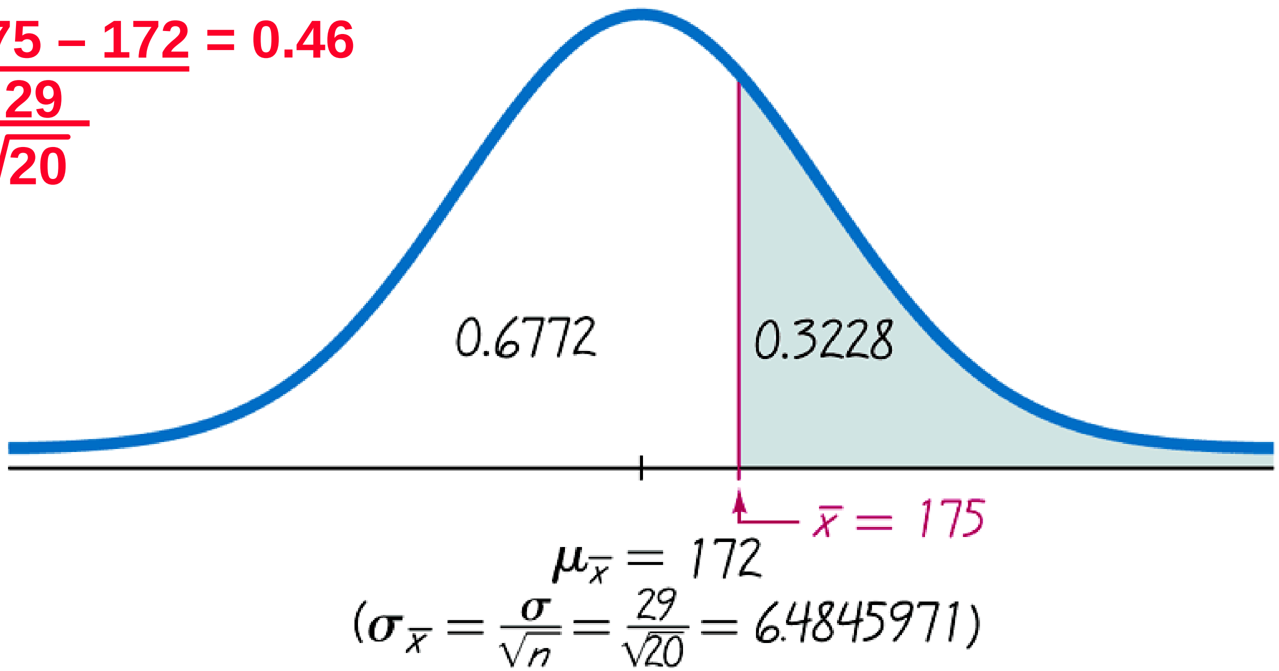


(a)

Example – cont

b) Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).

$$z = \frac{175 - 172}{\frac{29}{\sqrt{20}}} = 0.46$$



(b)

Example - cont

a) **Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.**

$$P(x > 175) = 0.4602$$

b) **Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).**

$$P(\bar{x} > 175) = 0.3228$$

It is much easier for an individual to deviate from the mean than it is for a group of 20 to deviate from the mean.

Interpretation of Results

Given that the safe capacity of the water taxi is 3500 pounds, there is a fairly good chance (with probability 0.3228) that it will be overloaded with 20 randomly selected men.

Correction for a Finite Population

When sampling without replacement and the sample size n is greater than 5% of the finite population of size N (that is, $n > 0.05N$), adjust the standard deviation of sample means by multiplying it by the *finite population correction factor*:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \underbrace{\sqrt{\frac{N-n}{N-1}}}_{\text{finite population correction factor}}$$

Recap

In this section we have discussed:

- **Central limit theorem.**
- **Practical rules.**
- **Effects of sample sizes.**
- **Correction for a finite population.**

Section 6-6
Normal as Approximation to Binomial

Key Concept

This section presents a method for using a normal distribution as an approximation to the binomial probability distribution.

If the conditions of *np* ≥ 5 and *nq* ≥ 5 are both satisfied, then probabilities from a binomial probability distribution can be approximated well by using a normal distribution with mean $\mu = np$ and standard deviation $\sigma = \sqrt{npq}$.

Review

Binomial Probability Distribution

1. The procedure must have a fixed number of trials.
2. The trials must be independent.
3. Each trial must have all outcomes classified into two categories (commonly, success and failure).
4. The probability of success remains the same in all trials.

Solve by binomial probability formula, Table A-1, or technology.

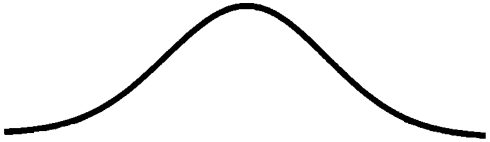
Approximation of a Binomial Distribution with a Normal Distribution

$np \geq 5$

$nq \geq 5$

then $\mu = np$ and $\sigma = \sqrt{npq}$

and the random variable has

a  distribution.

(normal)

Procedure for Using a Normal Distribution to Approximate a Binomial Distribution

1. **Verify that both $np \geq 5$ and $nq \geq 5$. If not, you must use software, a calculator, a table or calculations using the binomial probability formula.**

2. **Find the values of the parameters μ and σ by calculating $\mu = np$ and $\sigma = \sqrt{npq}$.**

3. **Identify the discrete whole number x that is relevant to the binomial probability problem. Focus on this value temporarily.**

Procedure for Using a Normal Distribution to Approximate a Binomial Distribution

4. **Draw a normal distribution centered about μ, then draw a *vertical strip area* centered over x. Mark the left side of the strip with the number equal to x – 0.5, and mark the right side with the number equal to x + 0.5. *Consider the entire area of the entire strip to represent the probability of the discrete whole number itself.***

5. **Determine whether the value of x itself is included in the probability. Determine whether you want the probability of at least x, at most x, more than x, fewer than x, or exactly x. Shade the area to the right or left of the strip; also shade the interior of the strip *if and only if x itself* is to be included. This total shaded region corresponds to the probability being sought.**

Procedure for Using a Normal Distribution to Approximate a Binomial Distribution

6. **Using x – 0.5 or x + 0.5 in place of x, find the area of the shaded region: find the z score; use that z score to find the area to the left of the adjusted value of x; use that cumulative area to identify the shaded area corresponding to the desired probability.**

Example – Number of Men Among Passengers

Finding the Probability of "At Least 122 Men" Among 213 Passengers

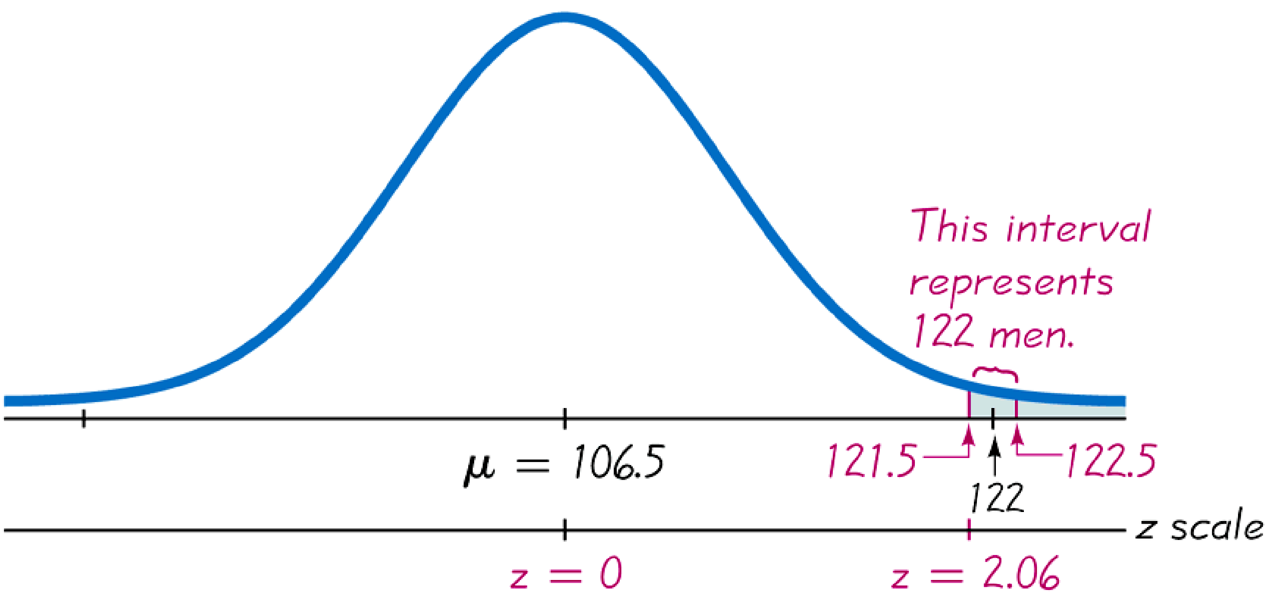


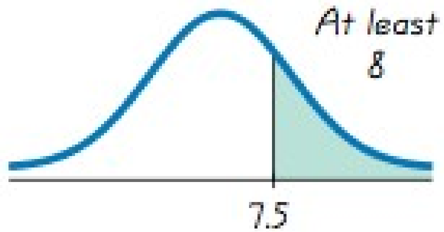
Figure 6-21

Definition

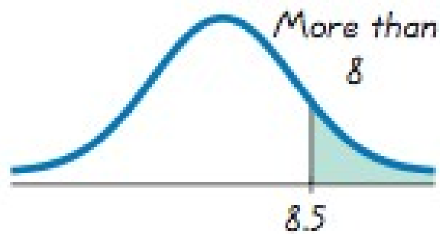
When we use the normal distribution (which is a continuous probability distribution) as an approximation to the binomial distribution (which is discrete), a continuity correction is made to a discrete whole number x in the binomial distribution by representing the discrete whole number x by the interval from

$$x - 0.5 \text{ to } x + 0.5$$

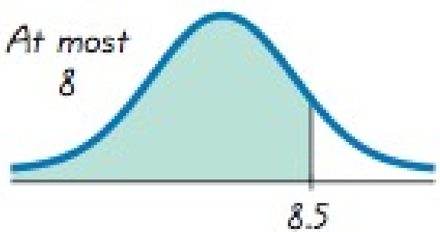
(that is, adding and subtracting 0.5).



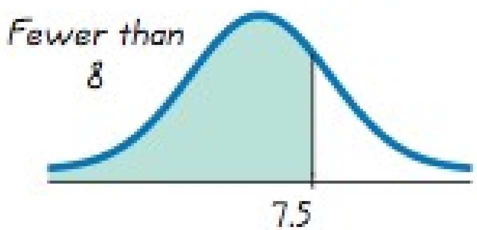
X = <u>**at least**</u> **8**
(includes 8 and above)



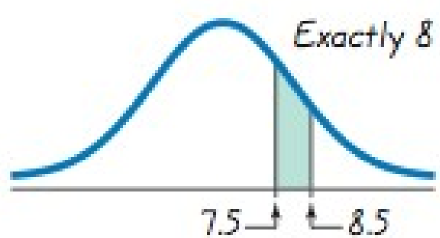
X = <u>**more than**</u> **8**
(doesn’t include 8)



X = <u>**at most**</u> **8**
(includes 8 and below)



X = <u>**fewer than**</u> **8**
(doesn’t include 8)



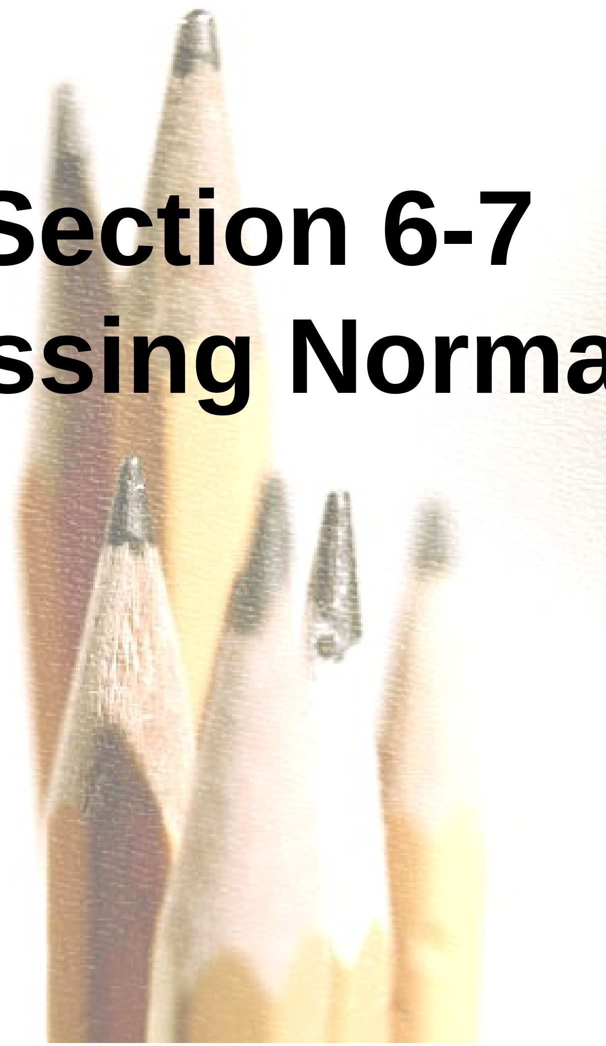
X = <u>**exactly**</u> **8**

Recap

In this section we have discussed:

- **Approximating a binomial distribution with a normal distribution.**
- **Procedures for using a normal distribution to approximate a binomial distribution.**
- **Continuity corrections.**

Section 6-7
Assessing Normality



Key Concept

This section presents criteria for determining whether the requirement of a normal distribution is satisfied.

The criteria involve visual inspection of a histogram to see if it is roughly bell shaped, identifying any outliers, and constructing a graph called a normal quantile plot.

Definition

A **normal quantile plot** (or **normal probability plot**) is a graph of points (x,y), where each x value is from the original set of sample data, and each y value is the corresponding z score that is a quantile value expected from the standard normal distribution.

Procedure for Determining Whether It Is Reasonable to Assume that Sample Data are From a Normally Distributed Population

1. **Histogram:** Construct a histogram. Reject normality if the histogram departs dramatically from a bell shape.

2. **Outliers:** Identify outliers. Reject normality if there is more than one outlier present.

3. **Normal Quantile Plot:** If the histogram is basically symmetric and there is at most one outlier, use technology to generate a **normal quantile plot**.

Procedure for Determining Whether It Is Reasonable to Assume that Sample Data are From a Normally Distributed Population

3. Continued

Use the following criteria to determine whether or not the distribution is normal.

Normal Distribution: The population distribution is normal if the pattern of the points is reasonably close to a straight line and the points do not show some systematic pattern that is not a straight-line pattern.

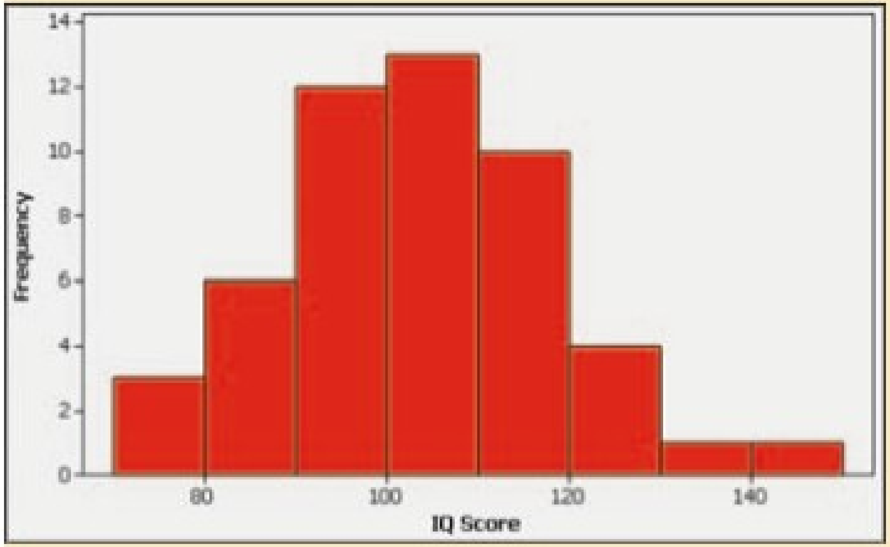
Procedure for Determining Whether It Is Reasonable to Assume that Sample Data are From a Normally Distributed Population

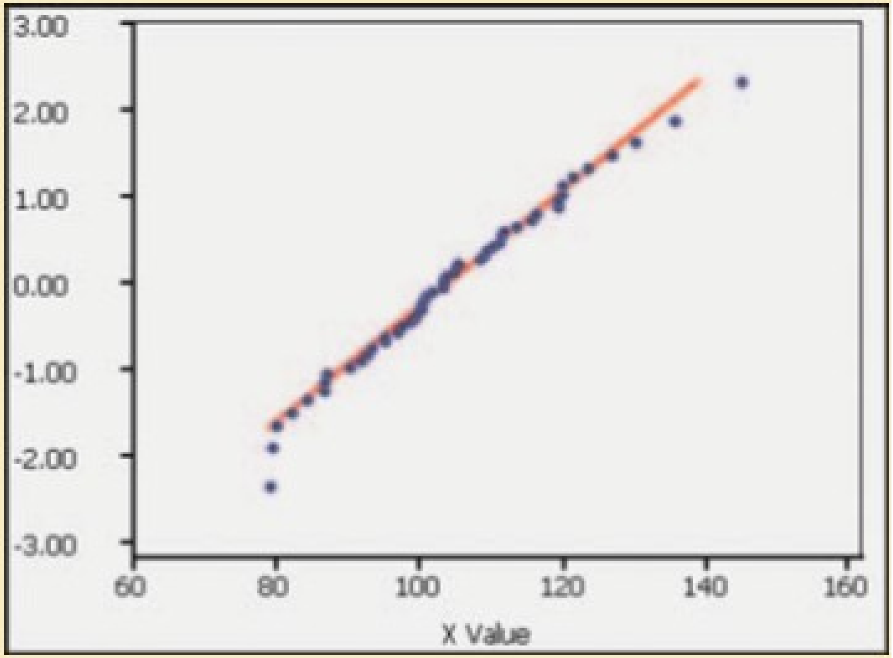
3. Continued

Not a Normal Distribution: **The population distribution is not normal if either or both of these two conditions applies:**

- **The points do not lie reasonably close to a straight line.**
- **The points show some systematic pattern that is not a straight-line pattern.**

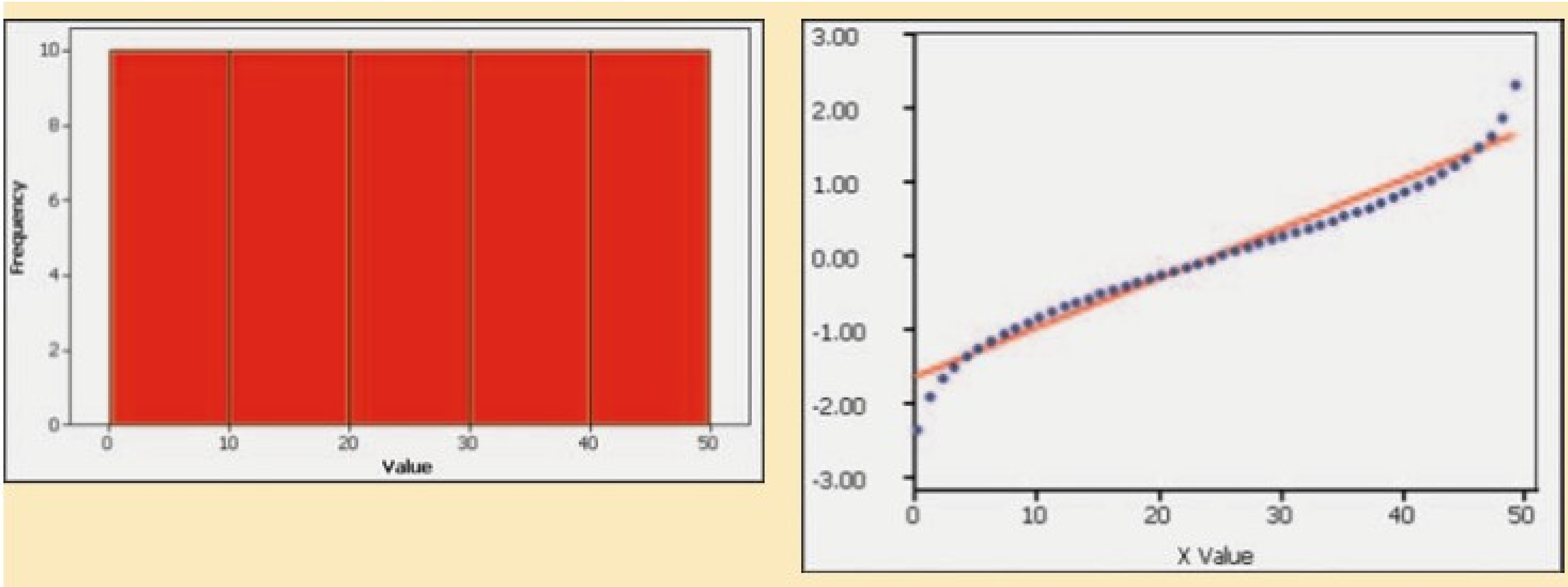
Example





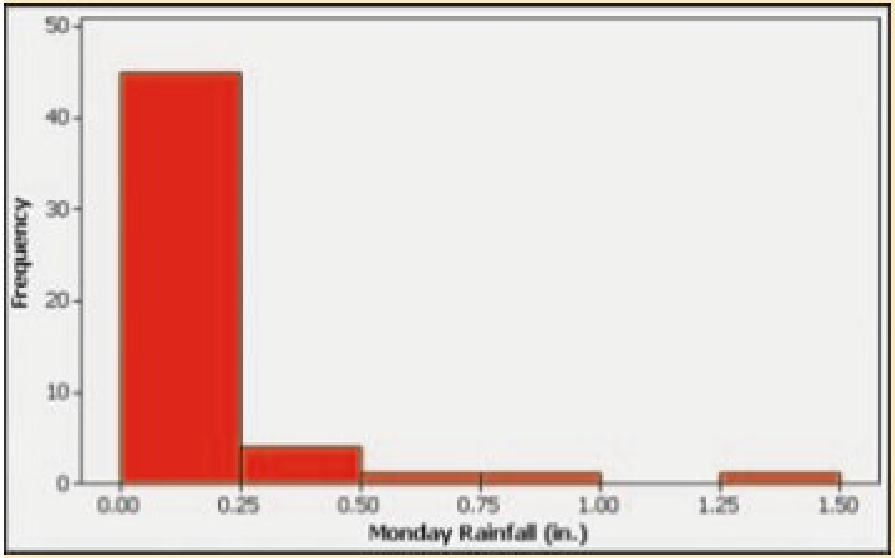
Normal: **Histogram of IQ scores is close to being bell-shaped, suggests that the IQ scores are from a normal distribution. The normal quantile plot shows points that are reasonably close to a straight-line pattern. It is safe to assume that these IQ scores are from a normally distributed population.**

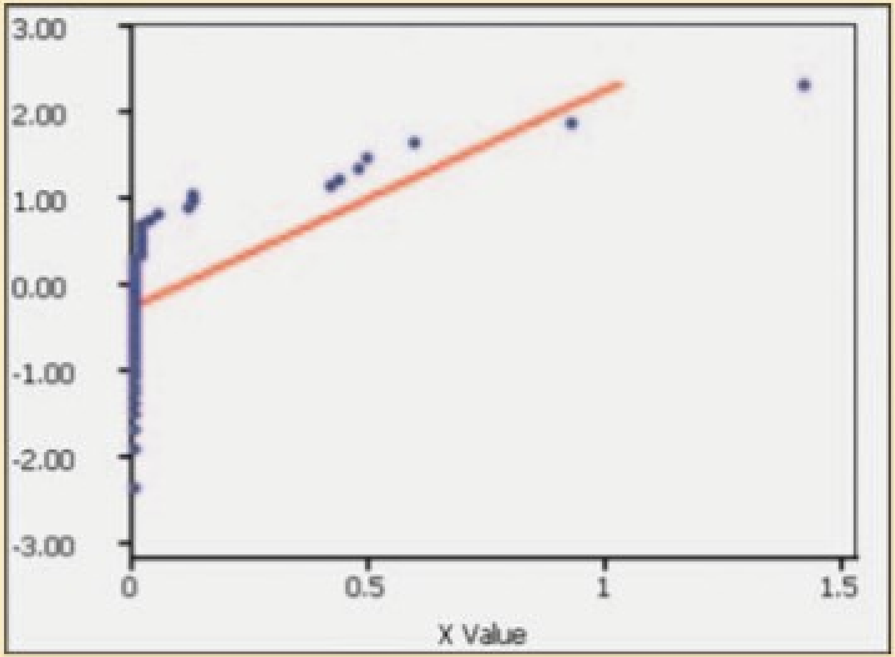
Example



Uniform: **Histogram of data having a uniform distribution. The corresponding normal quantile plot suggests that the points are not normally distributed because the points show a systematic pattern that is not a straight-line pattern. These sample values are not from a population having a normal distribution.**

Example





Skewed: **Histogram of the amounts of rainfall in Boston for every Monday during one year. The shape of the histogram is skewed, not bell-shaped. The corresponding normal quantile plot shows points that are not at all close to a straight-line pattern. These rainfall amounts are not from a population having a normal distribution.**

Manual Construction of a Normal Quantile Plot

Step 1. First sort the data by arranging the values in order from lowest to highest.

Step 2. With a sample of size *n*, each value represents a proportion of 1/*n* of the sample. Using the known sample size *n*, identify the areas of 1/2*n*, 3/2*n*, and so on. These are the cumulative areas to the left of the corresponding sample values.

Step 3. Use the standard normal distribution (Table A-2 or software or a calculator) to find the *z* scores corresponding to the cumulative left areas found in Step 2. (These are the *z* scores that are expected from a normally distributed sample.)

Manual Construction of a Normal Quantile Plot

Step 4. Match the original sorted data values with their corresponding *z* scores found in Step 3, then plot the points (*x*, *y*), where each *x* is an original sample value and *y* is the corresponding *z* score.

Step 5. Examine the normal quantile plot and determine whether or not the distribution is normal.

Ryan-Joiner Test

The Ryan-Joiner test is one of several formal tests of normality, each having their own advantages and disadvantages. STATDISK has a feature of Normality Assessment that displays a histogram, normal quantile plot, the number of potential outliers, and results from the Ryan-Joiner test. Information about the Ryan-Joiner test is readily available on the Internet.

Data Transformations

Many data sets have a distribution that is not normal, but we can transform the data so that the modified values have a normal distribution. One common transformation is to replace each value of x with log (*x* + 1). If the distribution of the log (*x* + 1) values is a normal distribution, the distribution of the x values is referred to as a lognormal distribution.

Other Data Transformations

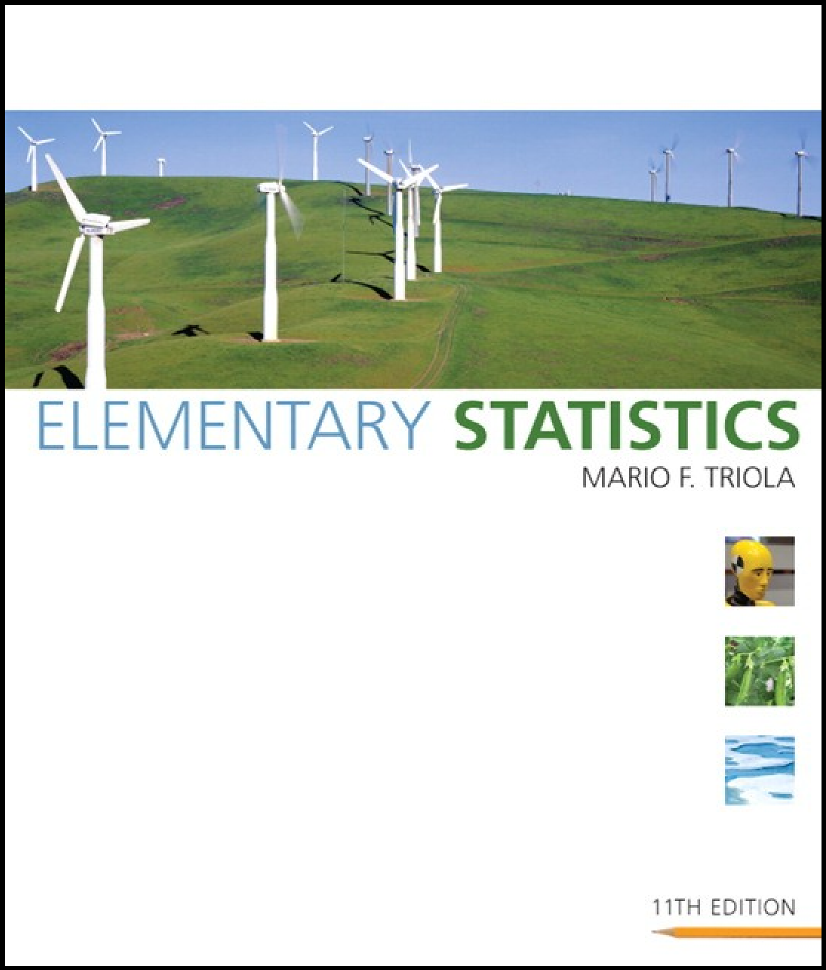
In addition to replacing each x value with the log (x + 1), there are other transformations, such as replacing each x value with $\sqrt{x}$, or 1/x, or x^2. In addition to getting a required normal distribution when the original data values are not normally distributed, such transformations can be used to correct other deficiencies, such as a requirement (found in later chapters) that different data sets have the same variance.

Recap

In this section we have discussed:

- **Normal quantile plot.**
- **Procedure to determine if data have a normal distribution.**

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

Chapter 7
Estimates and Sample Sizes

7-1 Review and Preview

7-2 Estimating a Population Proportion

7-3 Estimating a Population Mean: σ Known

7-4 Estimating a Population Mean: σ Not Known

7-5 Estimating a Population Variance

Section 7-1
Review and Preview

Review

- **Chapters 2 & 3 we used "descriptive statistics" when we summarized data using tools such as graphs, and statistics such as the mean and standard deviation.**
- **Chapter 6 we introduced critical values:**
  z_α denotes the z score with an area of α to its right.
  If $\alpha = 0.025$, the critical value is $z_{0.025} = 1.96$.
  That is, the critical value $z_{0.025} = 1.96$ has an area of 0.025 to its right.

Preview

This chapter presents the beginning of inferential statistics.

- **The two major activities of inferential statistics are (1) to use sample data to estimate values of a population parameters, and (2) to test hypotheses or claims made about population parameters.**
- **We introduce methods for estimating values of these important population parameters: proportions, means, and variances.**
- **We also present methods for determining sample sizes necessary to estimate those parameters.**

Section 7-2 Estimating a Population Proportion

Key Concept

In this section we present methods for using a sample proportion to estimate the value of a population proportion.

- **The sample proportion is the best point estimate of the population proportion.**
- **We can use a sample proportion to construct a confidence interval to estimate the true value of a population proportion, and we should know how to interpret such confidence intervals.**
- **We should know how to find the sample size necessary to estimate a population proportion.**

Definition

A **point estimate** is a single value (or point) used to approximate a population parameter.

Definition

The sample proportion $\hat{p}$ is the best point estimate of the population proportion p.

Example:

In the Chapter Problem we noted that in a Pew Research Center poll, 70% of 1501 randomly selected adults in the United States believe in global warming, so the sample proportion is $\hat{p}$ = 0.70. Find the best point estimate of the proportion of all adults in the United States who believe in global warming.

Because the sample proportion is the best point estimate of the population proportion, we conclude that the best point estimate of *p* is 0.70. When using the sample results to estimate the percentage of all adults in the United States who believe in global warming, the best estimate is 70%.

Definition

A **confidence interval** (or **interval estimate**) is a range (or an interval) of values used to estimate the true value of a population parameter. A confidence interval is sometimes abbreviated as CI.

Definition

A confidence level is the probability $1 - \alpha$ (often expressed as the equivalent percentage value) that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times. (The confidence level is also called degree of confidence, or the confidence coefficient.)

Most common choices are 90%, 95%, or 99%.

($\alpha = 10\%$), ($\alpha = 5\%$), ($\alpha = 1\%$)

Interpreting a Confidence Interval

We must be careful to interpret confidence intervals correctly. There is a correct interpretation and many different and creative incorrect interpretations of the confidence interval $0.677 < p < 0.723$.

"We are 95% confident that the interval from 0.677 to 0.723 actually does contain the true value of the population proportion p."

This means that if we were to select many different samples of size 1501 and construct the corresponding confidence intervals, 95% of them would actually contain the value of the population proportion p.

(Note that in this correct interpretation, the level of 95% refers to the success rate of the process being used to estimate the proportion.)

Caution

Know the correct interpretation of a confidence interval.

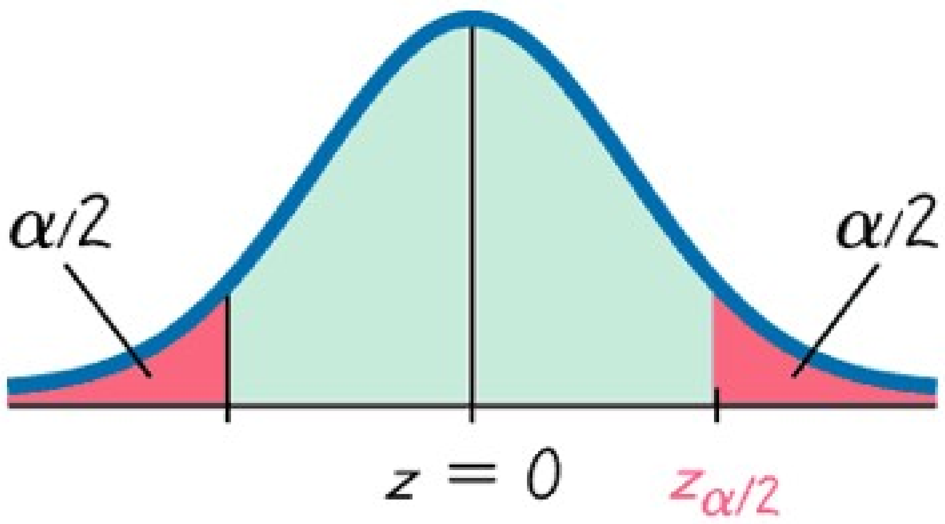
Caution

Confidence intervals can be used informally to compare different data sets, *but the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of proportions*.

Critical Values

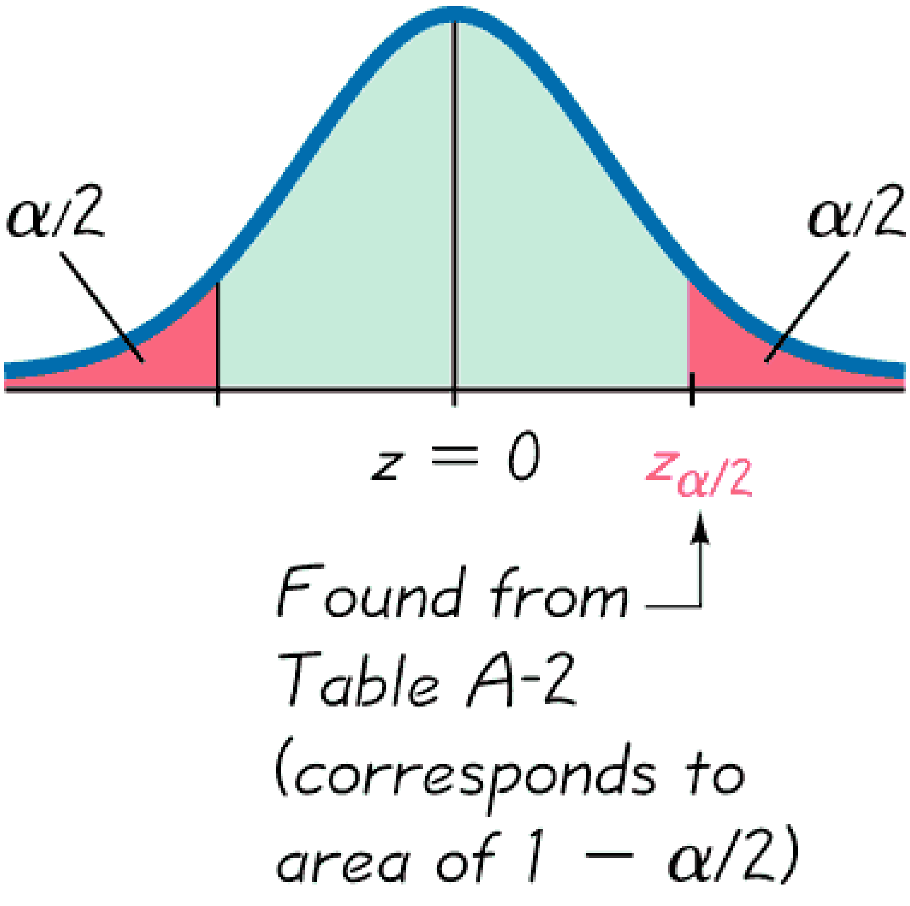
A standard *z* score can be used to distinguish between sample statistics that are likely to occur and those that are unlikely to occur. Such a *z* score is called a critical value. Critical values are based on the following observations:

1. **Under certain conditions, the sampling distribution of sample proportions can be approximated by a normal distribution.**



Critical Values

2. A *z* score associated with a sample proportion has a probability of $\alpha/2$ of falling in the right tail.



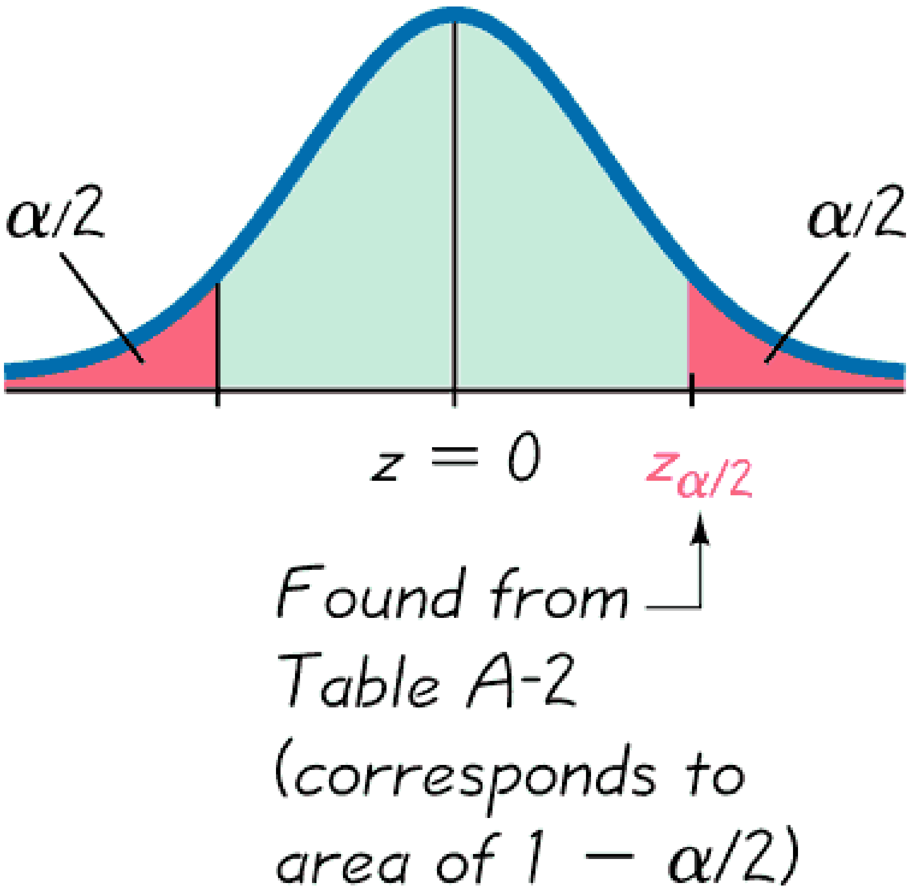
Critical Values

3. **The z score separating the right-tail region is commonly denoted by $z_{\alpha/2}$ and is referred to as a critical value because it is on the borderline separating z scores from sample proportions that are likely to occur from those that are unlikely to occur.**

Definition

A **critical value** is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely to occur. The number $z_{\alpha/2}$ is a critical value that is a z score with the property that it separates an area of $\alpha/2$ in the right tail of the standard normal distribution.

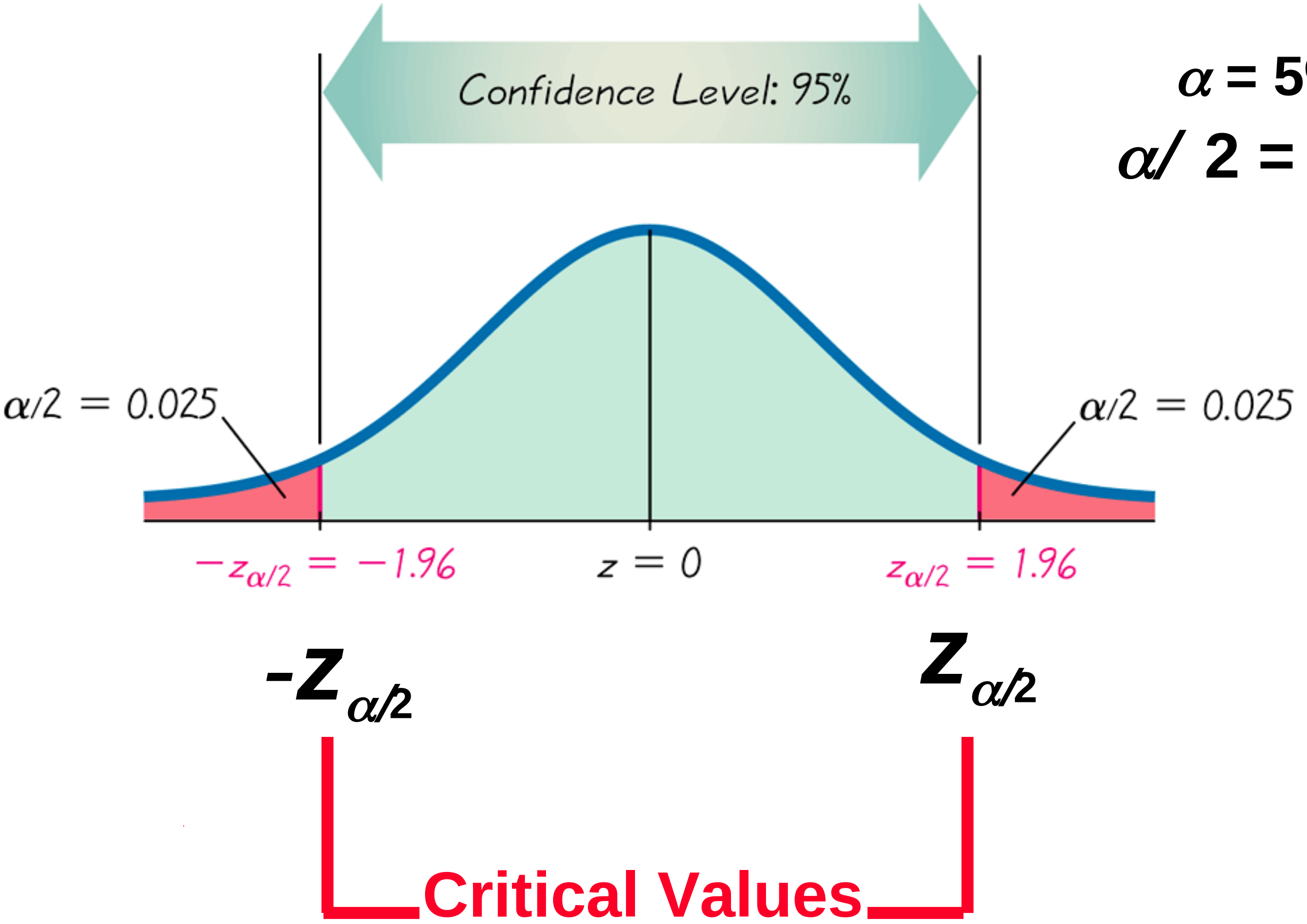
The Critical Value $z_{\alpha/2}$



Notation for Critical Value

The critical value $z_{\alpha/2}$ is the positive z value that is at the vertical boundary separating an area of $\alpha/2$ in the right tail of the standard normal distribution. (The value of $-z_{\alpha/2}$ is at the vertical boundary for the area of $\alpha/2$ in the left tail.) The subscript $\alpha/2$ is simply a reminder that the z score separates an area of $\alpha/2$ in the right tail of the standard normal distribution.

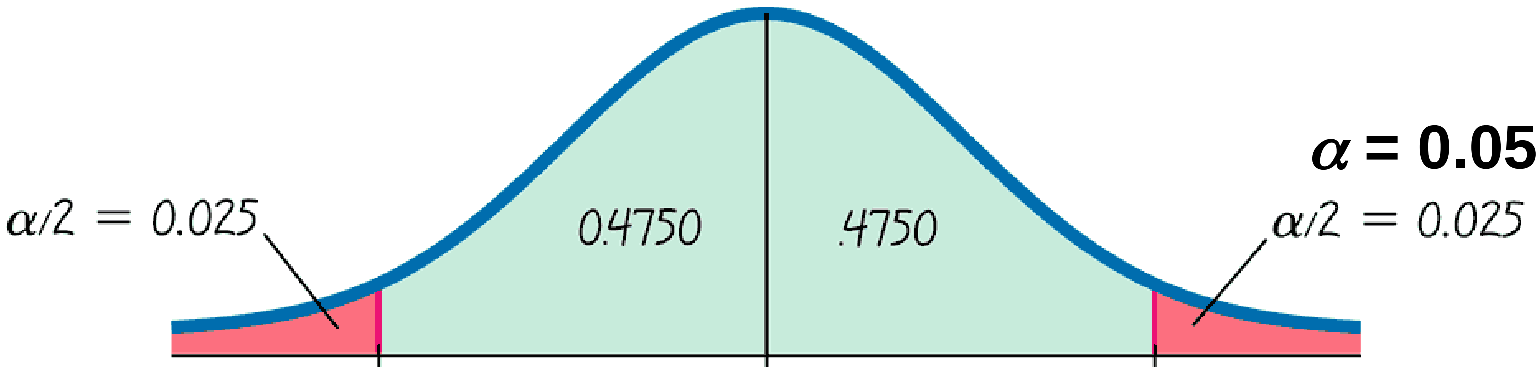
Finding $z_{\alpha/2}$ for a 95% Confidence Level



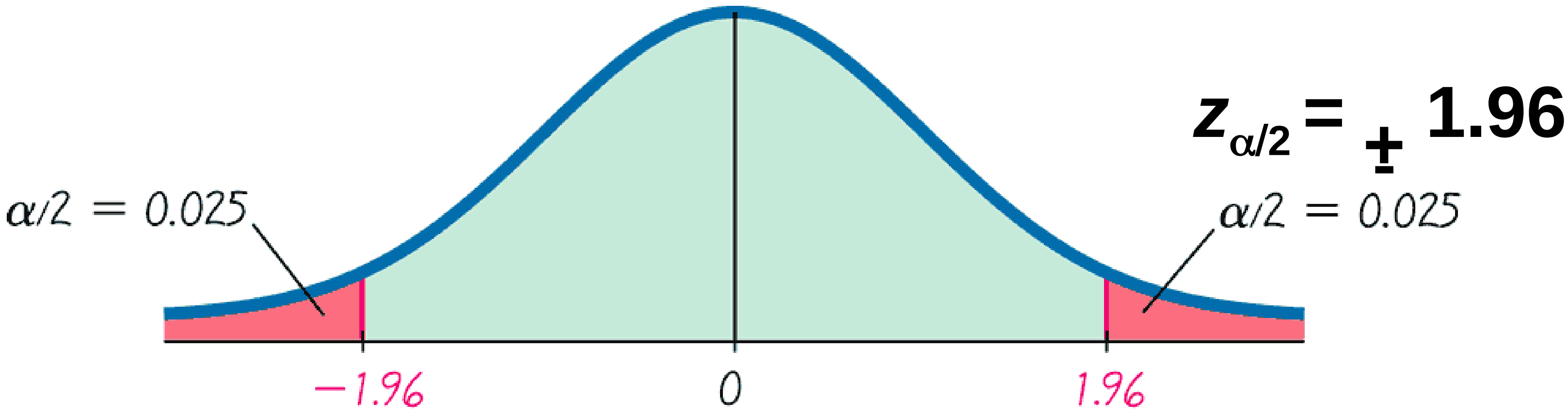
α = 5%

$\alpha/$ 2 = 2.5% = .025

Finding $z_{\alpha/2}$ for a 95% Confidence Level - cont



Use Table A-2 to find a *z* score of 1.96



Definition

When data from a simple random sample are used to estimate a population proportion *p*, the margin of error, denoted by *E*, is the maximum likely difference (with probability 1 – α, such as 0.95) between the observed proportion $\hat{p}$ and the true value of the population proportion *p*. The margin of error *E* is also called the maximum error of the estimate and can be found by multiplying the critical value and the standard deviation of the sample proportions:

Margin of Error for Proportions

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Confidence Interval for Estimating a Population Proportion p

p = population proportion

$\hat{p}$ = sample proportion

n = number of sample values

E = margin of error

$z_{\alpha/2}$ = z score separating an area of $\alpha/2$ in the right tail of the standard normal distribution

Confidence Interval for Estimating a Population Proportion *p*

1. The sample is a simple random sample.

2. The conditions for the binomial distribution are satisfied: there is a fixed number of trials, the trials are independent, there are two categories of outcomes, and the probabilities remain constant for each trial.

3. There are at least 5 successes and 5 failures.

Confidence Interval for Estimating a Population Proportion *p*

$$\hat{p} - E < p < \hat{p} + E$$

where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Confidence Interval for Estimating a Population Proportion *p*

$$\hat{p} - E < p < \hat{p} + E$$

$$\hat{p} \pm E$$

$$(\hat{p} - E, \hat{p} + E)$$

Round-Off Rule for Confidence Interval Estimates of *p*

Round the confidence interval limits for *p* to

three significant digits.

Procedure for Constructing a Confidence Interval for *p*

1. **Verify that the required assumptions are satisfied. (The sample is a simple random sample, the conditions for the binomial distribution are satisfied, and the normal distribution can be used to approximate the distribution of sample proportions because $np \geq 5$, and $nq \geq 5$ are both satisfied.)**

2. **Refer to Table A-2 and find the critical value $z_{\alpha/2}$ that corresponds to the desired confidence level.**

3. **Evaluate the margin of error** $E = z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$

Procedure for Constructing a Confidence Interval for *p* - cont

4. Using the value of the calculated margin of error, E and the value of the sample proportion, $\hat{p}$, find the values of $\hat{p} - E$ and $\hat{p} + E$. Substitute those values in the general format for the confidence interval:

$$\hat{p} - E < p < \hat{p} + E$$

5. Round the resulting confidence interval limits to three significant digits.

Example:

In the Chapter Problem we noted that a Pew Research Center poll of 1501 randomly selected U.S. adults showed that 70% of the respondents believe in global warming. The sample results are *n* = 1501, and $\hat{p} = 0.70$

a. **Find the margin of error E that corresponds to a 95% confidence level.**

b. **Find the 95% confidence interval estimate of the population proportion *p*.**

c. **Based on the results, can we safely conclude that the majority of adults believe in global warming?**

d. **Assuming that you are a newspaper reporter, write a brief statement that accurately describes the results and includes all of the relevant information.**

Example:

Requirement check: simple random sample; fixed number of trials, 1501; trials are independent; two categories of outcomes (believes or does not); probability remains constant. Note: number of successes and failures are both at least 5.

a) Use the formula to find the margin of error.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{(0.70)(0.30)}{1501}}$$

$$E = 0.023183$$

Example:

b) The 95% confidence interval:

$$\hat{p} - E < p < \hat{p} + E$$

$$0.70 - 0.023183 < p < 0.70 + 0.023183$$

$$0.677 < p < 0.723$$

Example:

c) **Based on the confidence interval obtained in part (b), it does appear that the proportion of adults who believe in global warming is greater than 0.5 (or 50%), so we can safely conclude that the majority of adults believe in global warming. Because the limits of 0.677 and 0.723 are likely to contain the true population proportion, it appears that the population proportion is a value greater than 0.5.**

Example:

d) Here is one statement that summarizes the results: 70% of United States adults believe that the earth is getting warmer. That percentage is based on a Pew Research Center poll of 1501 randomly selected adults in the United States. In theory, in 95% of such polls, the percentage should differ by no more than 2.3 percentage points in either direction from the percentage that would be found by interviewing all adults in the United States.

Analyzing Polls

When analyzing polls consider:

1. **The sample should be a simple random sample, not an inappropriate sample (such as a voluntary response sample).**
2. **The confidence level should be provided. (It is often 95%, but media reports often neglect to identify it.)**
3. **The sample size should be provided. (It is usually provided by the media, but not always.)**
4. **Except for relatively rare cases, the quality of the poll results depends on the sampling method and the size of the sample, but the size of the population is usually not a factor.**

Caution

Never follow the common misconception that poll results are unreliable if the sample size is a small percentage of the population size. The population size is usually not a factor in determining the reliability of a poll.

Sample Size

Suppose we want to collect sample data in order to estimate some population proportion. The question is how many sample items must be obtained?

Determining Sample Size

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

(solve for *n* by algebra)

$$n = \frac{(z_{\alpha/2})^2\,\hat{p}\,\hat{q}}{E^2}$$

Sample Size for Estimating Proportion *p*

When an estimate of $\hat{p}$ is known:

$$n = \frac{\left(z_{\alpha/2}\right)^2 \hat{p}\,\hat{q}}{E^2}$$

When no estimate of $\hat{p}$ is known:

$$n = \frac{\left(z_{\alpha/2}\right)^2 0.25}{E^2}$$

Round-Off Rule for Determining Sample Size

If the computed sample size *n* is not a whole number, round the value of *n* up to the next larger whole number.

Example:

The Internet is affecting us all in many different ways, so there are many reasons for estimating the proportion of adults who use it. Assume that a manager for E-Bay wants to determine the current percentage of U.S. adults who now use the Internet. How many adults must be surveyed in order to be 95% confident that the sample percentage is in error by no more than three percentage points?

a. In 2006, 73% of adults used the Internet.

b. No known possible value of the proportion.

Example:

a) Use $\hat{p} = 0.73$ and $\hat{q} = 1 - \hat{p} = 0.27$

$\alpha = 0.05$ so $z_{\alpha/2} = 1.96$

$E = 0.03$

$$n = \frac{\left(z_{\alpha/2}\right)^2 \cdot \hat{p}\hat{q}}{E^2}$$

$$= \frac{(1.96)^2 \cdot (0.73)(0.27)}{(0.03)^2}$$

$$= 841.3104$$

$$= 842$$

To be 95% confident that our sample percentage is within three percentage points of the true percentage for all adults, we should obtain a simple random sample of 842 adults.

Example:

b) Use $\alpha = 0.05$ so $z_{\alpha/2} = 1.96$

$E = 0.03$

$$n = \frac{\left(z_{\alpha/2}\right)^2 \cdot 0.25}{E^2}$$

$$= \frac{(1.96)^2 \cdot 0.25}{(0.03)^2}$$

$$= 1067.1111$$

$$= 1068$$

To be 95% confident that our sample percentage is within three percentage points of the true percentage for all adults, we should obtain a simple random sample of 1068 adults.

Finding the Point Estimate and *E* from a Confidence Interval

Point estimate of $\hat{p}$:

$$\hat{p} = \frac{\text{(upper confidence limit)} + \text{(lower confidence limit)}}{2}$$

Margin of Error:

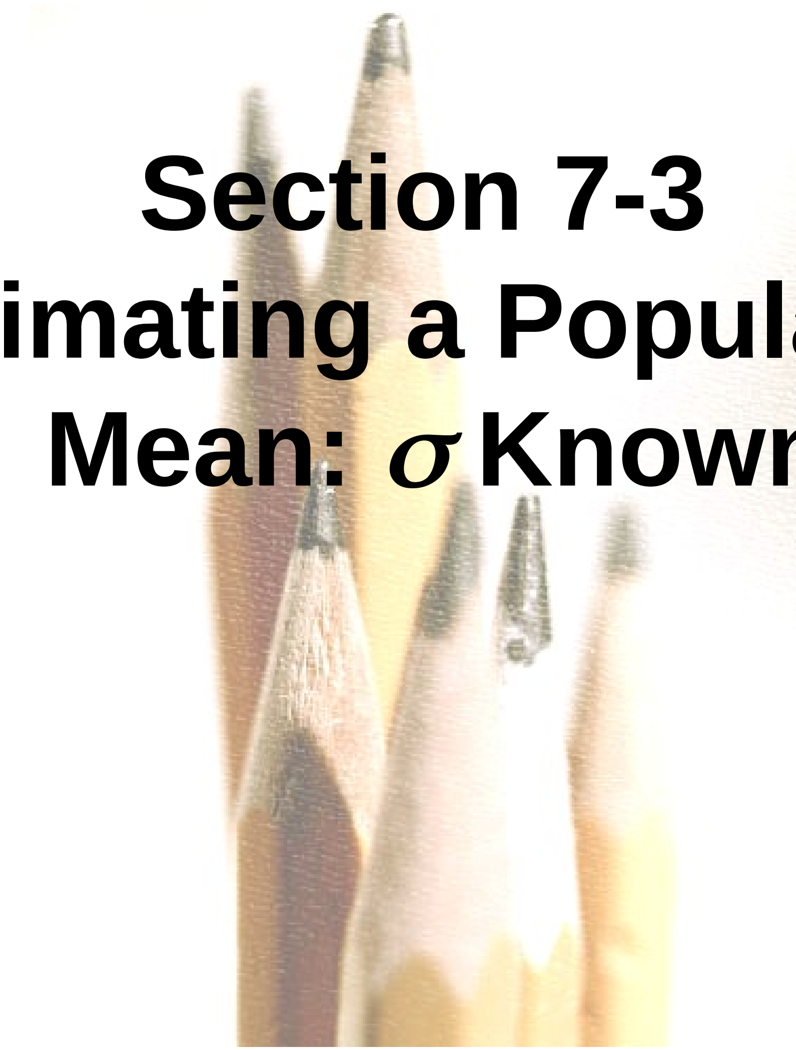
$$E = \frac{\text{(upper confidence limit)} - \text{(lower confidence limit)}}{2}$$

Recap

In this section we have discussed:

- Point estimates.
- Confidence intervals.
- Confidence levels.
- Critical values.
- Margin of error.
- Determining sample sizes.

Section 7-3
Estimating a Population Mean: σ Known



Key Concept

This section presents methods for estimating a population mean. In addition to knowing the values of the sample data or statistics, we must also know the value of the population standard deviation, σ.

Here are three key concepts that should be learned in this section:

Key Concept

1. **We should know that the sample mean $\bar{x}$ is the best point estimate of the population mean μ.**

2. **We should learn how to use sample data to construct a confidence interval for estimating the value of a population mean, and we should know how to interpret such confidence intervals.**

3. **We should develop the ability to determine the sample size necessary to estimate a population mean.**

Point Estimate of the Population Mean

The sample mean $\bar{x}$ is the best point estimate of the population mean μ.

Confidence Interval for Estimating a Population Mean (with σ Known)

μ = population mean

σ = population standard deviation

$\bar{x}$ = sample mean

n = number of sample values

E = margin of error

$z_{\alpha/2}$ = z score separating an area of $a/2$ in the right tail of the standard normal distribution

Confidence Interval for Estimating a Population Mean (with σ Known)

1. The sample is a simple random sample. (All samples of the same size have an equal chance of being selected.)

2. The value of the population standard deviation σ is known.

3. Either or both of these conditions is satisfied: The population is normally distributed or $n > 30$.

Confidence Interval for Estimating a Population Mean (with σ Known)

$$\bar{x} - E < \mu < \bar{x} + E \quad \text{where} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

or $$\bar{x} \pm E$$

or $$(\bar{x} - E, \bar{x} + E)$$

Definition

The two values $\bar{x} - E$ and $\bar{x} + E$ are called confidence interval limits.

Sample Mean

1. For all populations, the sample mean $\bar{x}$ is an unbiased estimator of the population mean μ, meaning that the distribution of sample means tends to center about the value of the population mean μ.

2. For many populations, the distribution of sample means $\bar{x}$ tends to be more consistent (with less variation) than the distributions of other sample statistics.

Procedure for Constructing a Confidence Interval for μ (with Known σ)

1. Verify that the requirements are satisfied.
2. Refer to Table A-2 or use technology to find the critical value $z_{\alpha/2}$ that corresponds to the desired confidence level.
3. Evaluate the margin of error $E = z_{\alpha/2} \cdot \sigma / \sqrt{n}$
4. Find the values of $\bar{x} - E$ and $\bar{x} + E$. Substitute those values in the general format of the confidence interval:

$$\bar{x} - E < \mu < \bar{x} + E$$

5. Round using the confidence intervals round-off rules.

Round-Off Rule for Confidence Intervals Used to Estimate μ

1. When using the **original set of data**, round the confidence interval limits to one more decimal place than used in original set of data.
2. When the original set of data is unknown and only the **summary statistics** (n, $\bar{x}$, s) are used, round the confidence interval limits to the same number of decimal places used for the sample mean.

Example:

People have died in boat and aircraft accidents because an obsolete estimate of the mean weight of men was used. In recent decades, the mean weight of men has increased considerably, so we need to update our estimate of that mean so that boats, aircraft, elevators, and other such devices do not become dangerously overloaded. Using the weights of men from Data Set 1 in Appendix B, we obtain these sample statistics for the simple random sample: n = 40 and $\bar{x}$ = 172.55 lb. Research from several other sources suggests that the population of weights of men has a standard deviation given by σ = 26 lb.

Example:

a. **Find the best point estimate of the mean weight of the population of all men.**

b. **Construct a 95% confidence interval estimate of the mean weight of all men.**

c. **What do the results suggest about the mean weight of 166.3 lb that was used to determine the safe passenger capacity of water vessels in 1960 (as given in the National Transportation and Safety Board safety recommendation M-04-04)?**

Example:

a. **The sample mean of 172.55 lb is the best point estimate of the mean weight of the population of all men.**

b. **A 95% confidence interval or 0.95 implies σ = 0.05, so $z_{\alpha/2}$ = 1.96.**
Calculate the margin of error.

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{26}{\sqrt{40}} = 8.0574835$$

Construct the confidence interval.

$$\bar{x} - E < \mu < \bar{x} - E$$

$$172.55 - 8.0574835 < \mu < 172.55 + 8.0574835$$

$$164.49 < \mu < 180.61$$

Example:

c. Based on the confidence interval, it is possible that the mean weight of 166.3 lb used in 1960 could be the mean weight of men today. However, the best point estimate of 172.55 lb suggests that the mean weight of men is now considerably greater than 166.3 lb. Considering that an underestimate of the mean weight of men could result in lives lost through overloaded boats and aircraft, these results strongly suggest that additional data should be collected. (Additional data have been collected, and the assumed mean weight of men has been increased.)

Finding a Sample Size for Estimating a Population Mean

μ = population mean

σ = population standard deviation

$\bar{x}$ = population standard deviation

E = desired margin of error

$z_{\alpha/2}$ = z score separating an area of $\alpha/2$ in the right tail of the standard normal distribution

$$n = \left[\frac{(z_{\alpha/2}) \cdot \sigma}{E}\right]^2$$

Round-Off Rule for Sample Size *n*

If the computed sample size *n* is not a whole number, round the value of *n* up to the next larger whole number.

Finding the Sample Size n When σ is Unknown

1. Use the range rule of thumb (see Section 3-3) to estimate the standard deviation as follows: $\sigma \approx$ range/4.

2. Start the sample collection process without knowing σ and, using the first several values, calculate the sample standard deviation s and use it in place of σ. The estimated value of σ can then be improved as more sample data are obtained, and the sample size can be refined accordingly.

3. Estimate the value of σ by using the results of some other study that was done earlier.

Example:

Assume that we want to estimate the mean IQ score for the population of statistics students. How many statistics students must be randomly selected for IQ tests if we want 95% confidence that the sample mean is within 3 IQ points of the population mean?

$\alpha = 0.05$

$\alpha/2 = 0.025$

$z_{\alpha/2} = 1.96$

$E = 3$

$\sigma = 15$

$$n = \left[\frac{1.96 \cdot 15}{3}\right]^2 = 96.04 = 97$$

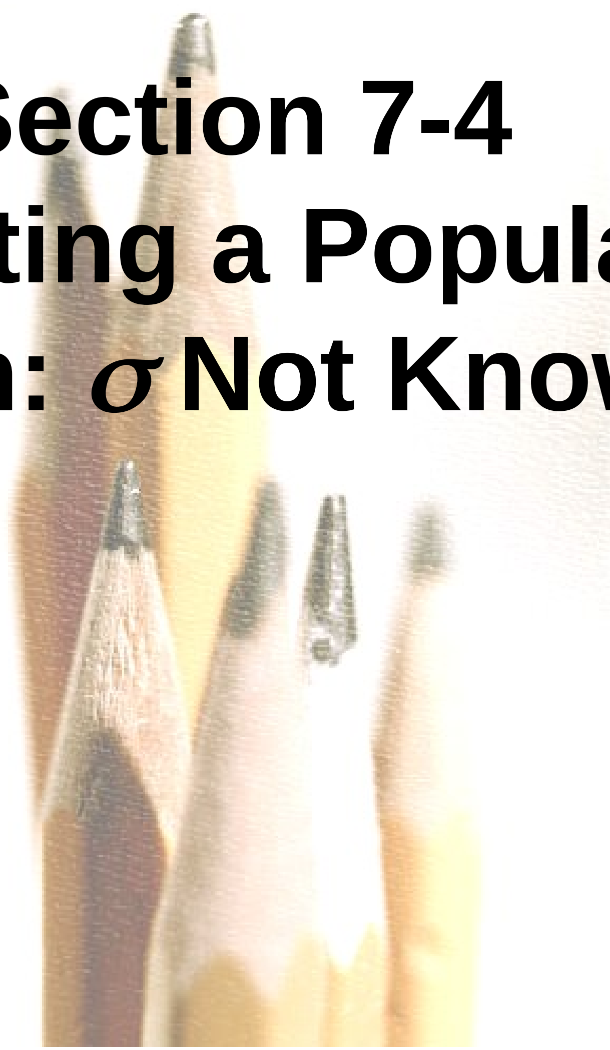
With a simple random sample of only 97 statistics students, we will be 95% confident that the sample mean is within 3 IQ points of the true population mean μ.

Recap

In this section we have discussed:

- **Margin of error.**
- **Confidence interval estimate of the population mean with σ known.**
- **Round off rules.**
- **Sample size for estimating the mean μ.**

Section 7-4
Estimating a Population Mean: σ Not Known



Key Concept

This section presents methods for estimating a population mean when the population standard deviation is not known. With σ unknown, we use the Student t distribution assuming that the relevant requirements are satisfied.

Sample Mean

The sample mean is the best point estimate of the population mean.

Student *t* Distribution

If the distribution of a population is essentially normal, then the distribution of

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

is a Student *t* Distribution for all samples of size *n*. It is often referred to as a *t* distribution and is used to find critical values denoted by $t_{\alpha/2}$.

Definition

The number of **degrees of freedom** for a collection of sample data is the number of sample values that can vary after certain restrictions have been imposed on all data values. The degree of freedom is often abbreviated **df**.

degrees of freedom = $n - 1$

in this section.

Margin of Error E for Estimate of μ (With σ Not Known)

Formula 7-6

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

where $t_{\alpha/2}$ has $n - 1$ degrees of freedom.

Table A-3 lists values for $t_{\alpha/2}$

Notation

μ = population mean

$\bar{x}$ = sample mean

s = sample standard deviation

n = number of sample values

E = margin of error

$t_{\alpha/2}$ = critical t value separating an area of $\alpha/2$ in the right tail of the t distribution

Confidence Interval for the Estimate of μ (With σ Not Known)

$$\bar{x} - E < \mu < \bar{x} + E$$

where $$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$ df = $n - 1$

$t_{\alpha/2}$ found in Table A-3

Procedure for Constructing a Confidence Interval for μ (With σ Unknown)

1. Verify that the requirements are satisfied.
2. Using $n - 1$ degrees of freedom, refer to Table A-3 or use technology to find the critical value $t_{\alpha/2}$ that corresponds to the desired confidence level.
3. Evaluate the margin of error $E = t_{\alpha/2} \cdot s / \sqrt{n}$.
4. Find the values of $\bar{x} - E$ and $\bar{x} + E$. Substitute those values in the general format for the confidence interval:

$$\bar{x} - E < \mu < \bar{x} + E$$

5. Round the resulting confidence interval limits.

Example:

A common claim is that garlic lowers cholesterol levels. In a test of the effectiveness of garlic, 49 subjects were treated with doses of raw garlic, and their cholesterol levels were measured before and after the treatment. The changes in their levels of LDL cholesterol (in mg/dL) have a mean of 0.4 and a standard deviation of 21.0. Use the sample statistics of n = 49, $\overline{x}$ = 0.4 and s = 21.0 to construct a 95% confidence interval estimate of the mean net change in LDL cholesterol after the garlic treatment. What does the confidence interval suggest about the effectiveness of garlic in reducing LDL cholesterol?

Example:

Requirements are satisfied: simple random sample and *n* = 49 (i.e., *n* > 30).

95% implies *a* = 0.05.
With *n* = 49, the df = 49 – 1 = 48
Closest df is 50, two tails, so $t_{\alpha/2}$ = 2.009

Using $t_{\alpha/2}$ = 2.009, *s* = 21.0 and *n* = 49 the margin of error is:

$$E = t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 2.009 \cdot \frac{21.0}{\sqrt{49}} = 6.027$$

Example:

Construct the confidence interval: $\bar{x} = 0.4, E = 6.027$

$$\bar{x} - E < \mu < \bar{x} + E$$

$$0.4 - 6.027 < \mu < 0.4 + 6.027$$

$$-5.6 < \mu < 6.4$$

We are 95% confident that the limits of –5.6 and 6.4 actually do contain the value of μ, the mean of the changes in LDL cholesterol for the population. Because the confidence interval limits contain the value of 0, it is very possible that the mean of the changes in LDL cholesterol is equal to 0, suggesting that the garlic treatment did not affect the LDL cholesterol levels. It does not appear that the garlic treatment is effective in lowering LDL cholesterol.

Important Properties of the Student *t* Distribution

1. The Student *t* distribution is different for different sample sizes (see the following slide, for the cases $n = 3$ and $n = 12$).

2. The Student *t* distribution has the same general symmetric bell shape as the standard normal distribution but it reflects the greater variability (with wider distributions) that is expected with small samples.

3. The Student *t* distribution has a mean of $t = 0$ (just as the standard normal distribution has a mean of $z = 0$).

4. The standard deviation of the Student *t* distribution varies with the sample size and is greater than 1 (unlike the standard normal distribution, which has a $\sigma = 1$).

5. As the sample size *n* gets larger, the Student *t* distribution gets closer to the normal distribution.

Student *t* Distributions for *n* = 3 and *n* = 12

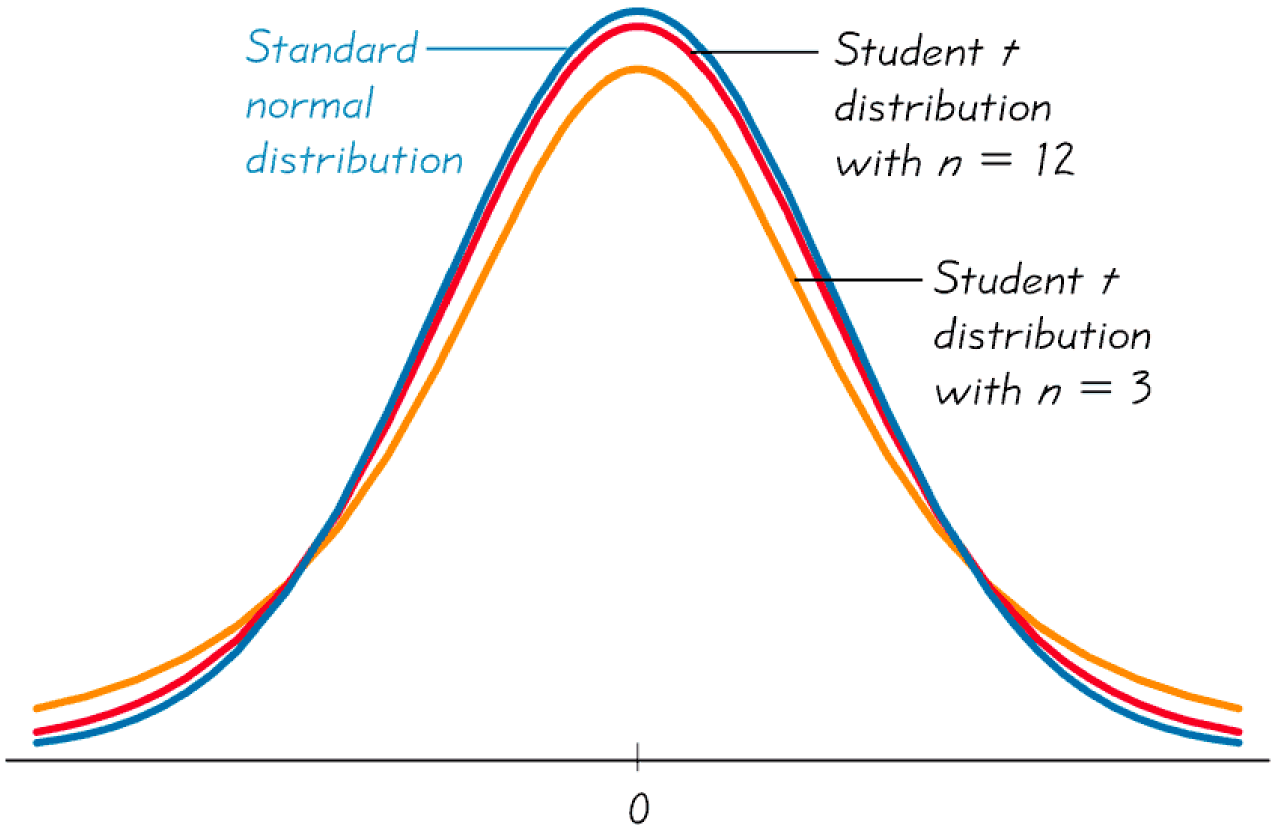
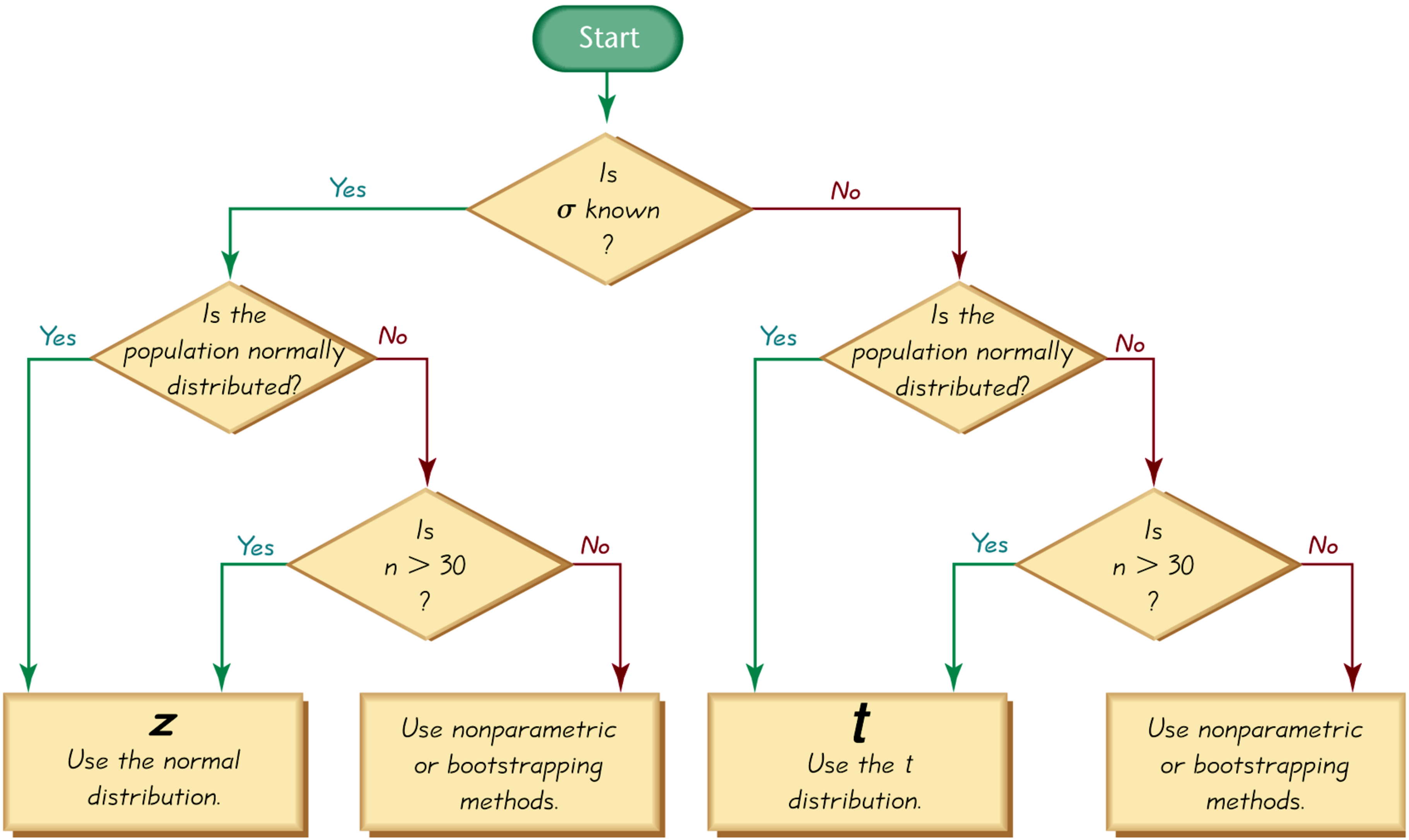


Figure 7-5

Choosing the Appropriate Distribution



Choosing the Appropriate Distribution

Use the normal (z) distribution	σ known and normally distributed population or σ known and $n > 30$
Use t distribution	σ not known and normally distributed population or σ not known and $n > 30$
Use a nonparametric method or bootstrapping	Population is not normally distributed and $n \leq 30$

Finding the Point Estimate and *E* from a Confidence Interval

Point estimate of μ:

$$\bar{x} = \frac{\text{(upper confidence limit)} + \text{(lower confidence limit)}}{2}$$

Margin of Error:

$$E = \frac{\text{(upper confidence limit)} - \text{(lower confidence limit)}}{2}$$

Confidence Intervals for Comparing Data

As in Sections 7-2 and 7-3, confidence intervals can be used informally to compare different data sets, but the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of means.

Recap

In this section we have discussed:

- Student t distribution.
- Degrees of freedom.
- Margin of error.
- Confidence intervals for μ with σ unknown.
- Choosing the appropriate distribution.
- Point estimates.
- Using confidence intervals to compare data.

Section 7-5
Estimating a Population Variance

Key Concept

This section we introduce the chi-square probability distribution so that we can construct confidence interval estimates of a population standard deviation or variance. We also present a method for determining the sample size required to estimate a population standard deviation or variance.

Chi-Square Distribution

In a normally distributed population with variance σ^2 assume that we randomly select independent samples of size n and, for each sample, compute the sample variance s^2 (which is the square of the sample standard deviation s). The sample statistic χ^2 (pronounced chi-square) has a sampling distribution called the chi-square distribution.

Chi-Square Distribution

$$\chi^2 = \frac{(n-1)\, s^2}{\sigma^2}$$

where

n = sample size

s^2 = sample variance

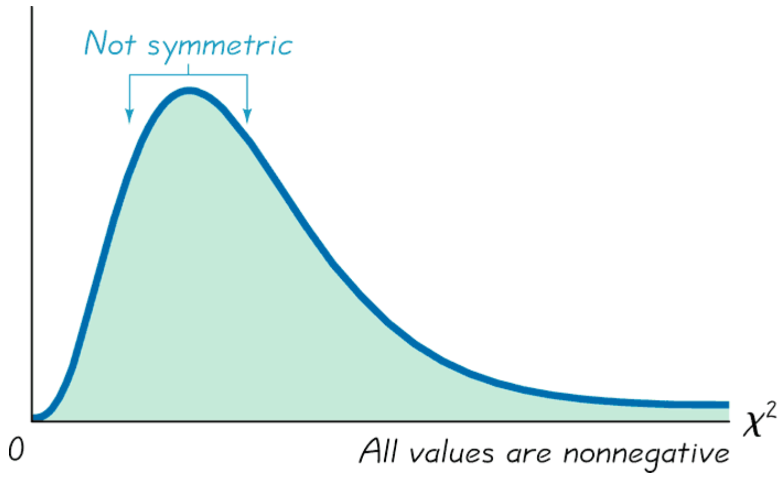
σ^2 = population variance

degrees of freedom = $n - 1$

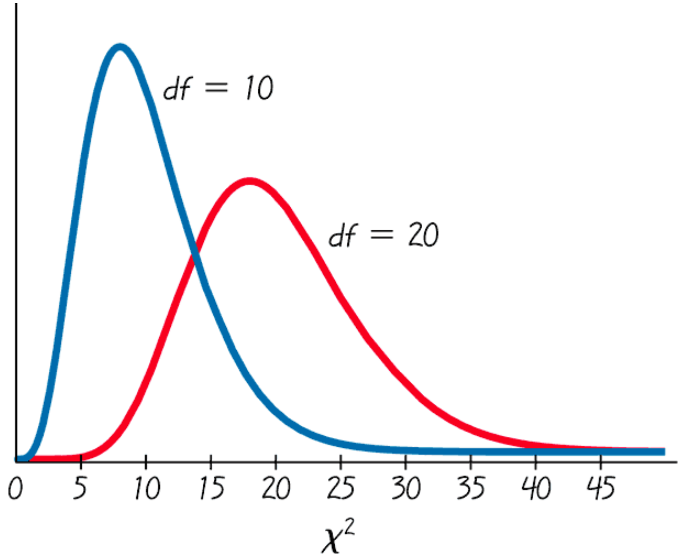
Properties of the Distribution of the Chi-Square Statistic

1. The chi-square distribution is not symmetric, unlike the normal and Student *t* distributions.

As the number of degrees of freedom increases, the distribution becomes more symmetric.



Chi-Square Distribution



Chi-Square Distribution for df = 10 and df = 20

Properties of the Distribution of the Chi-Square Statistic – cont.

2. **The values of chi-square can be zero or positive, but they cannot be negative.**

3. **The chi-square distribution is different for each number of degrees of freedom, which is df = *n* – 1. As the number of degrees of freedom increases, the chi-square distribution approaches a normal distribution.**

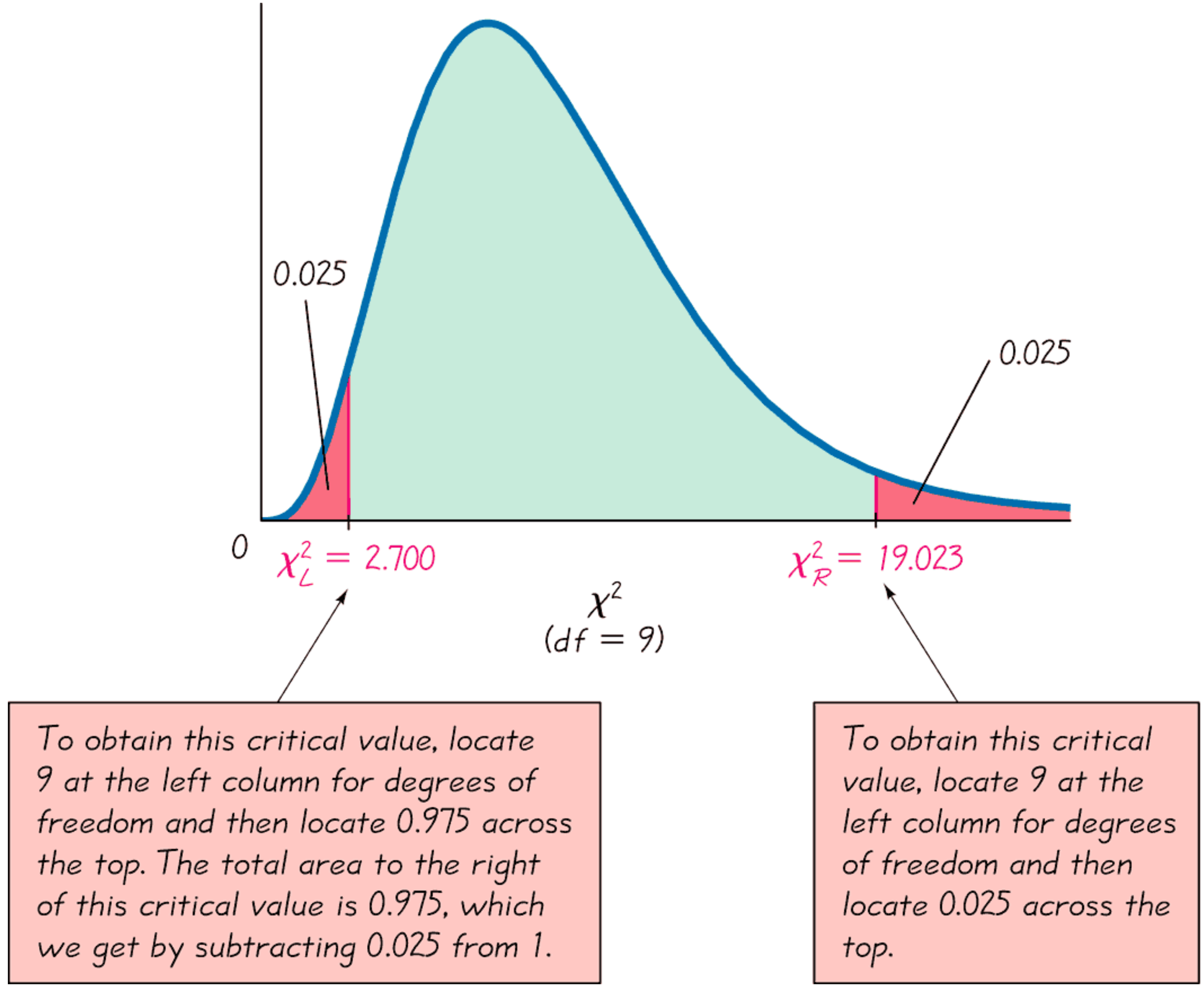
   In Table A-4, each critical value of χ^2 corresponds to an area given in the top row of the table, and that area represents the cumulative area located to the right of the critical value.

Example

A simple random sample of ten voltage levels is obtained. Construction of a confidence interval for the population standard deviation σ requires the left and right critical values of χ^2 corresponding to a confidence level of 95% and a sample size of n = 10. Find the critical value of χ^2 separating an area of 0.025 in the left tail, and find the critical value of χ^2 separating an area of 0.025 in the right tail.

Example

Critical Values of the Chi-Square Distribution



Example

For a sample of 10 values taken from a normally distributed population, the chi-square statistic $\chi^2 = (n - 1)s^2/\sigma^2$ has a 0.95 probability of falling between the chi-square critical values of 2.700 and 19.023.

Instead of using Table A-4, technology (such as STATDISK, Excel, and Minitab) can be used to find critical values of χ^2. A major advantage of technology is that it can be used for any number of degrees of freedom and any confidence level, not just the limited choices included in Table A-4.

Estimators of σ^2

The sample variance s^2 is the best point estimate of the population variance σ^2.

Estimators of σ

The sample standard deviation *s* is a commonly used point estimate of σ (even though it is a biased estimate).

Confidence Interval for Estimating a Population Standard Deviation or Variance

σ = population standard deviation

s = sample standard deviation

n = number of sample values

χ_L^2 = left-tailed critical value of χ^2

σ^2 = population variance

s^2 = sample variance

E = margin of error

χ_R^2 = right-tailed critical value of χ^2

Confidence Interval for Estimating a Population Standard Deviation or Variance

Requirements:

1. **The sample is a simple random sample.**

2. **The population must have normally distributed values (even if the sample is large).**

Confidence Interval for Estimating a Population Standard Deviation or Variance

Confidence Interval for the Population Variance σ^2

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

Confidence Interval for Estimating a Population Standard Deviation or Variance

Confidence Interval for the Population Standard Deviation σ

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

Procedure for Constructing a Confidence Interval for σ or σ^2

1. Verify that the required assumptions are satisfied.

2. Using $n - 1$ degrees of freedom, refer to Table A-4 or use technology to find the critical values χ^2_R and χ^2_L that correspond to the desired confidence level.

3. Evaluate the upper and lower confidence interval limits using this format of the confidence interval:

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

Procedure for Constructing a Confidence Interval for σ or σ^2 - cont

4. If a confidence interval estimate of σ is desired, take the square root of the upper and lower confidence interval limits and change σ^2 to σ.

5. Round the resulting confidence level limits. If using the original set of data to construct a confidence interval, round the confidence interval limits to one more decimal place than is used for the original set of data. If using the sample standard deviation or variance, round the confidence interval limits to the same number of decimals places.

Confidence Intervals for Comparing Data
Caution

Confidence intervals can be used informally to compare the variation in different data sets, but the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of variances or standard deviations.

Example:

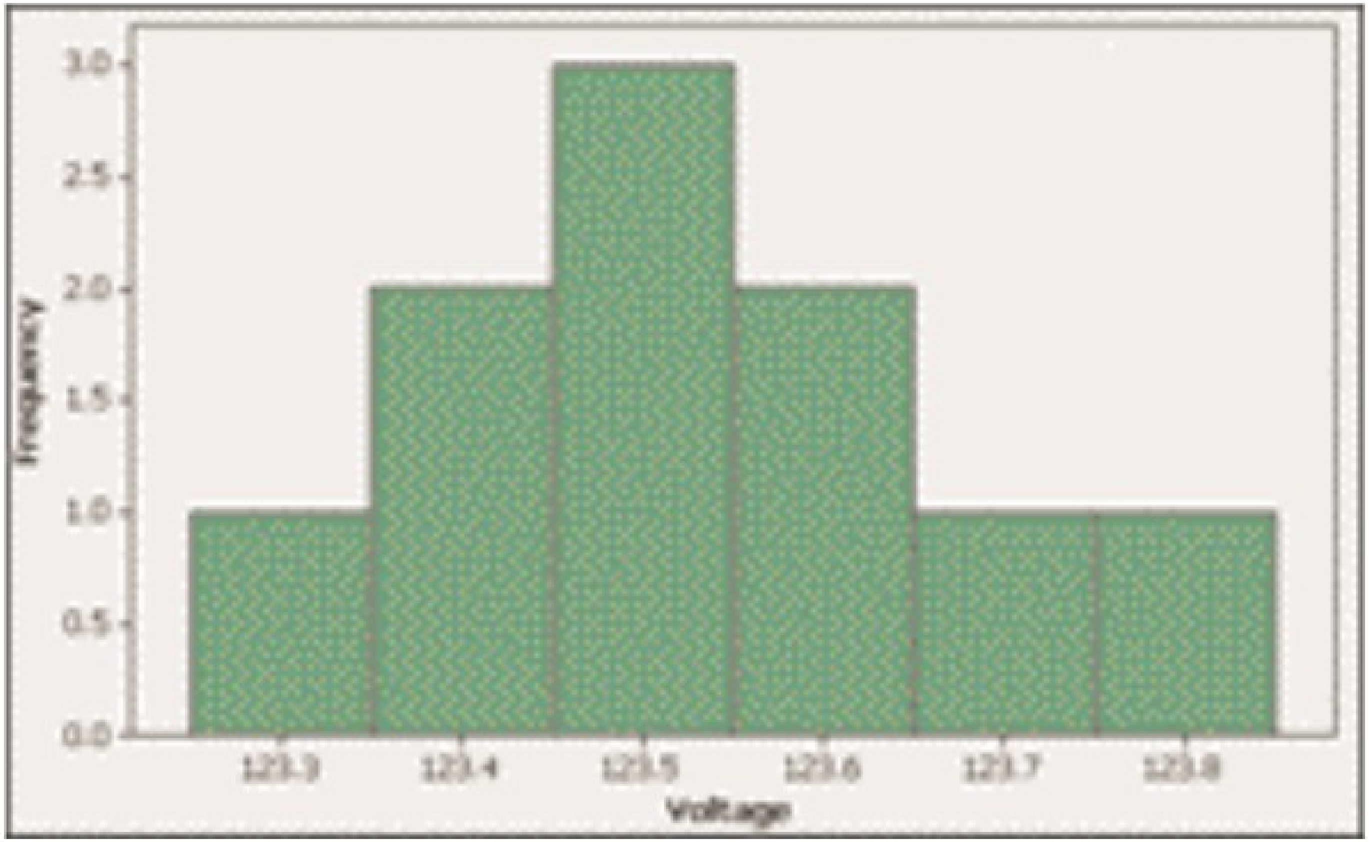
The proper operation of typical home appliances requires voltage levels that do not vary much. Listed below are ten voltage levels (in volts) recorded in the author's home on ten different days. These ten values have a standard deviation of $s = 0.15$ volt. Use the sample data to construct a 95% confidence interval estimate of the standard deviation of all voltage levels.

123.3 123.5 123.7 123.4 123.6 123.5 123.5 123.4 123.6 123.8

Example:

Requirements are satisfied: simple random sample and normality

MINITAB



Example:

n = 10 so df = 10 – 1 = 9

Use table A-4 to find:

$$\chi_L^2 = 2.700 \text{ and } \chi_R^2 = 19.023$$

Construct the confidence interval: n = 10, s = 0.15

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

$$\frac{(10-1)(0.15)^2}{19.023} < \sigma^2 < \frac{(10-1)(0.15)^2}{2.700}$$

Example:

Evaluation the preceding expression yields:

$$0.010645 < \sigma^2 < 0.075000$$

Finding the square root of each part (before rounding), then rounding to two decimal places, yields this 95% confidence interval estimate of the population standard deviation:

$$0.10 \text{ volt} < \sigma < 0.27 \text{ volt.}$$

Based on this result, we have 95% confidence that the limits of 0.10 volt and 0.27 volt contain the true value of σ.

Determining Sample Sizes

The procedures for finding the sample size necessary to estimate σ^2 are much more complex than the procedures given earlier for means and proportions. Instead of using very complicated procedures, we will use Table 7-2.

STATDISK also provides sample sizes. With STATDISK, select Analysis, Sample Size Determination, and then Estimate St Dev.

Minitab, Excel, and the TI-83/84 Plus calculator do not provide such sample sizes.

Determining Sample Sizes

Sample Size for σ^2		Sample Size for σ	
To be 95% confident that s^2 is within	of the value of σ^2, the sample size n should be at least	To be 95% confident that s is within	of the value of σ, the sample size n should be at least
1%	77,208	1%	19,205
5%	3,149	5%	768
10%	806	10%	192
20%	211	20%	48
30%	98	30%	21
40%	57	40%	12
50%	38	50%	8
To be 99% confident that s^2 is within	of the value of σ^2, the sample size n should be at least	To be 99% confident that s is within	of the value of σ, the sample size n should be at least
1%	133,449	1%	33,218
5%	5,458	5%	1,336
10%	1,402	10%	336
20%	369	20%	85
30%	172	30%	38
40%	101	40%	22
50%	68	50%	14

Example:

We want to estimate the standard deviation σ of all voltage levels in a home. We want to be 95% confident that our estimate is within 20% of the true value of σ. How large should the sample be? Assume that the population is normally distributed.

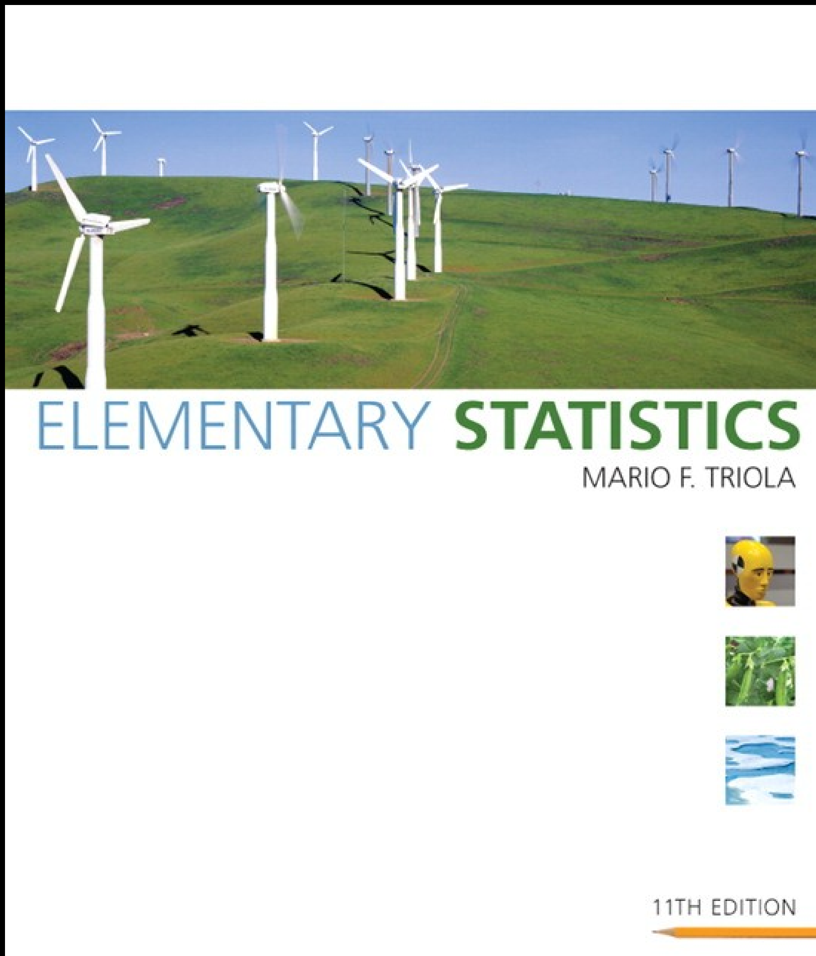
From Table 7-2, we can see that 95% confidence and an error of 20% for σ correspond to a sample of size 48. We should obtain a simple random sample of 48 voltage levels form the population of voltage levels.

Recap

In this section we have discussed:

- **The chi-square distribution.**
- **Using Table A-4.**
- **Confidence intervals for the population variance and standard deviation.**
- **Determining sample sizes.**

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 8
Hypothesis Testing

8-1 Review and Preview

8-2 Basics of Hypothesis Testing

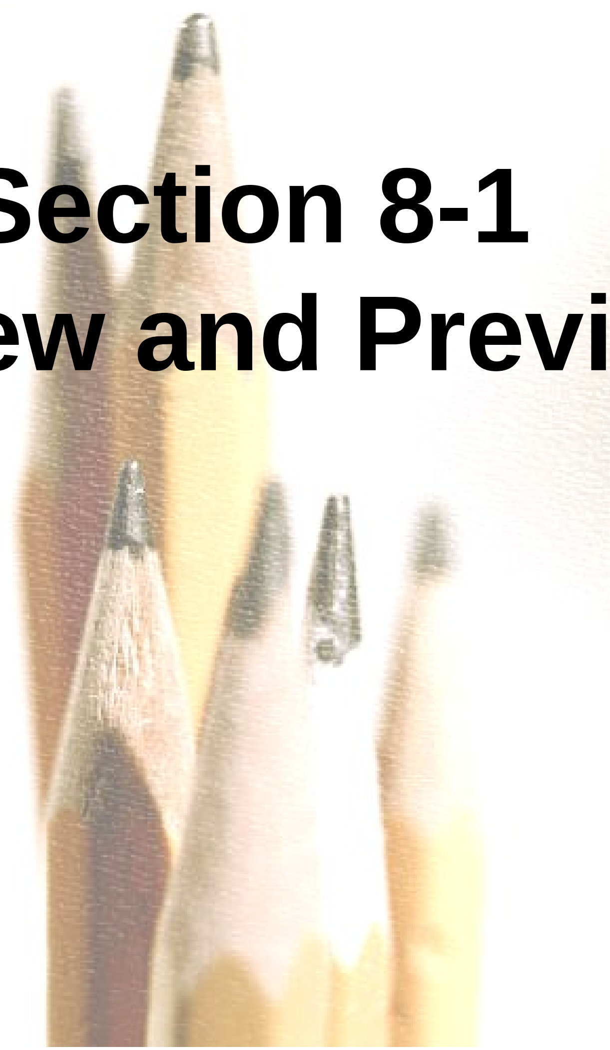
8-3 Testing a Claim about a Proportion

8-4 Testing a Claim About a Mean: σ Known

8-5 Testing a Claim About a Mean: σ Not Known

8-6 Testing a Claim About a Standard Deviation or Variance

Section 8-1
Review and Preview



Review

In Chapters 2 and 3 we used "descriptive statistics" when we summarized data using tools such as graphs, and statistics such as the mean and standard deviation. Methods of inferential statistics use sample data to make an inference or conclusion about a population. The two main activities of inferential statistics are using sample data to (1) estimate a population parameter (such as estimating a population parameter with a confidence interval), and (2) test a hypothesis or claim about a population parameter. In Chapter 7 we presented methods for estimating a population parameter with a confidence interval, and in this chapter we present the method of hypothesis testing.

Definitions

In statistics, a hypothesis is a claim or statement about a property of a population.

A hypothesis test (or test of significance) is a standard procedure for testing a claim about a property of a population.

Main Objective

The main objective of this chapter is to develop the ability to conduct hypothesis tests for claims made about a population proportion p, a population mean μ, or a population standard deviation σ.

Examples of Hypotheses that can be Tested

- **Genetics:** The Genetics & IVF Institute claims that its XSORT method allows couples to increase the probability of having a baby girl.
- **Business:** A newspaper headline makes the claim that most workers get their jobs through networking.
- **Medicine:** Medical researchers claim that when people with colds are treated with echinacea, the treatment has no effect.

Examples of Hypotheses that can be Tested

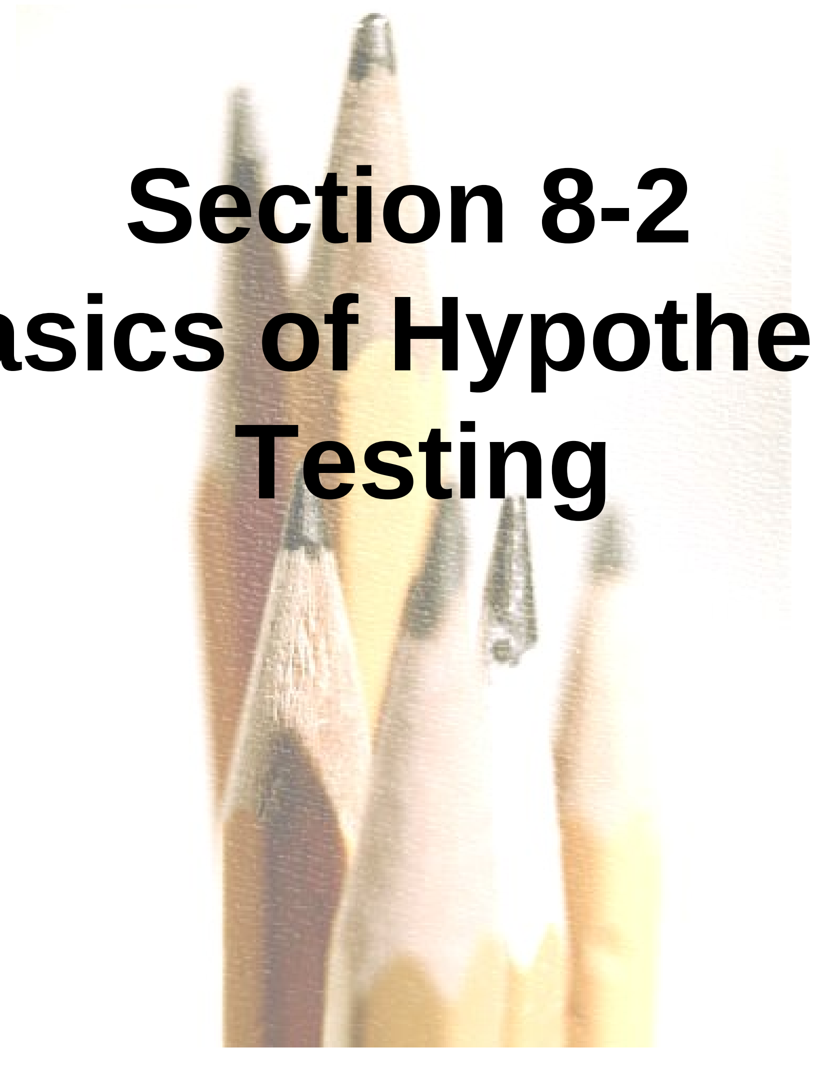
- **Aircraft Safety:** The Federal Aviation Administration claims that the mean weight of an airline passenger (including carry-on baggage) is greater than 185 lb, which it was 20 years ago.

- **Quality Control:** When new equipment is used to manufacture aircraft altimeters, the new altimeters are better because the variation in the errors is reduced so that the readings are more consistent. (In many industries, the quality of goods and services can often be improved by reducing variation.)

Caution

When conducting hypothesis tests as described in this chapter and the following chapters, instead of jumping directly to procedures and calculations, be sure to consider the context of the data, the source of the data, and the sampling method used to obtain the sample data.

Section 8-2
Basics of Hypothesis Testing



Key Concept

This section presents individual components of a hypothesis test. We should know and understand the following:

- **How to identify the null hypothesis and alternative hypothesis from a given claim, and how to express both in symbolic form**
- **How to calculate the value of the test statistic, given a claim and sample data**
- **How to identify the critical value(s), given a significance level**
- **How to identify the P-value, given a value of the test statistic**
- **How to state the conclusion about a claim in simple and nontechnical terms**

Part 1:

The Basics of Hypothesis Testing

Rare Event Rule for Inferential Statistics

If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.

Components of a Formal Hypothesis Test

Null Hypothesis: H_0

- The null hypothesis (denoted by H_0) is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.

- We test the null hypothesis directly.

- Either reject H_0 or fail to reject H_0.

Alternative Hypothesis: H_1

- **The alternative hypothesis (denoted by H_1 or H_a or H_A) is the statement that the parameter has a value that somehow differs from the null hypothesis.**
- **The symbolic form of the alternative hypothesis must use one of these symbols: $\neq$, $<$, $>$.**

Note about Forming Your Own Claims (Hypotheses)

If you are conducting a study and want to use a hypothesis test to support your claim, the claim must be worded so that it becomes the alternative hypothesis.

Note about Identifying H_0 and H_1

Figure 8-2

Start

Identify the specific claim or hypothesis to be tested, and express it in symbolic form.

Give the symbolic form that must be true when the original claim is false.

Of the two symbolic expressions obtained so far, let the alternative hypothesis H_1 be the one not containing equality, so that H_1 uses the symbol $<$ or $>$ or $\neq$. Let the null hypothesis H_0 be the symbolic expression that the parameter equals the fixed value being considered.

Example:

Consider the claim that the mean weight of airline passengers (including carry-on baggage) is at most 195 lb (the current value used by the Federal Aviation Administration). Follow the three-step procedure outlined in Figure 8-2 to identify the null hypothesis and the alternative hypothesis.

Example:

Step 1: Express the given claim in symbolic form. The claim that the mean is at most 195 lb is expressed in symbolic form as $\mu \le 195$ lb.

Step 2: If $\mu \le 195$ lb is false, then $\mu > 195$ lb must be true.

Example:

Step 3: Of the two symbolic expressions $\mu \leq 195$ lb and $\mu > 195$ lb, we see that $\mu > 195$ lb does not contain equality, so we let the alternative hypothesis H_1 be $\mu > 195$ lb. Also, the null hypothesis must be a statement that the mean equals 195 lb, so we let H_0 be $\mu = 195$ lb.

Note that the original claim that the mean is at most 195 lb is neither the alternative hypothesis nor the null hypothesis. (However, we would be able to address the original claim upon completion of a hypothesis test.)

Test Statistic

The test statistic is a value used in making a decision about the null hypothesis, and is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.

Test Statistic - Formulas

Test statistic for proportion

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test statistic for mean

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{or} \quad t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Test statistic for standard deviation

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

Example:

Let's again consider the claim that the XSORT method of gender selection increases the likelihood of having a baby girl. Preliminary results from a test of the XSORT method of gender selection involved 14 couples who gave birth to 13 girls and 1 boy. Use the given claim and the preliminary results to calculate the value of the test statistic. Use the format of the test statistic given above, so that a normal distribution is used to approximate a binomial distribution. (There are other exact methods that do not use the normal approximation.)

Example:

The claim that the XSORT method of gender selection increases the likelihood of having a baby girl results in the following null and alternative hypotheses H_0: p = 0.5 and H_1: p > 0.5. We work under the assumption that the null hypothesis is true with $p = 0.5$. The sample proportion of 13 girls in 14 births results in $\hat{p} = 13/14 = 0.929$. Using p = 0.5, $\hat{p} = 0.929$ and n = 14, we find the value of the test statistic as follows:

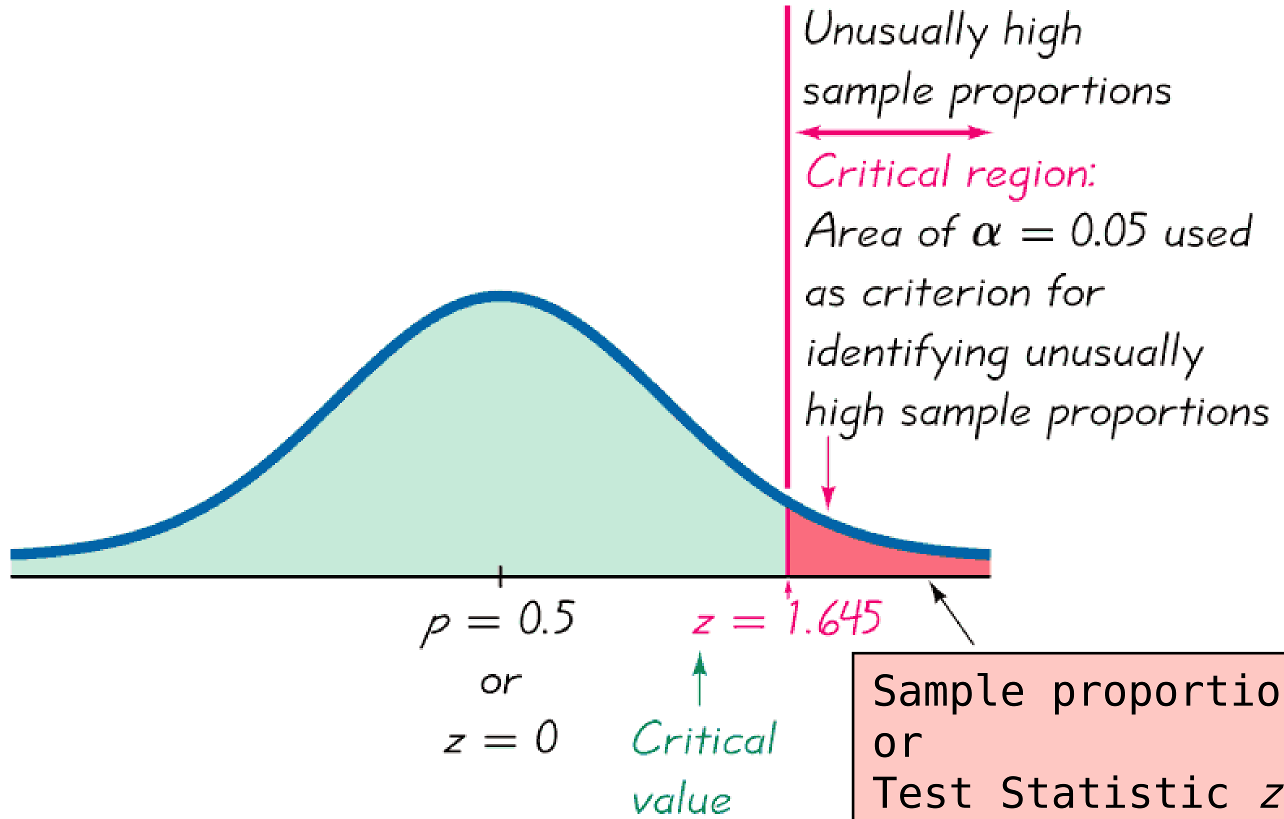
Example:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.929 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{14}}} = 3.21$$

We know from previous chapters that a z score of 3.21 is "unusual" (because it is greater than 2). It appears that in addition to being greater than 0.5, the sample proportion of 13/14 or 0.929 is significantly greater than 0.5. The figure on the next slide shows that the sample proportion of 0.929 does fall within the range of values considered to be significant because

Example:

they are so far above 0.5 that they are not likely to occur by chance (assuming that the population proportion is *p* = 0.5).



Sample proportion of: $\hat{p} = 0.929$
or
Test Statistic $z = 3.21$

Critical Region

The critical region (or rejection region) is the set of all values of the test statistic that cause us to reject the null hypothesis. For example, see the red-shaded region in the previous figure.

Significance Level

The significance level (denoted by α) is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true. This is the same α introduced in Section 7-2. Common choices for α are 0.05, 0.01, and 0.10.

Critical Value

A critical value is any value that separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis. The critical values depend on the nature of the null hypothesis, the sampling distribution that applies, and the significance level α. See the previous figure where the critical value of $z = 1.645$ corresponds to a significance level of $\alpha = 0.05$.

P-Value

The ***P*-value** (or ***p*-value** or **probability value**) is the probability of getting a value of the test statistic that is **at least as extreme** as the one representing the sample data, assuming that the null hypothesis is true.

Critical region in the **left** tail:	*P*-value = area to the **left** of the test statistic
Critical region in the **right** tail:	*P*-value = area to the **right** of the test statistic
Critical region in **two** tails:	*P*-value = **twice** the area in the tail beyond the test statistic

P-Value

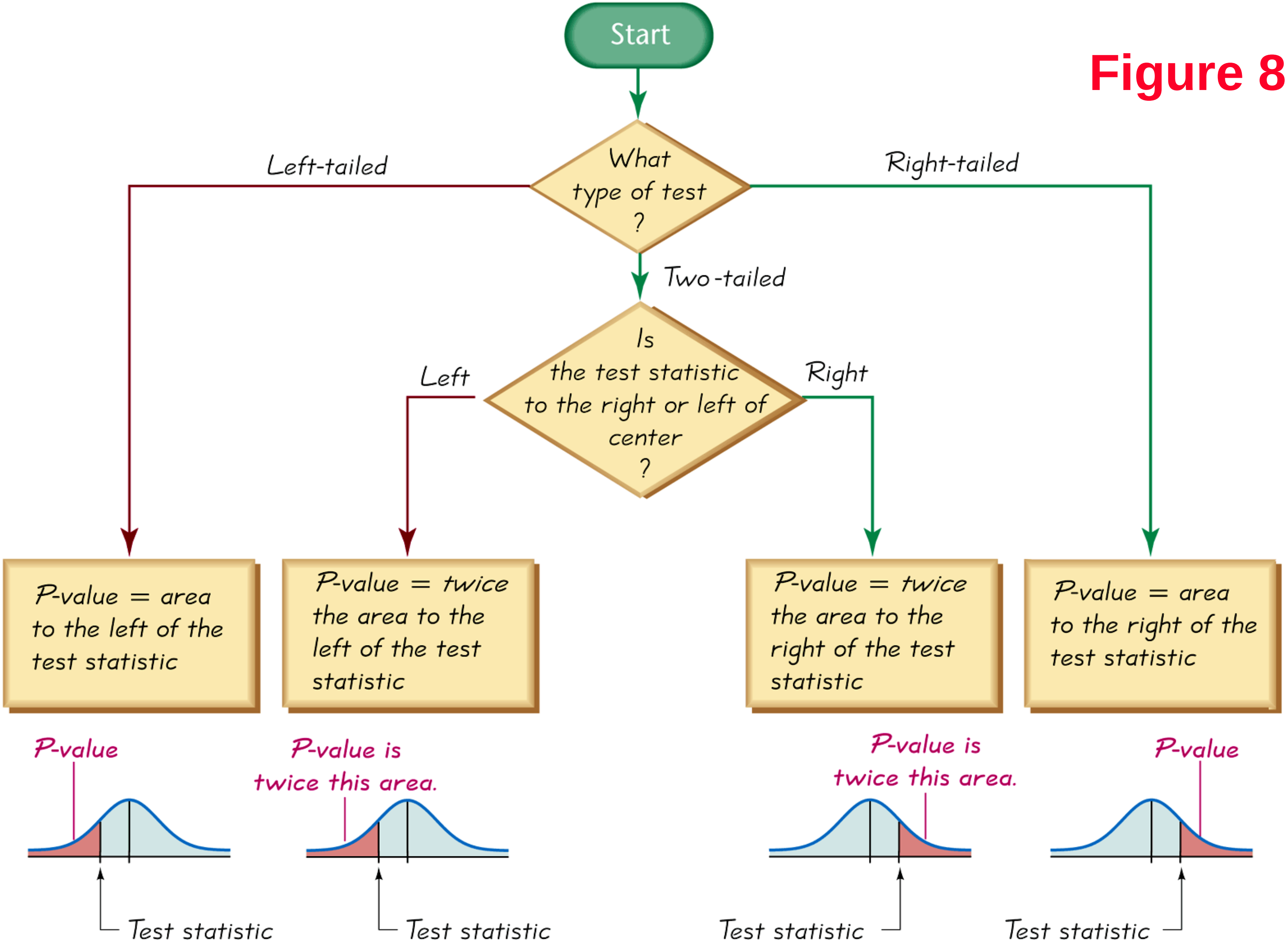
The null hypothesis is rejected if the *P*-value is very small, such as 0.05 or less.

Here is a memory tool useful for interpreting the P-value:

If the P is low, the null must go.
If the P is high, the null will fly.

Procedure for Finding *P*-Values

Figure 8-5



Caution

Don't confuse a *P*-value with a proportion *p*.
Know this distinction:

***P*-value = probability of getting a test statistic at least as extreme as the one representing sample data**

***p* = population proportion**

Example

Consider the claim that with the XSORT method of gender selection, the likelihood of having a baby girl is different from *p* = 0.5, and use the test statistic *z* = 3.21 found from 13 girls in 14 births. First determine whether the given conditions result in a critical region in the right tail, left tail, or two tails, then use Figure 8-5 to find the *P*-value. Interpret the *P*-value.

Example

The claim that the likelihood of having a baby girl is different from p = 0.5 can be expressed as $p \neq 0.5$ so the critical region is in two tails. Using Figure 8-5 to find the P-value for a two-tailed test, we see that the P-value is *twice* the area to the right of the test statistic z = 3.21. We refer to Table A-2 (or use technology) to find that the area to the right of z = 3.21 is 0.0007. In this case, the P-value is twice the area to the right of the test statistic, so we have:

$$P\text{-value} = 2 \times 0.0007 = 0.0014$$

Example

The *P*-value is 0.0014 (or 0.0013 if greater precision is used for the calculations). The small *P*-value of 0.0014 shows that there is a very small chance of getting the sample results that led to a test statistic of z = 3.21. This suggests that with the XSORT method of gender selection, the likelihood of having a baby girl is different from 0.5.

Types of Hypothesis Tests: Two-tailed, Left-tailed, Right-tailed

The tails in a distribution are the extreme regions bounded by critical values.

Determinations of *P*-values and critical values are affected by whether a critical region is in two tails, the left tail, or the right tail. It therefore becomes important to correctly characterize a hypothesis test as two-tailed, left-tailed, or right-tailed.

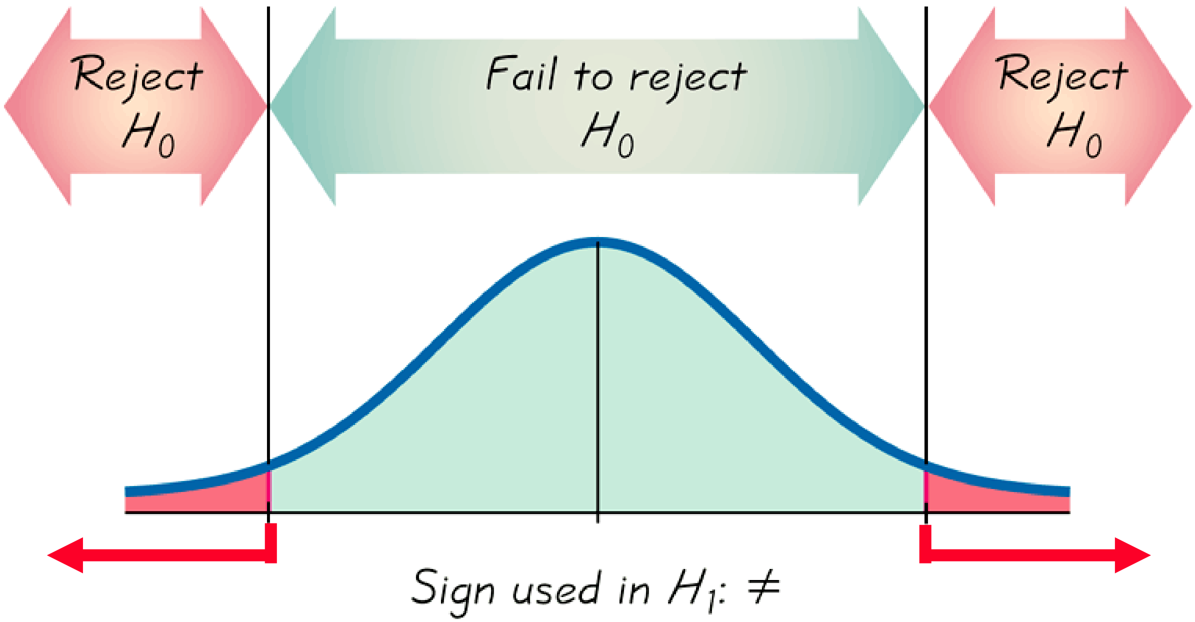
Two-tailed Test

H_0: =

H_1: ≠

Means less than or greater than

α is divided equally between the two tails of the critical region



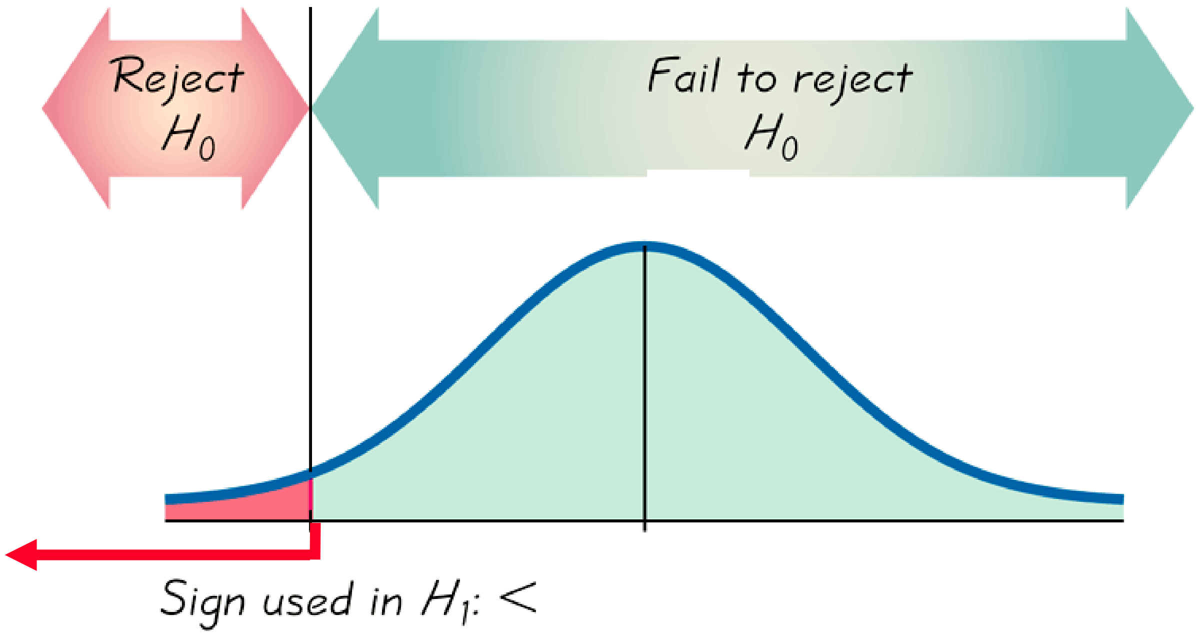
Left-tailed Test

α the left tail

H_0: =

H_1: <

Points Left

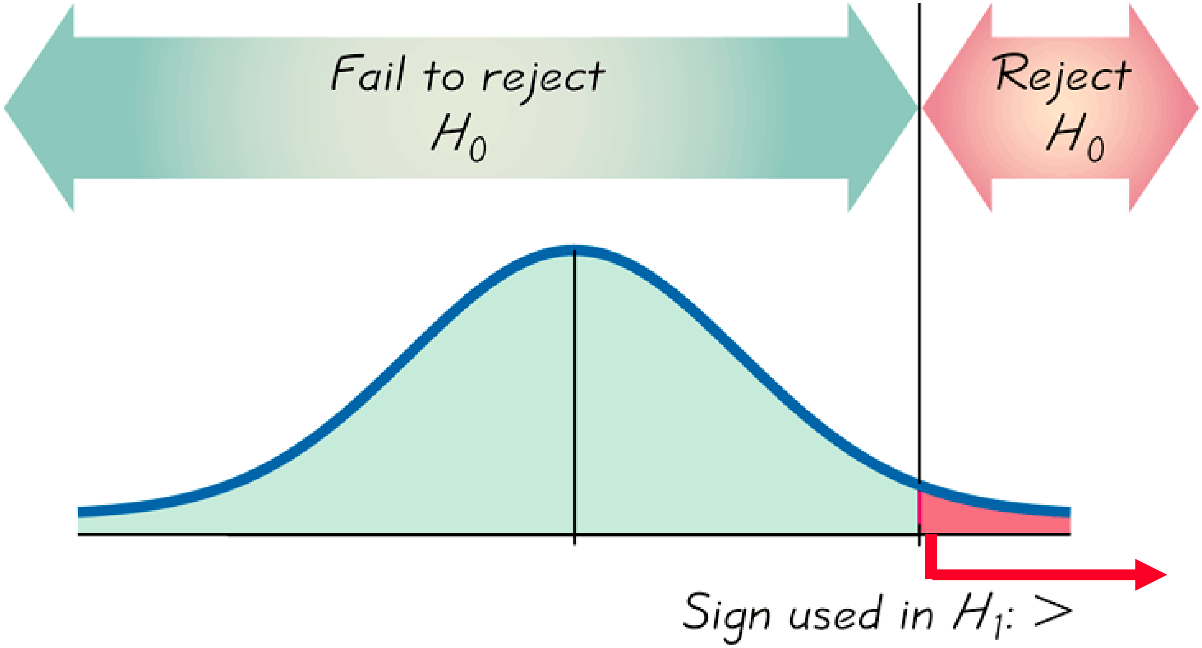


Right-tailed Test

H_0: =

H_1: >

Points Right



Conclusions in Hypothesis Testing

We always test the null hypothesis. The initial conclusion will always be one of the following:

1. Reject the null hypothesis.

2. Fail to reject the null hypothesis.

Decision Criterion

***P*-value method**:
Using the significance level α:

If *P*-value $\leq \alpha$, reject H_0.

If *P*-value $> \alpha$, fail to reject H_0.

Decision Criterion

Traditional method:

If the test statistic falls within the critical region, reject H_0.

If the test statistic does not fall within the critical region, fail to reject H_0.

Decision Criterion

Another option:

Instead of using a significance level such as 0.05, simply identify the *P*-value and leave the decision to the reader.

Decision Criterion

Confidence Intervals:

A confidence interval estimate of a population parameter contains the likely values of that parameter.
If a confidence interval does not include a claimed value of a population parameter, reject that claim.

Wording of Final Conclusion

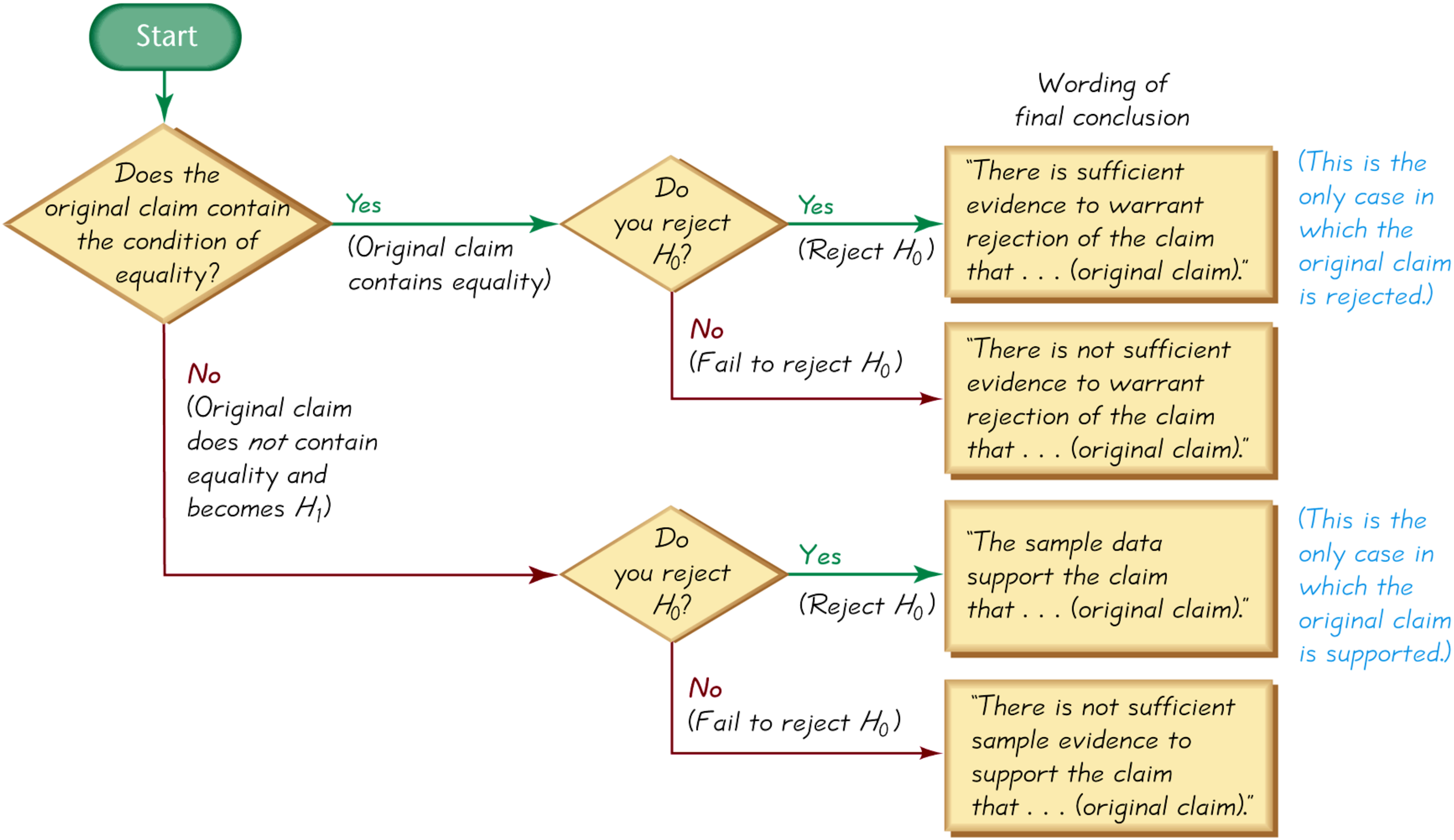


Figure 8-7

Caution

Never conclude a hypothesis test with a statement of “reject the null hypothesis” or “fail to reject the null hypothesis.” Always make sense of the conclusion with a statement that uses simple nontechnical wording that addresses the original claim.

Accept Versus Fail to Reject

- Some texts use "accept the null hypothesis."
- We are not proving the null hypothesis.
- Fail to reject says more correctly
- The available evidence is not strong enough to warrant rejection of the null hypothesis (such as not enough evidence to convict a suspect).

Type I Error

- A **Type I error** is the mistake of rejecting the null hypothesis when it is actually true.

- The symbol α (alpha) is used to represent the probability of a type I error.

Type II Error

- **A Type II error is the mistake of failing to reject the null hypothesis when it is actually false.**

- **The symbol β (beta) is used to represent the probability of a type II error.**

Type I and Type II Errors

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	**Type I error** (rejecting a true null hypothesis) $P(\text{type I error}) = \alpha$	Correct decision
	We fail to reject the null hypothesis	Correct decision	**Type II error** (failing to reject a false null hypothesis) $P(\text{type II error}) = \beta$

Example:

Assume that we are conducting a hypothesis test of the claim that a method of gender selection increases the likelihood of a baby girl, so that the probability of a baby girls is $p > 0.5$. Here are the null and alternative hypotheses: H_0: $p = 0.5$, and H_1: $p > 0.5$.

a) **Identify a type I error.**
b) **Identify a type II error.**

Example:

a) A type I error is the mistake of rejecting a true null hypothesis, so this is a type I error: Conclude that there is sufficient evidence to support $p > 0.5$, when in reality $p = 0.5$.

b) A type II error is the mistake of failing to reject the null hypothesis when it is false, so this is a type II error: Fail to reject $p = 0.5$ (and therefore fail to support $p > 0.5$) when in reality $p > 0.5$.

Controlling Type I and Type II Errors

- For any fixed α, an increase in the sample size n will cause a decrease in β.
- For any fixed sample size n, a decrease in α will cause an increase in β. Conversely, an increase in α will cause a decrease in β.
- To decrease both α and β, increase the sample size.

Comprehensive Hypothesis Test – *P*-Value Method

Start

1 Identify the specific claim or hypothesis to be tested, and put it in symbolic form.

2 Give the symbolic form that must be true when the original claim is false.

3 Of the two symbolic expressions obtained so far, let the alternative hypothesis H_1 be the one not containing equality, so that H_1 uses the symbol $>$ or $<$ or $\neq$. Let the null hypothesis H_0 be the symbolic expression that the parameter *equals* the fixed value being considered.

4 Select the significance level α based on the seriousness of a type 1 error. Make α small if the consequences of rejecting a true H_0 are severe. The values of 0.05 and 0.01 are very common.

5 Identify the statistic that is relevant to this test and determine its sampling distribution (such as normal, *t*, chi-square).

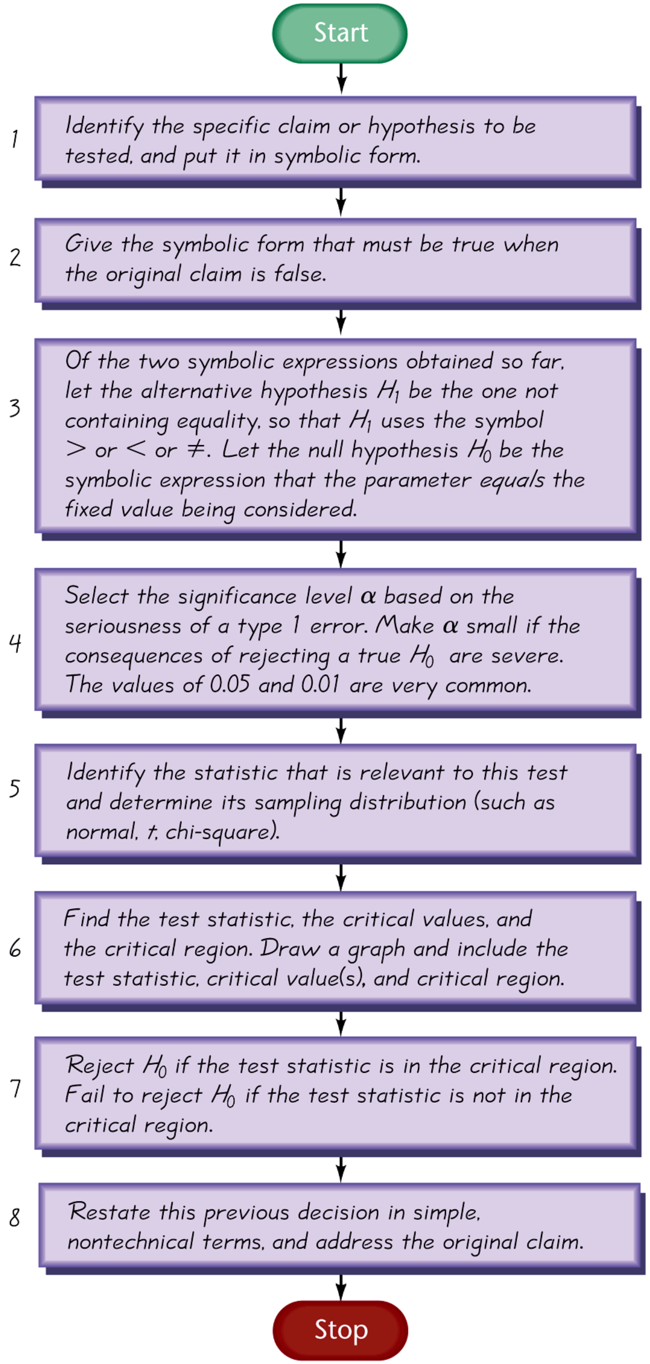
6 Find the test statistic and find the *P*-value (see Figure 8-6). Draw a graph and show the test statistic and *P*-value.

7 Reject H_0 if the *P*-value is less than or equal to the significance level α. Fail to reject H_0 if the *P*-value is greater than α.

8 Restate this previous decision in simple, nontechnical terms, and address the original claim.

Stop

Comprehensive Hypothesis Test – Traditional Method



Comprehensive Hypothesis Test - cont

A confidence interval estimate of a population parameter contains the likely values of that parameter. We should therefore reject a claim that the population parameter has a value that is not included in the confidence interval.

Table 8-2 Confidence Level for Confidence Interval

		Two-Tailed Test	One-Tailed Test
Significance Level for Hypothesis Test	0.01	99%	98%
	0.05	95%	90%
	0.10	90%	80%

Caution

In some cases, a conclusion based on a confidence interval may be different from a conclusion based on a hypothesis test. See the comments in the individual sections.

Part 2:

Beyond the Basics of Hypothesis Testing:

The *Power* of a Test

Definition

The power of a hypothesis test is the probability ($1 - \beta$) of rejecting a false null hypothesis. The value of the power is computed by using a particular significance level α and a particular value of the population parameter that is an alternative to the value assumed true in the null hypothesis.

That is, the power of the hypothesis test is the probability of supporting an alternative hypothesis that is true.

Power and the Design of Experiments

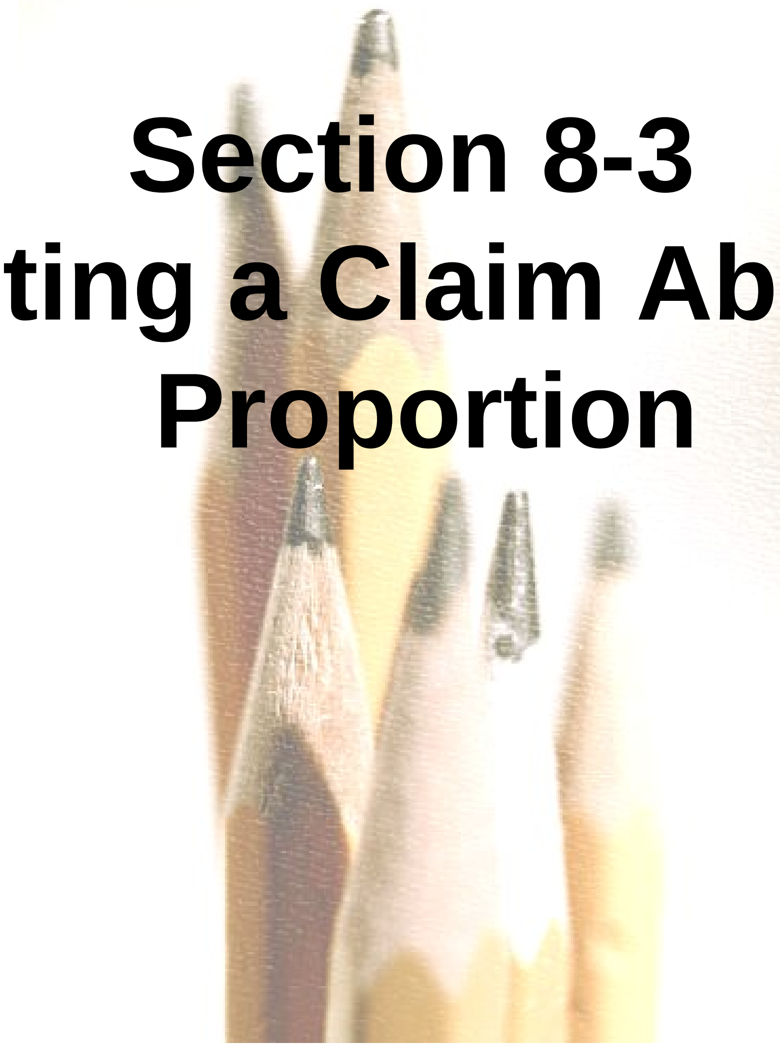
Just as 0.05 is a common choice for a significance level, a power of at least 0.80 is a common requirement for determining that a hypothesis test is effective. (Some statisticians argue that the power should be higher, such as 0.85 or 0.90.) When designing an experiment, we might consider how much of a difference between the claimed value of a parameter and its true value is an important amount of difference. When designing an experiment, a goal of having a power value of at least 0.80 can often be used to determine the minimum required sample size.

Recap

In this section we have discussed:

- **Null and alternative hypotheses.**
- **Test statistics.**
- **Significance levels.**
- ***P*-values.**
- **Decision criteria.**
- **Type I and II errors.**
- **Power of a hypothesis test.**

Section 8-3
Testing a Claim About a Proportion



Key Concept

This section presents complete procedures for testing a hypothesis (or claim) made about a population proportion. This section uses the components introduced in the previous section for the *P*-value method, the traditional method or the use of confidence intervals.

Key Concept

Two common methods for testing a claim about a population proportion are (1) to use a normal distribution as an approximation to the binomial distribution, and (2) to use an exact method based on the binomial probability distribution. Part 1 of this section uses the approximate method with the normal distribution, and Part 2 of this section briefly describes the exact method.

Part 1:

Basic Methods of Testing Claims about a Population Proportion *p*

Notation

n = number of trials

$\hat{p} = \dfrac{x}{n}$ (sample proportion)

p = population proportion (used in the null hypothesis)

$q = 1 - p$

Requirements for Testing Claims About a Population Proportion *p*

1) The sample observations are a simple random sample.

2) The conditions for a binomial distribution are satisfied.

3) The conditions $np \geq 5$ and $nq \geq 5$ are both satisfied, so the binomial distribution of sample proportions can be approximated by a normal distribution with $\mu = np$ and $\sigma = \sqrt{npq}$. Note: *p* is the assumed proportion not the sample proportion.

Test Statistic for Testing a Claim About a Proportion

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

***P*-values:** **Use the standard normal distribution (Table A-2) and refer to Figure 8-5**

Critical Values: **Use the standard normal distribution (Table A-2).**

Caution

Don't confuse a *P*-value with a proportion *p*.

***P*-value = probability of getting a test statistic at least as extreme as the one representing sample data**

***p* = population proportion**

P-Value Method:

Use the same method as described in Section 8-2 and in Figure 8-8.

Use the standard normal distribution (Table A-2).

Traditional Method

Use the same method as described in Section 8-2 and in Figure 8-9.

Confidence Interval Method

Use the same method as described in Section 8-2 and in Table 8-2.

CAUTION

When testing claims about a population proportion, the traditional method and the *P*-value method are equivalent and will yield the same result since they use the same standard deviation based on the claimed proportion *p*. However, the confidence interval uses an estimated standard deviation based upon the sample proportion $\hat{p}$. Consequently, it is possible that the traditional and *P*-value methods may yield a different conclusion than the confidence interval method.

A good strategy is to use a confidence interval to estimate a population proportion, but use the *P*-value or traditional method for testing a claim about the proportion.

Example:

The text refers to a study in which 57 out of 104 pregnant women correctly guessed the sex of their babies. Use these sample data to test the claim that the success rate of such guesses is no different from the 50% success rate expected with random chance guesses. Use a 0.05 significance level.

Example:

Requirements are satisfied: simple random sample; fixed number of trials (104) with two categories (guess correctly or do not); np = (104)(0.5) = 52 ≥ 5 and nq = (104)(0.5) = 52 ≥ 5

Step 1: original claim is that the success rate is no different from 50%: $p = 0.50$

Step 2: opposite of original claim is $p \neq 0.50$

Step 3: $p \neq 0.50$ does not contain equality so it is H_1.

H_0: $p = 0.50$ null hypothesis and original claim

H_1: $p \neq 0.50$ alternative hypothesis

Example:

Step 4: significance level is $\alpha = 0.50$

Step 5: sample involves proportion so the relevant statistic is the sample proportion, $\hat{p}$

Step 6: calculate *z*:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{57}{104} - 0.50}{\sqrt{\frac{(0.50)(0.50)}{104}}} = 0.98$$

two-tailed test, *P*-value is twice the area to the right of test statistic

Example:

Table A-2: *z* = 0.98 has an area of 0.8365 to its left, so area to the right is 1 – 0.8365 = 0.1635, doubles yields 0.3270 (technology provides a more accurate *P*-value of 0.3268

Step 7: the *P*-value of 0.3270 is greater than the significance level of 0.50, so fail to reject the null hypothesis

Here is the correct conclusion: There is not sufficient evidence to warrant rejection of the claim that women who guess the sex of their babies have a success rate equal to 50%.

Obtaining $\hat{P}$

$\hat{p}$ sometimes is given directly

"10% of the observed sports cars are red" is expressed as

$$\hat{p} = 0.10$$

$\hat{p}$ sometimes must be calculated

"96 surveyed households have cable TV and 54 do not" is calculated using

$$\hat{p} = \frac{x}{n} = \frac{96}{(96+54)} = 0.64$$

(determining the sample proportion of households with cable TV)

Part 2:

Exact Method for Testing Claims about a Proportion *p*

Testing Claims

We can get exact results by using the binomial probability distribution. Binomial probabilities are a nuisance to calculate manually, but technology makes this approach quite simple. Also, this exact approach does not require that $np \geq 5$ and $nq \geq 5$ so we have a method that applies when that requirement is not satisfied. To test hypotheses using the exact binomial distribution, use the binomial probability distribution with the *P*-value method, use the value of p assumed in the null hypothesis, and find *P*-values as follows:

Testing Claims

Left-tailed test:

The *P*-value is the probability of getting x or fewer successes among n trials.

Right-tailed test:

The *P*-value is the probability of getting x or more successes among n trials.

Testing Claims

Two-tailed test:

If $\hat{p} > p$, **the *P*-value is twice the probability of getting *x* or more successes**

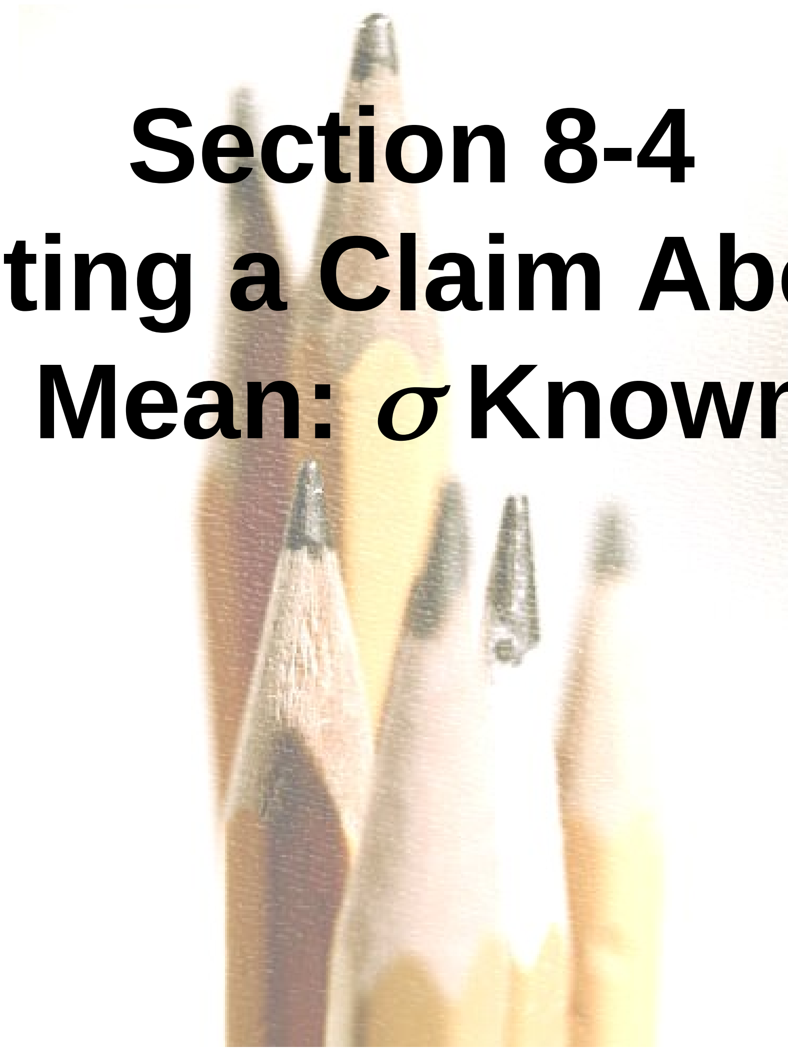
If $\hat{p} < p$, **the *P*-value is twice the probability of getting *x* or fewer successes**

Recap

In this section we have discussed:

- **Test statistics for claims about a proportion.**
- ***P*-value method.**
- **Confidence interval method.**
- **Obtaining $\hat{p}$.**

Section 8-4
Testing a Claim About a Mean: σ Known



Key Concept

This section presents methods for testing a claim about a population mean, given that the population standard deviation is a known value. This section uses the normal distribution with the same components of hypothesis tests that were introduced in Section 8-2.

Notation

n = sample size

$\bar{x}$ = sample mean

$\mu_{\bar{x}}$ = population mean of all sample means from samples of size n

σ = known value of the population standard deviation

Requirements for Testing Claims About a Population Mean (with σ Known)

1) The sample is a simple random sample.

2) The value of the population standard deviation σ is known.

3) Either or both of these conditions is satisfied: The population is normally distributed or $n > 30$.

Test Statistic for Testing a Claim About a Mean (with σ Known)

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Example:

People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: *n* = 40 and $\bar{x}$ = 172.55 lb. Research from several other sources suggests that the population of weights of men has a standard deviation given by σ = 26 lb. Use these results to test the claim that men have a mean weight greater than 166.3 lb, which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and use the *P*-value method outlined in Figure 8-8.

Example:

Requirements are satisfied: simple random sample, σ is known (26 lb), sample size is 40 (n > 30)

Step 1: Express claim as $\mu > 166.3$ lb

Step 2: alternative to claim is $\mu \leq 166.3$ lb

Step 3: $\mu > 166.3$ lb does not contain equality, it is the alternative hypothesis:

H_0: $\mu = 166.3$ lb null hypothesis

H_1: $\mu > 166.3$ lb alternative hypothesis and original claim

Example:

Step 4: significance level is α = 0.05

Step 5: claim is about the population mean, so the relevant statistic is the sample mean (172.55 lb), σ is known (26 lb), sample size greater than 30

Step 6: calculate *z*

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{172.55 - 166.3}{\frac{26}{\sqrt{40}}} = 1.52$$

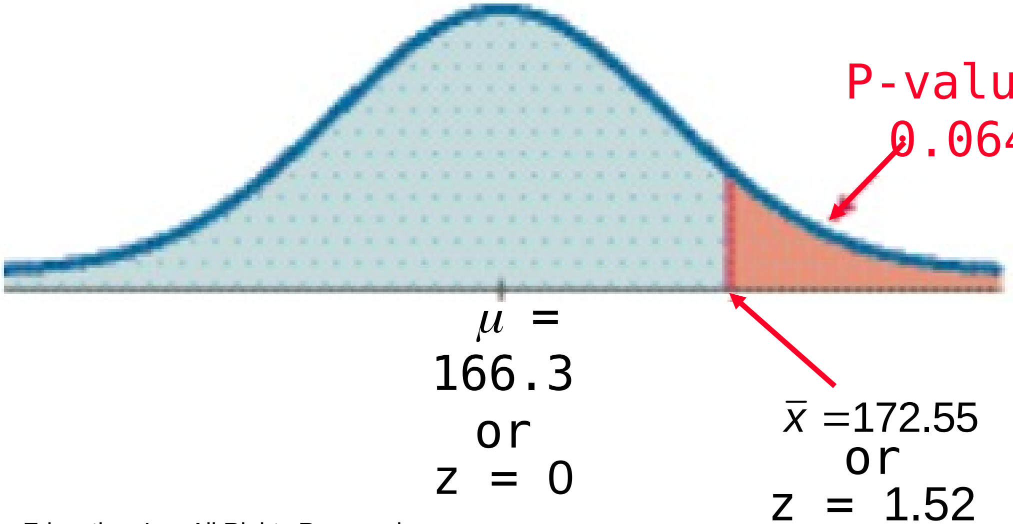
right-tailed test, so *P*-value is the area is to the right of *z* = 1.52;

Example:

Table A-2: area to the left of $z = 1.52$ is 0.9357, so the area to the right is $1 - 0.9357 = 0.0643$.

The *P*-value is 0.0643

Step 7: The *P*-value of 0.0643 is greater than the significance level of $\alpha = 0.05$, we fail to reject the null hypothesis.



Example:

The *P*-value of 0.0643 tells us that if men have a mean weight given by μ = 166.3 lb, there is a good chance (0.0643) of getting a sample mean of 172.55 lb. A sample mean such as 172.55 lb could easily occur by chance. There is not sufficient evidence to support a conclusion that the population mean is greater than 166.3 lb, as in the National Transportation and Safety Board's recommendation.

Example:

The traditional method: Use *z* = 1.645 instead of finding the *P*-value. Since *z* = 1.52 does not fall in the critical region, again fail to reject the null hypothesis.

Confidence Interval method: Use a one-tailed test with a = 0.05, so construct a 90% confidence interval:

$$165.8 < \mu < 179.3$$

The confidence interval contains 166.3 lb, we cannot support a claim that μ is greater than 166.3. Again, fail to reject the null hypothesis.

Underlying Rationale of Hypothesis Testing

If, under a given assumption, there is an extremely small probability of getting sample results at least as extreme as the results that were obtained, we conclude that the assumption is probably not correct.

When testing a claim, we make an assumption (null hypothesis) of equality. We then compare the assumption and the sample results and we form one of the following conclusions:

Underlying Rationale of Hypotheses Testing - cont

- **If the sample results (or more extreme results) can easily occur when the assumption (null hypothesis) is true, we attribute the relatively small discrepancy between the assumption and the sample results to chance.**

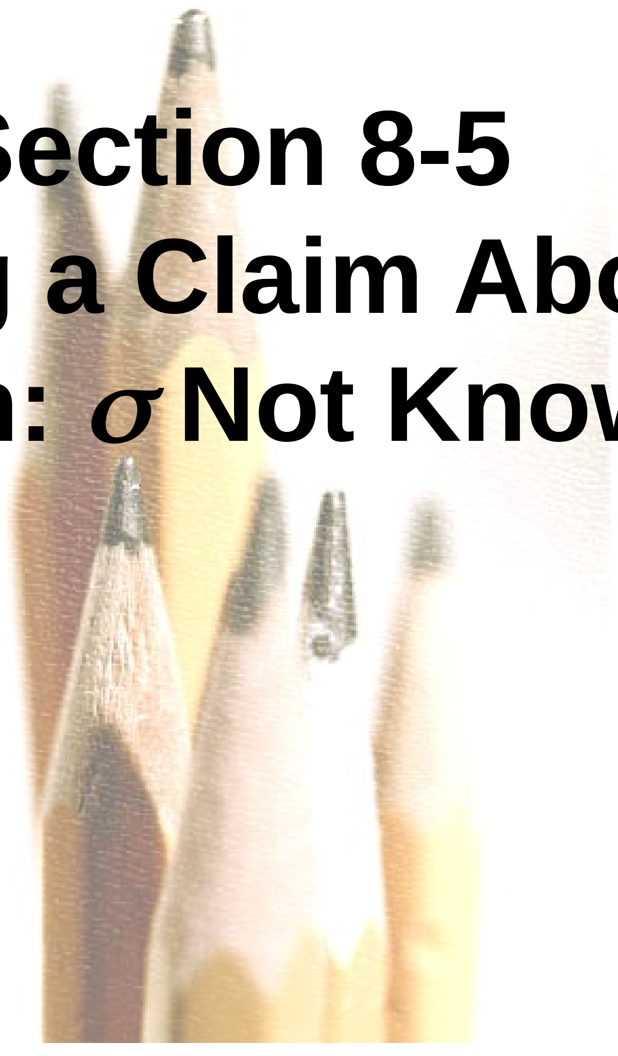
- **If the sample results cannot easily occur when that assumption (null hypothesis) is true, we explain the relatively large discrepancy between the assumption and the sample results by concluding that the assumption is not true, so we reject the assumption.**

Recap

In this section we have discussed:

- **Requirements for testing claims about population means, σ known.**
- ***P*-value method.**
- **Traditional method.**
- **Confidence interval method.**
- **Rationale for hypothesis testing.**

Section 8-5
Testing a Claim About a Mean: σ Not Known



Key Concept

This section presents methods for testing a claim about a population mean when we do not know the value of σ. The methods of this section use the Student t distribution introduced earlier.

Notation

n **= sample size**

$\bar{x}$ **= sample mean**

$\mu_{\bar{x}}$ **= population mean of all sample means from samples of size *n***

Requirements for Testing Claims About a Population Mean (with σ Not Known)

1) The sample is a simple random sample.

2) The value of the population standard deviation σ is not known.

3) Either or both of these conditions is satisfied: The population is normally distributed or $n > 30$.

Test Statistic for Testing a Claim About a Mean (with σ Not Known)

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$

P-values and Critical Values

- Found in Table A-3
- Degrees of freedom (df) = $n - 1$

Important Properties of the Student *t* Distribution

1. **The Student *t* distribution is different for different sample sizes (see Figure 7-5 in Section 7-4).**
2. **The Student *t* distribution has the same general bell shape as the normal distribution; its wider shape reflects the greater variability that is expected when *s* is used to estimate σ .**
3. **The Student *t* distribution has a mean of *t* = 0 (just as the standard normal distribution has a mean of *z* = 0).**
4. **The standard deviation of the Student *t* distribution varies with the sample size and is greater than 1 (unlike the standard normal distribution, which has σ = 1).**
5. **As the sample size *n* gets larger, the Student *t* distribution gets closer to the standard normal distribution.**

Choosing between the Normal and Student *t* Distributions when Testing a Claim about a Population Mean μ

Use the Student *t* distribution when σ is not known and either or both of these conditions is satisfied:
The population is normally distributed or $n > 30$.

Example:

People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: *n* = 40 and $\bar{x}$ = 172.55 lb, and σ = 26.33 lb. Do not assume that the value of σ is known. Use these results to test the claim that men have a mean weight greater than 166.3 lb, which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and the traditional method outlined in Figure 8-9.

Example:

Requirements are satisfied: simple random sample, population standard deviation is not known, sample size is 40 ($n > 30$)

Step 1: Express claim as $\mu > 166.3$ lb

Step 2: alternative to claim is $\mu \leq 166.3$ lb

Step 3: $\mu > 166.3$ lb does not contain equality, it is the alternative hypothesis:

H_0: $\mu = 166.3$ lb null hypothesis

H_1: $\mu > 166.3$ lb alternative hypothesis and original claim

Example:

Step 4: significance level is $\alpha = 0.05$

Step 5: claim is about the population mean, so the relevant statistic is the sample mean, 172.55 lb

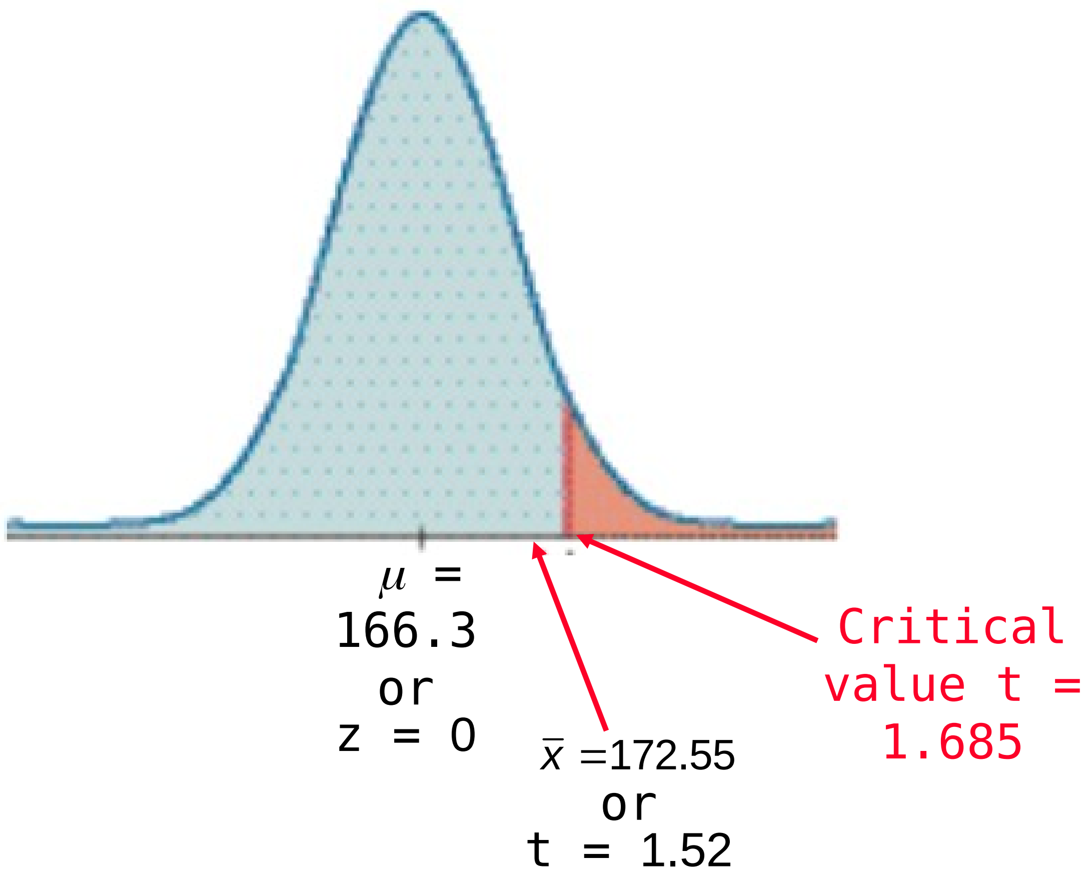
Step 6: calculate *t*

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{172.55 - 166.3}{\frac{26.33}{\sqrt{40}}} = 1.501$$

df = *n* – 1 = 39, area of 0.05, one-tail yields *t* = 1.685;

Example:

Step 7: ***t*** **= 1.501 does not fall in the critical region bounded by *t* = 1.685, we fail to reject the null hypothesis.**



Example:

Because we fail to reject the null hypothesis, we conclude that there is not sufficient evidence to support a conclusion that the population mean is greater than 166.3 lb, as in the National Transportation and Safety Board's recommendation.

Normal Distribution Versus Student *t* Distribution

The critical value in the preceding example was $t = 1.782$, but if the normal distribution were being used, the critical value would have been $z = 1.645$.

The Student *t* critical value is larger (farther to the right), showing that with the Student *t* distribution, the sample evidence must be more extreme before we can consider it to be significant.

P-Value Method

- Use software or a TI-83/84 Plus calculator.
- If technology is not available, use Table A-3 to identify a range of *P*-values.

Example: **Assuming that neither software nor a TI-83 Plus calculator is available, use Table A-3 to find a range of values for the *P*-value corresponding to the given results.**

a) **In a left-tailed hypothesis test, the sample size is $n = 12$, and the test statistic is $t = -2.007$.**
b) **In a right-tailed hypothesis test, the sample size is $n = 12$, and the test statistic is $t = 1.222$.**
c) **In a two-tailed hypothesis test, the sample size is n = 12, and the test statistic is $t = -3.456$.**

Example: **Assuming that neither software nor a TI-83 Plus calculator is available, use Table A-3 to find a range of values for the *P*-value corresponding to the given results.**

Table A-3 Finding *P*-Values from Table A-3

	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
	Area in Two Tails				
Degrees of Freedom	0.01	0.02	0.05	0.10	0.20
•					
•					
•					
11	3.106	2.718	2.201	1.796	1.363
	•				
	•				

Example: **Assuming that neither software nor a TI-83 Plus calculator is available, use Table A-3 to find a range of values for the *P*-value corresponding to the given results.**

a) The test is a left-tailed test with test statistic $t = -2.007$, so the *P*-value is the area to the left of –2.007. Because of the symmetry of the t distribution, that is the same as the area to the right of +2.007. Any test statistic between 2.201 and 1.796 has a right-tailed *P*-value that is between 0.025 and 0.05. We conclude that

$$0.025 < P\text{-value} < 0.05.$$

Example: **Assuming that neither software nor a TI-83 Plus calculator is available, use Table A-3 to find a range of values for the *P*-value corresponding to the given results.**

b) The test is a right-tailed test with test statistic $t = 1.222$, so the *P*-value is the area to the right of 1.222. Any test statistic less than 1.363 has a right-tailed *P*-value that is greater than 0.10. We conclude that *P*-value > 0.10.

Example: **Assuming that neither software nor a TI-83 Plus calculator is available, use Table A-3 to find a range of values for the *P*-value corresponding to the given results.**

c) The test is a two-tailed test with test statistic

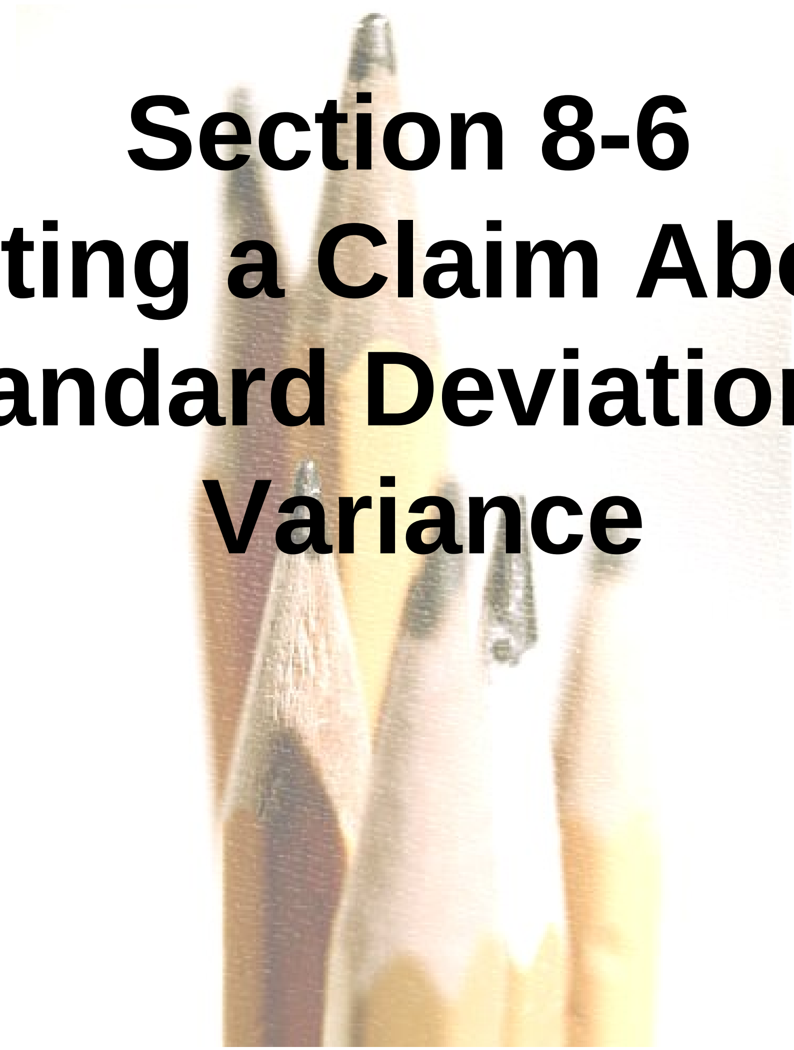
t = –3.456. The *P*-value is twice the area to the right of –3.456. Any test statistic greater than 3.106 has a two-tailed *P*-value that is less than 0.01. We conclude that *P*-value < 0.01.

Recap

In this section we have discussed:

- **Assumptions for testing claims about population means, σ unknown.**
- **Student *t* distribution.**
- ***P*-value method.**

Section 8-6
Testing a Claim About a Standard Deviation or Variance



Key Concept

This section introduces methods for testing a claim made about a population standard deviation σ or population variance σ^2. The methods of this section use the chi-square distribution that was first introduced in Section 7-5.

Requirements for Testing Claims About σ or σ^2

n = sample size

s = *sample* standard deviation

s^2 = *sample* variance

σ = claimed value of the *population* standard deviation

σ^2 = claimed value of the *population* variance

Requirements for Testing Claims About σ or σ^2

1. The sample is a simple random sample.

2. The population has a normal distribution. (This is a much stricter requirement than the requirement of a normal distribution when testing claims about means.)

Chi-Square Distribution

Test Statistic

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

P-Values and Critical Values for Chi-Square Distribution

- **Use Table A-4.**
- **The degrees of freedom = n –1.**

Caution

The χ^2 test of this section is not *robust* against a departure from normality, meaning that the test does not work well if the population has a distribution that is far from normal. The condition of a normally distributed population is therefore a much stricter requirement in this section than it was in Sections 8-4 and 8-5.

Properties of Chi-Square Distribution

- **All values of χ^2 are nonnegative, and the distribution is not symmetric (see Figure 8-13, following).**

- **There is a different distribution for each number of degrees of freedom (see Figure 8-14, following).**

- **The critical values are found in Table A-4 using n – 1 degrees of freedom.**

Properties of Chi-Square Distribution - cont

Properties of the Chi-Square Distribution

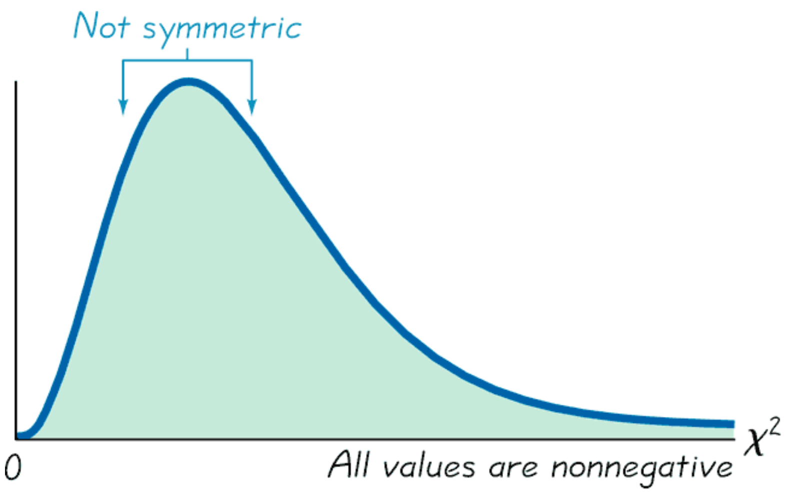
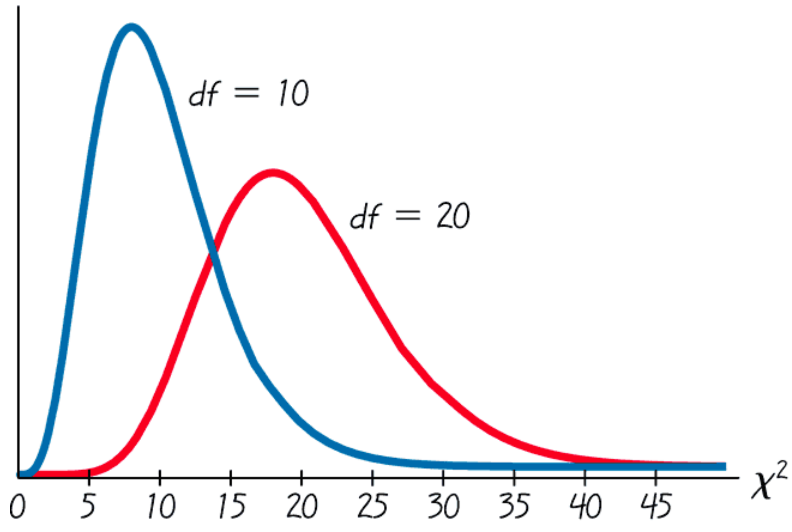


Figure 8-13

Chi-Square Distribution for 10 and 20 df



Different distribution for each number of df.

Figure 8-14

Table A-4

Table A-4 is based on cumulative areas from the right (unlike the entries in Table A-2, which are cumulative areas from the left). Critical values are found in Table A-4 by first locating the row corresponding to the appropriate number of degrees of freedom (where df = n –1). Next, the significance level α is used to determine the correct column. The following examples are based on a significance level of α = 0.05, but any other significance level can be used in a similar manner.

Table A-4

Right-tailed test:

Because the area to the right of the critical value is 0.05, locate 0.05 at the top of Table A-4.

Left-tailed test:

With a left-tailed area of 0.05, the area to the right of the critical value is 0.95, so locate 0.95 at the top of Table A-4.

Table A-4

Two-tailed test:

Unlike the normal and Student *t* distributions, the critical values in this χ^2 test will be two different positive values (instead of something like ±1.96). Divide a significance level of 0.05 between the left and right tails, so the areas to the right of the two critical values are 0.975 and 0.025, respectively. Locate 0.975 and 0.025 at the top of Table A-4

Example:

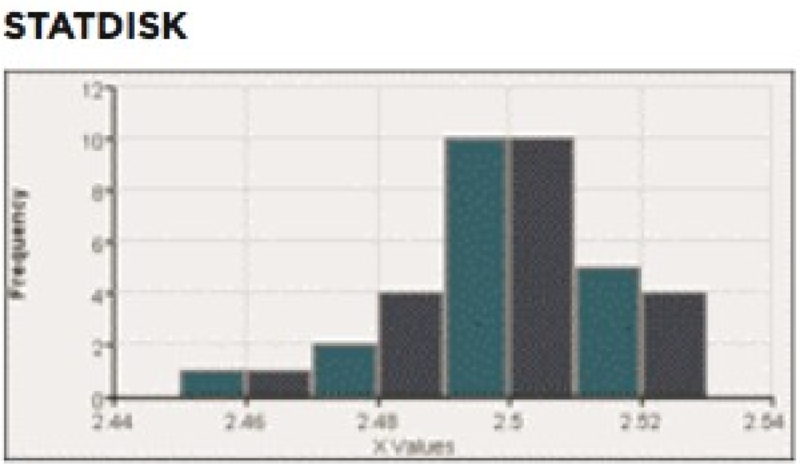
A common goal in business and industry is to improve the quality of goods or services by reducing variation. Quality control engineers want to ensure that a product has an acceptable mean, but they also want to produce items of consistent quality so that there will be few defects. If weights of coins have a specified mean but too much variation, some will have weights that are too low or too high, so that vending machines will not work correctly (unlike the stellar performance that they now provide).

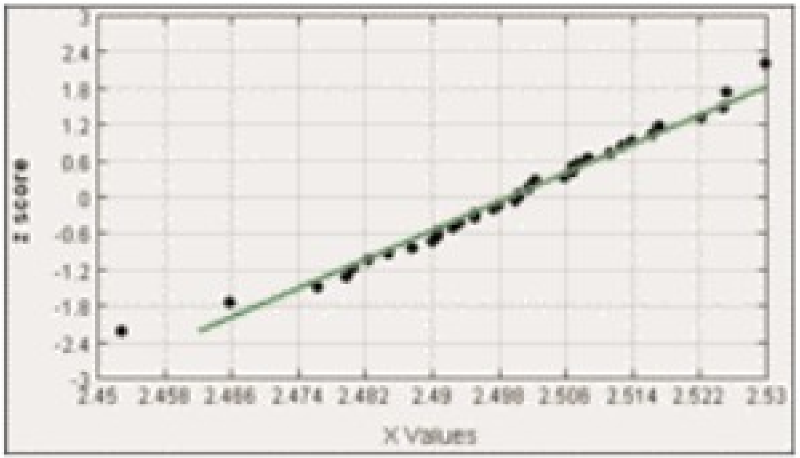
Example:

Consider the simple random sample of the 37 weights of post-1983 pennies listed in Data Set 20 in Appendix B. Those 37 weights have a mean of 2.49910 g and a standard deviation of 0.01648 g. U.S. Mint specifications require that pennies be manufactured so that the mean weight is 2.500 g. A hypothesis test will verify that the sample appears to come from a population with a mean of 2.500 g as required, but use a 0.05 significance level to test the claim that the population of weights has a standard deviation less than the specification of 0.0230 g.

Example:

Requirements are satisfied: simple random sample; and STATDISK generated the histogram and quantile plot - sample appears to come from a population having a normal distribution.





Example:

Step 1: Express claim as $\sigma < 0.0230$ g

Step 2: If $\sigma < 0.0230$ g is false, then $\sigma \geq 0.0230$ g

Step 3: $\sigma < 0.0230$ g does not contain equality so it is the alternative hypothesis; null hypothesis is $\sigma = 0.0230$ g

$$H_0: \sigma = 0.0230 \text{ g}$$
$$H_1: \sigma < 0.0230 \text{ g}$$

Step 4: significance level is $\alpha = 0.05$

Step 5: Claim is about σ so use chi-square

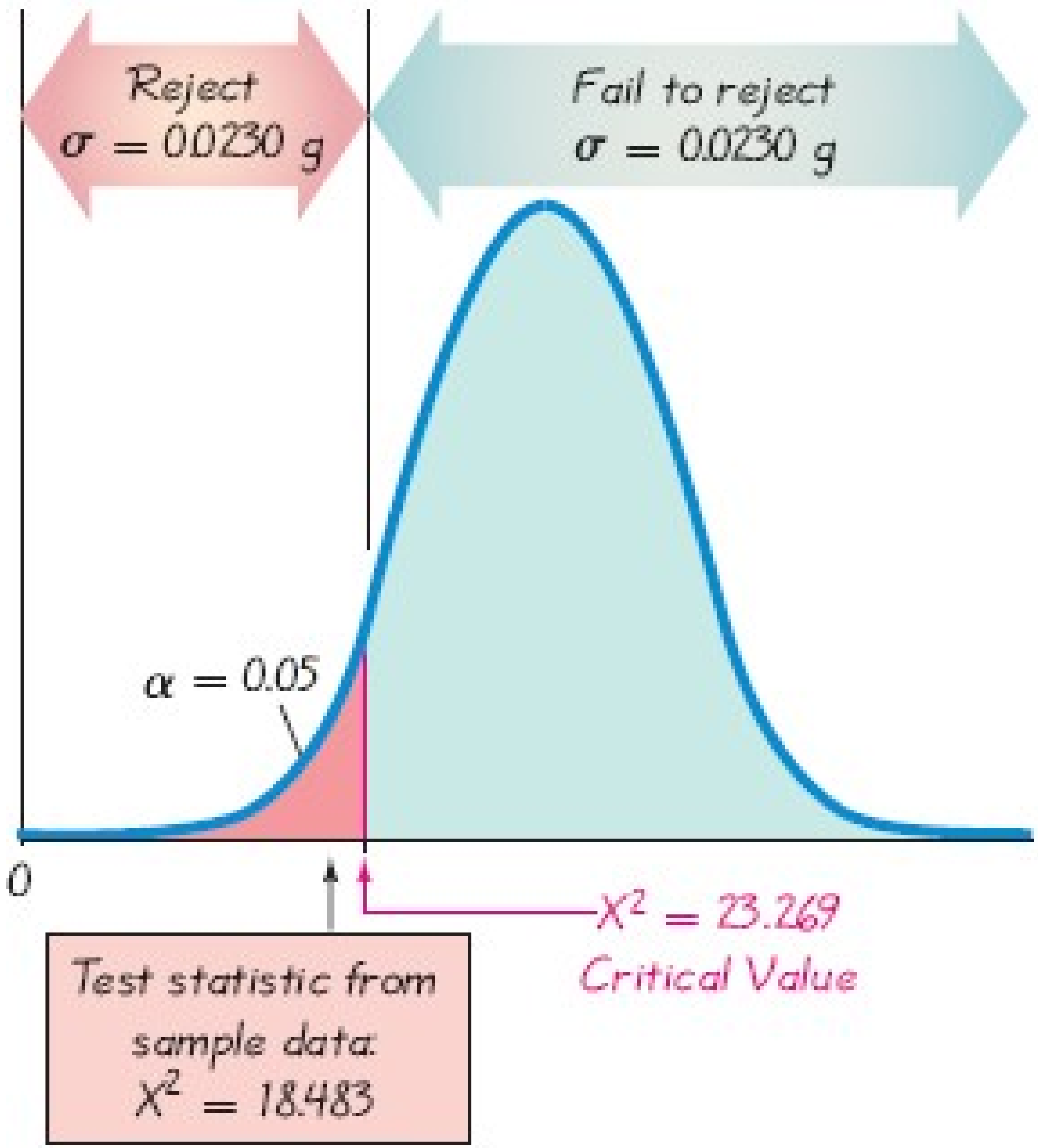
Example:

Step 6: The test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(37-1)(0.01648)^2}{0.0230^2} = 18.483$$

The critical value from Table A-4 corresponds to 36 degrees of freedom and an "area to the right" of 0.95 (based on the significance level of 0.05 for a left-tailed test). Table A-4 does not include 36 degrees of freedom, but Table A-4 shows that the critical value is between 18.493 and 26.509. (Using technology, the critical value is 23.269.)

Example:



Reject
$\sigma = 0.0230$ g
Fail to reject
$\sigma = 0.0230$ g
$\alpha = 0.05$
0
$X^2 = 23.269$
Critical Value
Test statistic from sample data:
$X^2 = 18.483$

Example:

Step 7: Because the test statistic is in the critical region, reject the null hypothesis.

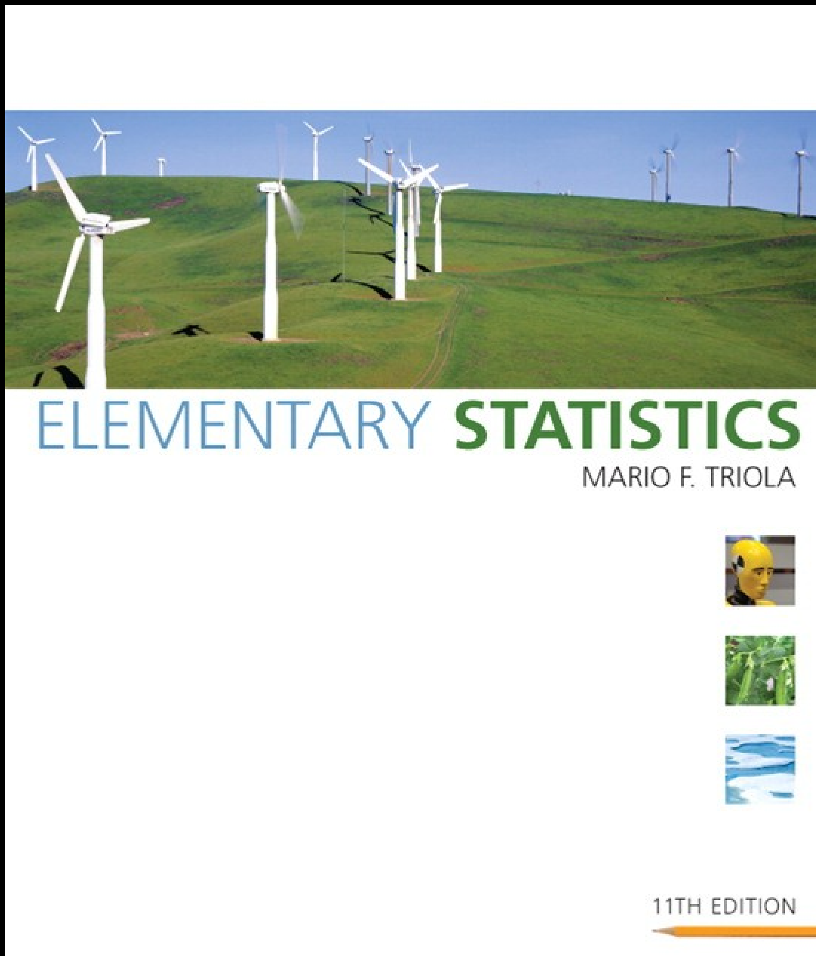
There is sufficient evidence to support the claim that the standard deviation of weights is less than 0.0230 g. It appears that the variation is less than 0.0230 g as specified, so the manufacturing process is acceptable.

Recap

In this section we have discussed:

- **Tests for claims about standard deviation and variance.**
- **Test statistic.**
- **Chi-square distribution.**
- **Critical values.**

Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 9
Inferences from Two Samples

9-1 Review and Preview

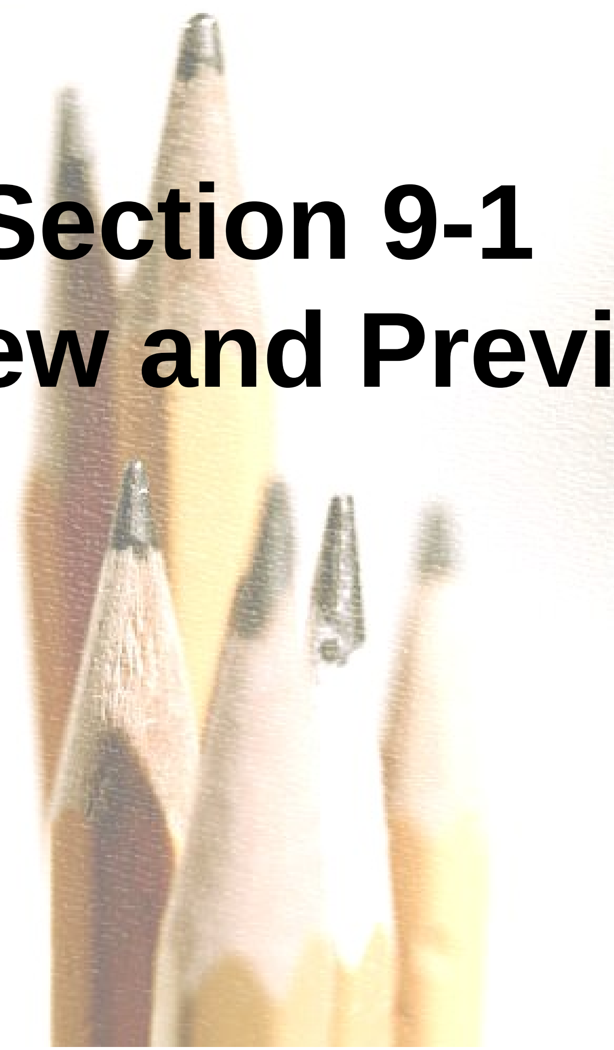
9-2 Inferences About Two Proportions

9-3 Inferences About Two Means: Independent Samples

9-4 Inferences from Dependent Samples

9-5 Comparing Variation in Two Samples

Section 9-1
Review and Preview



Review

In Chapters 7 and 8 we introduced methods of *inferential statistics*. In Chapter 7 we presented methods of constructing confidence interval estimates of population parameters. In Chapter 8 we presented methods of testing claims made about population parameters. Chapters 7 and 8 both involved methods for dealing with a sample from a single population.

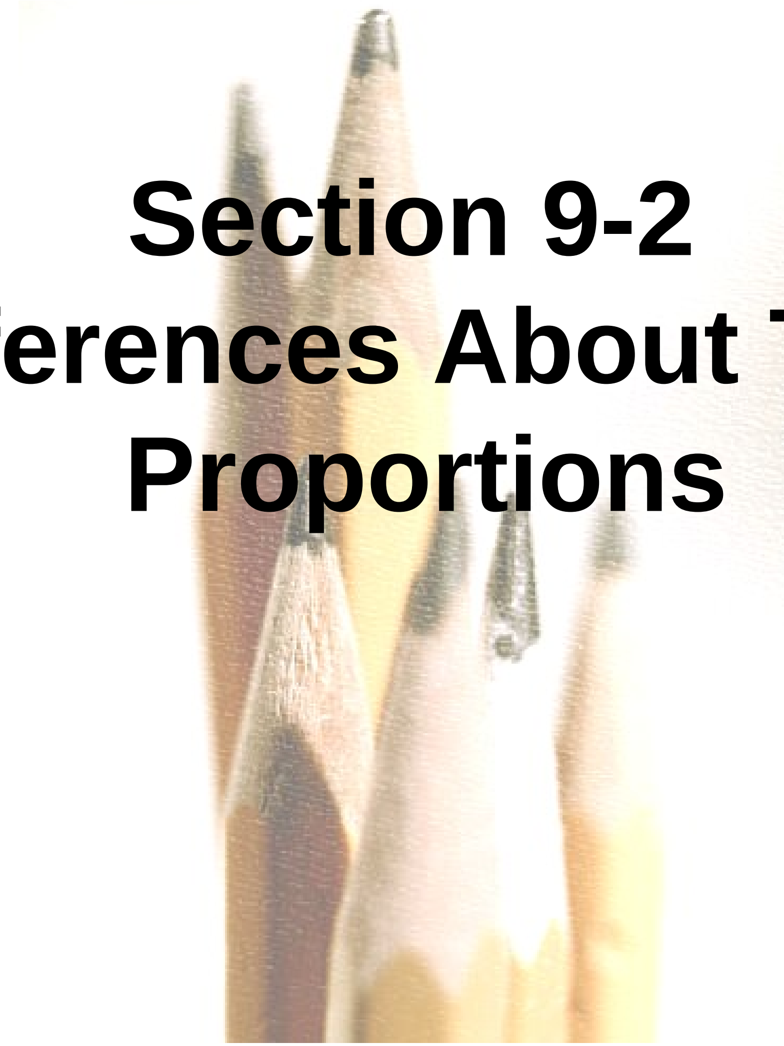
Preview

The objective of this chapter is to extend the methods for estimating values of population parameters and the methods for testing hypotheses to situations involving two sets of sample data instead of just one.
The following are examples typical of those found in this chapter, which presents methods for using sample data from two populations so that inferences can be made about those populations.

Preview

- **Test the claim that when college students are weighed at the beginning and end of their freshman year, the differences show a mean weight gain of 15 pounds (as in the "Freshman 15" belief).**
- **Test the claim that the proportion of children who contract polio is less for children given the Salk vaccine than for children given a placebo.**
- **Test the claim that subjects treated with Lipitor have a mean cholesterol level that is lower than the mean cholesterol level for subjects given a placebo.**

Section 9-2
Inferences About Two Proportions



Key Concept

In this section we present methods for (1) testing a claim made about the two population proportions and (2) constructing a confidence interval estimate of the difference between the two population proportions. This section is based on proportions, but we can use the same methods for dealing with probabilities or the decimal equivalents of percentages.

Notation for Two Proportions

For population 1, we let:

p_1 = **population proportion**

n_1 = **size of the sample**

x_1 = **number of successes in the sample**

$\hat{p}_1 = \dfrac{x_1}{n_1}$ **(the sample proportion)**

$\hat{q}_1 = 1 - \hat{p}_1$

The corresponding notations apply to p_2, n_2, x_2, $\hat{p}_2$. and $\hat{q}_2$, which come from population 2.

Pooled Sample Proportion

- The **pooled sample proportion** is denoted by $\bar{p}$ and is given by:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

- We denote the complement of $\bar{p}$ by $\bar{q}$, so $\bar{q} = 1 - \bar{p}$

Requirements

1. We have proportions from two **independent** simple random samples.

2. For each of the two samples, the number of successes is at least 5 and the number of failures is at least 5.

Test Statistic for Two Proportions

For $H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$, $H_1: p_1 < p_2$, $H_1: p_1 > p_2$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1} + \dfrac{\bar{p}\bar{q}}{n_2}}}$$

Test Statistic for Two Proportions - cont

For $H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$, $H_1: p_1 < p_2$, $H_1: p_1 > p_2$

where $p_1 - p_2 = 0$ (assumed in the null hypothesis)

$$\hat{p}_1 = \frac{x_1}{n_1} \quad \text{and} \quad \hat{p}_2 = \frac{x_2}{n_2}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \quad \text{and} \quad \bar{q} = 1 - \bar{p}$$

Test Statistic for Two Proportions - cont

P-value: Use Table A-2. (Use the computed value of the test statistic z and find its *P*-value by following the procedure summarized by Figure 8-5 in the text.)

Critical values: Use Table A-2. (Based on the significance level α, find critical values by using the procedures introduced in Section 8-2 in the text.)

Confidence Interval Estimate of $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

where $$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Example:

The table below lists results from a simple random sample of front-seat occupants involved in car crashes. Use a 0.05 significance level to test the claim that the fatality rate of occupants is lower for those in cars equipped with airbags.

	Airbag Available	No Airbag Available
Occupant Fatalities	41	52
Total number of occupants	11,541	9,853

Example:

Requirements are satisfied: two simple random samples, two samples are independent; Each has at least 5 successes and 5 failures (11,500, 41; 9801, 52).
Use the *P*-value method.

Step 1: Express the claim as $p_1 < p_2$.

Step 2: If $p_1 < p_2$ is false, then $p_1 \geq p_2$.

Step 3: $p_1 < p_2$ does not contain equality so it is the alternative hypothesis. The null hypothesis is the statement of equality.

Pearson Education

Example:

H_0: $p_1 = p_2$ H_a: $p_1 < p_2$ (original claim)

Step 4: Significance level is 0.05

Step 5: Use normal distribution as an approximation to the binomial distribution. Estimate the common values of p_1 and p_2 as follows:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{41 + 52}{11{,}541 + 9{,}853} = 0.004347$$

With $\bar{p} = 0.004347$ **it follows** $\bar{q} = 0.995653$

Example:

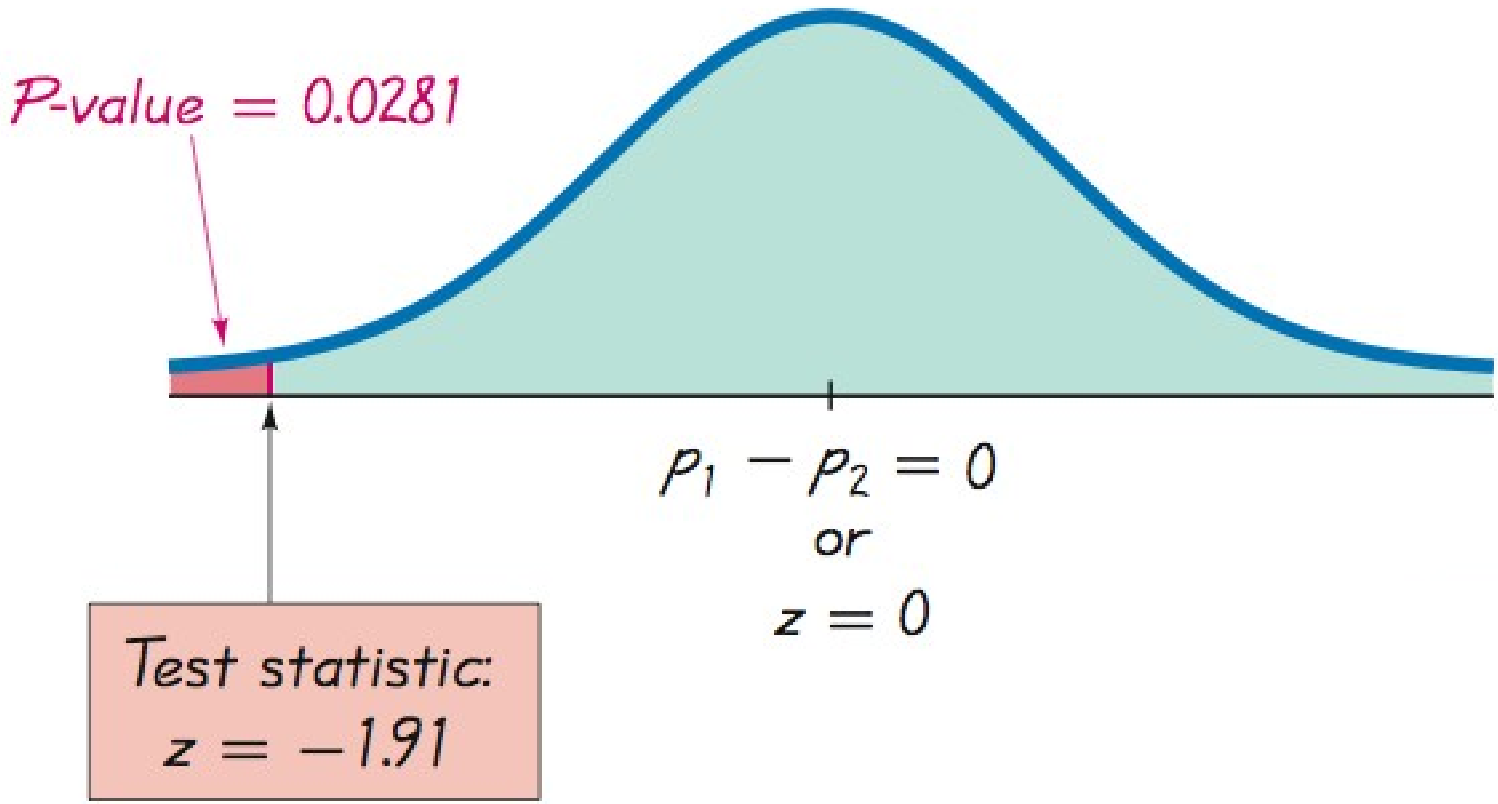
Step 6: Find the value of the test statistic.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1} + \dfrac{\bar{p}\bar{q}}{n_2}}}$$

$$= \frac{\left(\dfrac{41}{11{,}541} - \dfrac{52}{9{,}853}\right) - 0}{\sqrt{\dfrac{(0.004347)(0.995653)}{11{,}541} + \dfrac{(0.004347)(0.995653)}{9{,}853}}}$$

$$z = -1.91$$

Example:



Left-tailed test. Area to left of *z* = –1.91 is 0.0281 (Table A-2), so the *P*-value is 0.0281.

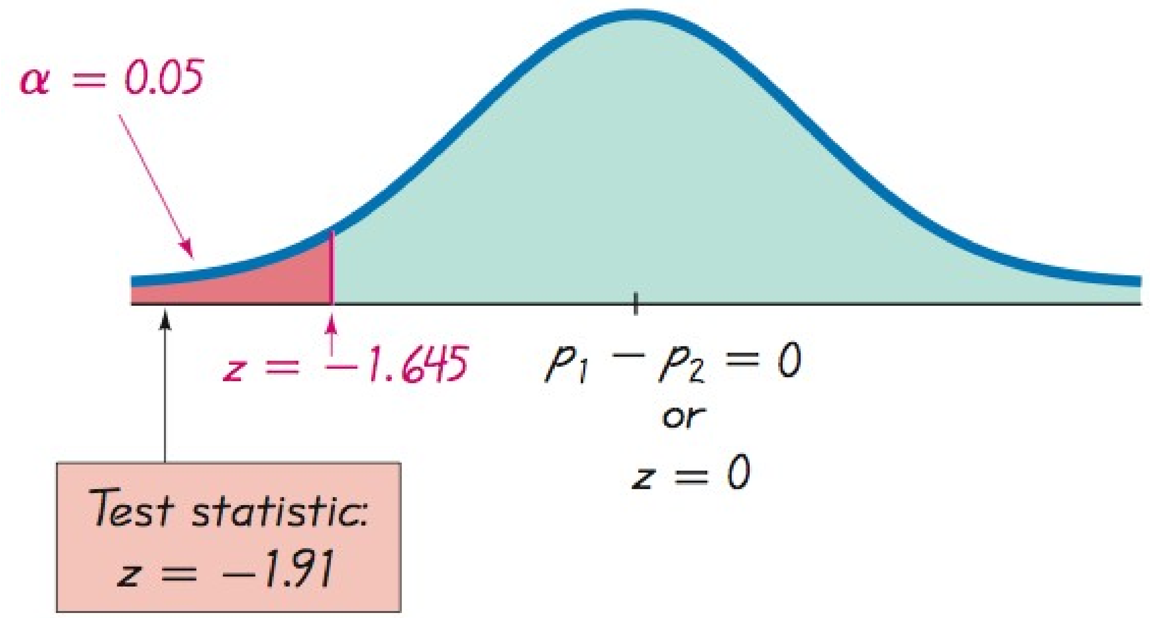
Example:

Step 7: Because the *P*-value of 0.0281 is less than the significance level of $\alpha = 0.05$, we reject the null hypothesis of $p_1 = p_2$.

Because we reject the null hypothesis, we conclude that there is sufficient evidence to support the claim that the proportion of accident fatalities for occupants in cars with airbags is less than the proportion of fatalities for occupants in cars without airbags. Based on these results, it appears that airbags are effective in saving lives.

Example: Using the Traditional Method

With a significance level of $\alpha = 0.05$ in a left- tailed test based on the normal



distribution, we refer to Table A-2 and find that an area of $\alpha = 0.05$ in the left tail corresponds to the critical value of $z = –1.645$. The test statistic of does fall in the critical region bounded by the critical value of $z = –1.645$. We again reject the null hypothesis.

Caution

When testing a claim about two population proportions, the *P*-value method and the traditional method are equivalent, but they are *not* equivalent to the confidence interval method. If you want to test a claim about two population proportions, use the *P*-value method or traditional method; if you want to estimate the difference between two population proportions, use a confidence interval.

Example:

Use the sample data given in the preceding Example to construct a 90% confidence interval estimate of the difference between the two population proportions. (As shown in Table 8-2 on page 406, the confidence level of 90% is comparable to the significance level of $\alpha = 0.05$ used in the preceding left-tailed hypothesis test.) What does the result suggest about the effectiveness of airbags in an accident?

Example:

Requirements are satisfied as we saw in the preceding example.

90% confidence interval: $z_{\alpha/2} = 1.645$

Calculate the margin of error, *E*

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$= 1.645\sqrt{\frac{\left(\frac{41}{11,541}\right)\left(\frac{11,500}{11,541}\right)}{11,541} + \frac{\left(\frac{52}{9,853}\right)\left(\frac{9801}{9,853}\right)}{9,853}}$$

$$= 0.001507$$

Example:

Construct the confidence interval

$$\left(\hat{p}_1 - \hat{p}_2\right) - E < \left(p_1 - p_2\right) < \left(\hat{p}_1 - \hat{p}_2\right) + E$$

$$\left(0.003553 - 0.005278\right) - 0.001507$$
$$< \left(p_1 - p_2\right) <$$
$$\left(0.003553 - 0.005278\right) + 0.001507$$

$$-0.00323 < \left(p_1 - p_2\right) < -0.000218$$

Example:

The confidence interval limits do not contain 0, implying that there is a significant difference between the two proportions. The confidence interval suggests that the fatality rate is lower for occupants in cars with air bags than for occupants in cars without air bags. The confidence interval also provides an estimate of the amount of the difference between the two fatality rates.

Why Do the Procedures of This Section Work?

The distribution of $\hat{p}_1$ can be approximated by a normal distribution with mean p_1, standard deviation $\sqrt{p_1 q_1 / n_1}$, and variance $p_1 q_1 / n_1$.

The difference $\hat{p}_1 - \hat{p}_2$ can be approximated by a normal distribution with mean $p_1 - p_2$ and variance

$$\sigma^2_{(\hat{p}_1 - \hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

The variance of the *differences* between two independent random variables is the *sum* of their individual variances.

Why Do the Procedures of This Section Work?

The preceding variance leads to

$$\sigma_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{\bar{p}_1 \bar{q}_1}{n_1} + \frac{\bar{p}_2 \bar{q}_2}{n_2}}$$

We now know that the distribution of $p_1 - p_2$ is approximately normal, with mean $p_1 - p_2$ and standard deviation as shown above, so the z test statistic has the form given earlier.

Why Do the Procedures of This Section Work?

When constructing the confidence interval estimate of the difference between two proportions, we don't assume that the two proportions are equal, and we estimate the standard deviation as

$$\sigma = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Pearson Education

Why Do the Procedures of This Section Work?

In the test statistic

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}}$$

use the positive and negative values of *z* (for two tails) and solve for $p_1 - p_2$. The results are the limits of the confidence interval given earlier.

Recap

In this section we have discussed:

- **Requirements for inferences about two proportions.**
- **Notation.**
- **Pooled sample proportion.**
- **Hypothesis tests.**

Section 9-3
Inferences About Two Means: Independent Samples

Key Concept

This section presents methods for using sample data from two independent samples to test hypotheses made about two population means or to construct confidence interval estimates of the difference between two population means.

Key Concept

In Part 1 we discuss situations in which the standard deviations of the two populations are unknown and are not assumed to be equal. In Part 2 we discuss two other situations: (1) The two population standard deviations are both known; (2) the two population standard deviations are unknown but are assumed to be equal. Because is typically unknown in real situations, most attention should be given to the methods described in Part 1.

Pearson Education

Part 1: Independent Samples with σ_1 and σ_2 Unknown and Not Assumed Equal

Pearson Education

Definitions

Two samples are independent if the sample values selected from one population are not related to or somehow paired or matched with the sample values from the other population.

Two samples are dependent if the sample values are *paired*. (That is, each pair of sample values consists of two measurements from the same subject (such as before/after data), or each pair of sample values consists of matched pairs (such as husband/wife data), where the matching is based on some inherent relationship.)

Pearson Education

Notation

μ_1 = population mean

σ_1 = population standard deviation

n_1 = size of the first sample

$\bar{x}_1$ = sample mean

s_1 = sample standard deviation

Corresponding notations for μ_2, σ_2, s_2, $\bar{x}_2$ and n_2 apply to population 2.

Requirements

1. σ_1 an σ_2 are unknown and no assumption is made about the equality of σ_1 and σ_2 .

2. The two samples are independent.

3. Both samples are simple random samples.

4. Either or both of these conditions are satisfied: The two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

Hypothesis Test for Two Means: Independent Samples

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(where $\mu_1 - \mu_2$ is often assumed to be 0)

Hypothesis Test - cont

Test Statistic for Two Means: Independent Samples

Degrees of freedom: In this book we use this simple and conservative estimate: df = smaller of $n_1 - 1$ and $n_2 - 1$.

***P*-values:** Refer to Table A-3. Use the procedure summarized in Figure 8-5.

Critical values: Refer to Table A-3.

Confidence Interval Estimate of $\mu_1 - \mu_2$: Independent Samples

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

where $E = t_{\alpha/2} \sqrt{\dfrac{s_1^{\,2}}{n_1} + \dfrac{s_2^{\,2}}{n_2}}$

where df = smaller $n_1 - 1$ and $n_2 - 1$

Caution

Before conducting a hypothesis test, consider the context of the data, the source of the data, the sampling method, and explore the data with graphs and descriptive statistics. Be sure to verify that the requirements are satisfied.

Example:

A headline in *USA Today* proclaimed that "Men, women are equal talkers." That headline referred to a study of the numbers of words that samples of men and women spoke in a day. Given below are the results from the study. Use a 0.05 significance level to test the claim that men and women speak the same mean number of words in a day. Does there appear to be a difference?

Number of Words Spoken in a Day

Men	Women
$n_1 = 186$	$n_2 = 210$
$\bar{x}_1 = 15{,}668.5$	$\bar{x}_2 = 16{,}215.0$
$s_1 = 8632.5$	$s_2 = 7301.2$

Example:

Requirements are satisfied: two population standard deviations are not known and not assumed to be equal, independent samples, simple random samples, both samples are large.

Step 1: Express claim as $\mu_1 = \mu_2$.

Step 2: If original claim is false, then $\mu_1 \neq \mu_2$.

Step 3: Alternative hypothesis does not contain equality, null hypothesis does.

$H_0 : \mu_1 = \mu_2$ **(original claim)** $\quad H_a : \mu_1 \neq \mu_2$

Proceed assuming $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$.

Example:

Step 4: Significance level is 0.05

Step 5: Use a *t* distribution

Step 6: Calculate the test statistic

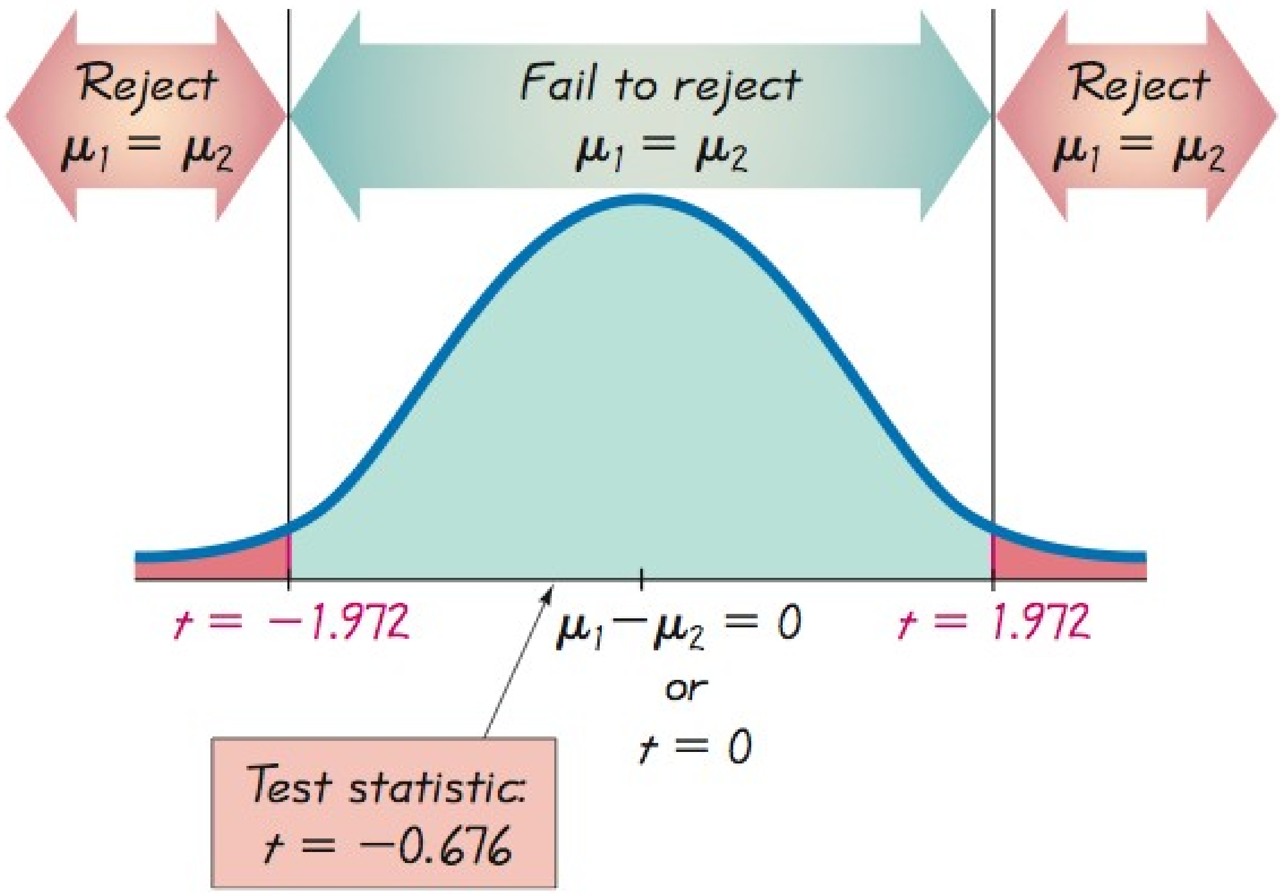
$$t = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{\left(15,668.5 - 16,215.0\right) - 0}{\sqrt{\frac{8632.5^2}{186} + \frac{7301.2^2}{210}}} = -0.676$$

Example:

Use Table A-3: area in two tails is 0.05, df = 185, which is not in the table, the closest value is

$t = \pm 1.972$



Example:

Step 7: **Because the test statistic does not fall within the critical region, fail to reject the null hypothesis:**
$\mu_1 = \mu_2$ (or $\mu_1 - \mu_2 = 0$).

There is not sufficient evidence to warrant rejection of the claim that men and women speak the same mean number of words in a day. There does not appear to be a significant difference between the two means.

Example:

Using the sample data given in the previous Example, construct a 95% confidence interval estimate of the difference between the mean number of words spoken by men and the mean number of words spoken by women.

Number of Words Spoken in a Day					
Men			Women		
n_1	=	186	n_2	=	210
$\bar{x}_1$	=	15,668.5	$\bar{x}_2$	=	16,215.0
s_1	=	8632.5	s_2	=	7301.2

Example:

Requirements are satisfied as it is the same data as the previous example.

Find the margin of Error, *E*; use $t_{\alpha/2}$ = 1.972

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1.972\sqrt{\frac{8632.5^2}{186} + \frac{7301.2^2}{210}} = 1595.4$$

Construct the confidence interval use *E* = 1595.4 and $\bar{x}_1 = 15{,}668.5$ and $\bar{x}_2 = 16{,}215.0.$

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

$$-2141.9 < (\mu_1 - \mu_2) < 1048.9$$

Example:

Step 4: Significance level is 0.05

Step 5: Use a *t* distribution

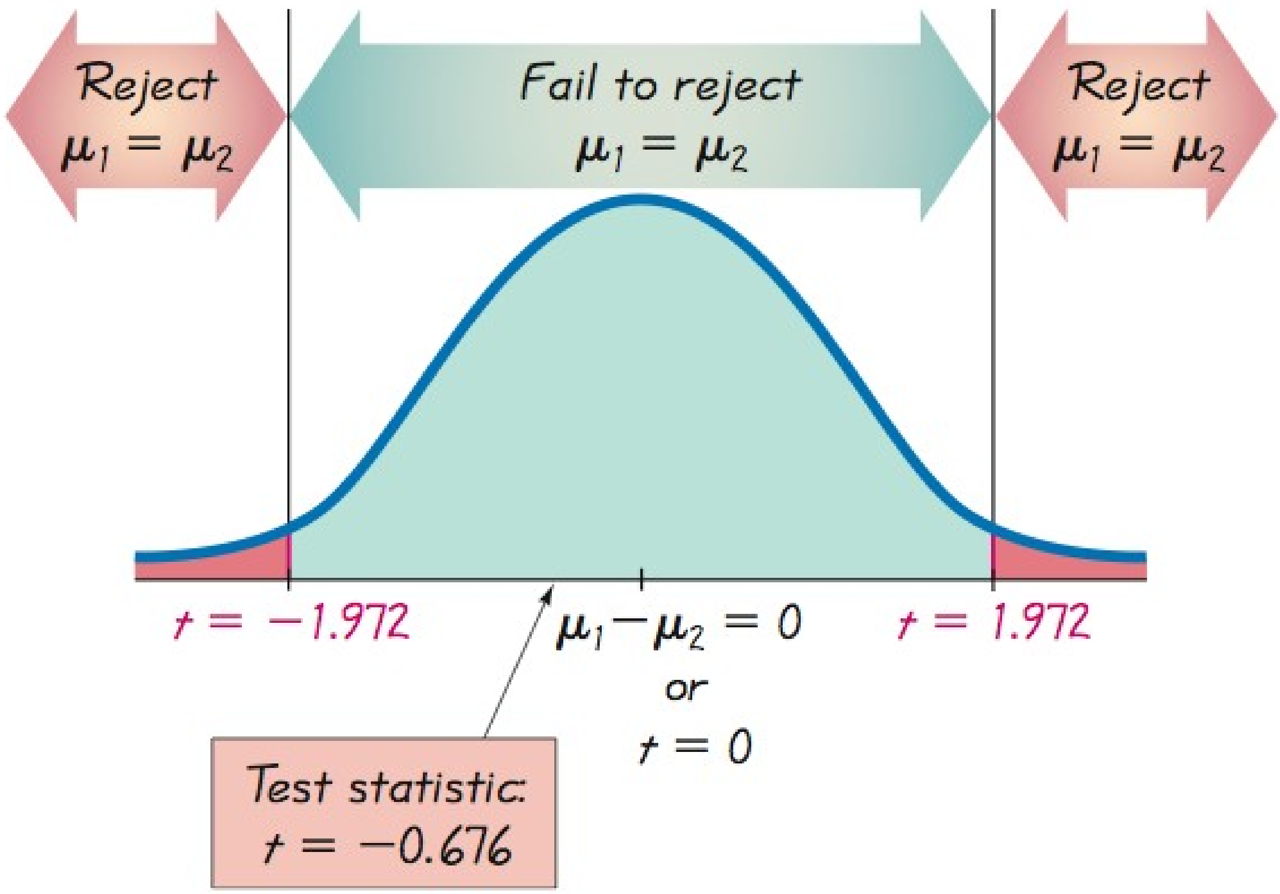
Step 6: Calculate the test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

$$= \frac{(15{,}668.5 - 16{,}215.0) - 0}{\sqrt{\dfrac{8632.5^2}{186} + \dfrac{7301.2^2}{210}}} = -0.676$$

Example:

Use Table A-3: area in two tails is 0.05, df = 185, which is not in the table, the closest value is
$t = \pm 1.972$



Example:

Step 7: Because the test statistic does not fall within the critical region, fail to reject the null hypothesis:
$\mu_1 = \mu_2$ (or $\mu_1 - \mu_2 = 0$).

There is not sufficient evidence to warrant rejection of the claim that men and women speak the same mean number of words in a day. There does not appear to be a significant difference between the two means.

Part 2: Alternative Methods

Independent Samples with σ_1 and σ_2 Known.

Requirements

1. The two population standard deviations are both known.

2. The two samples are independent.

3. Both samples are simple random samples.

4. Either or both of these conditions are satisfied: The two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

Hypothesis Test for Two Means: Independent Samples with σ_1 and σ_2 Both Known

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

***P*-values** **and critical values: Refer to Table A-2.**

Confidence Interval: Independent Samples with σ_1 and σ_2 Both Known

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

where $E = z_{\alpha/2} \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Methods for Inferences About Two Independent Means

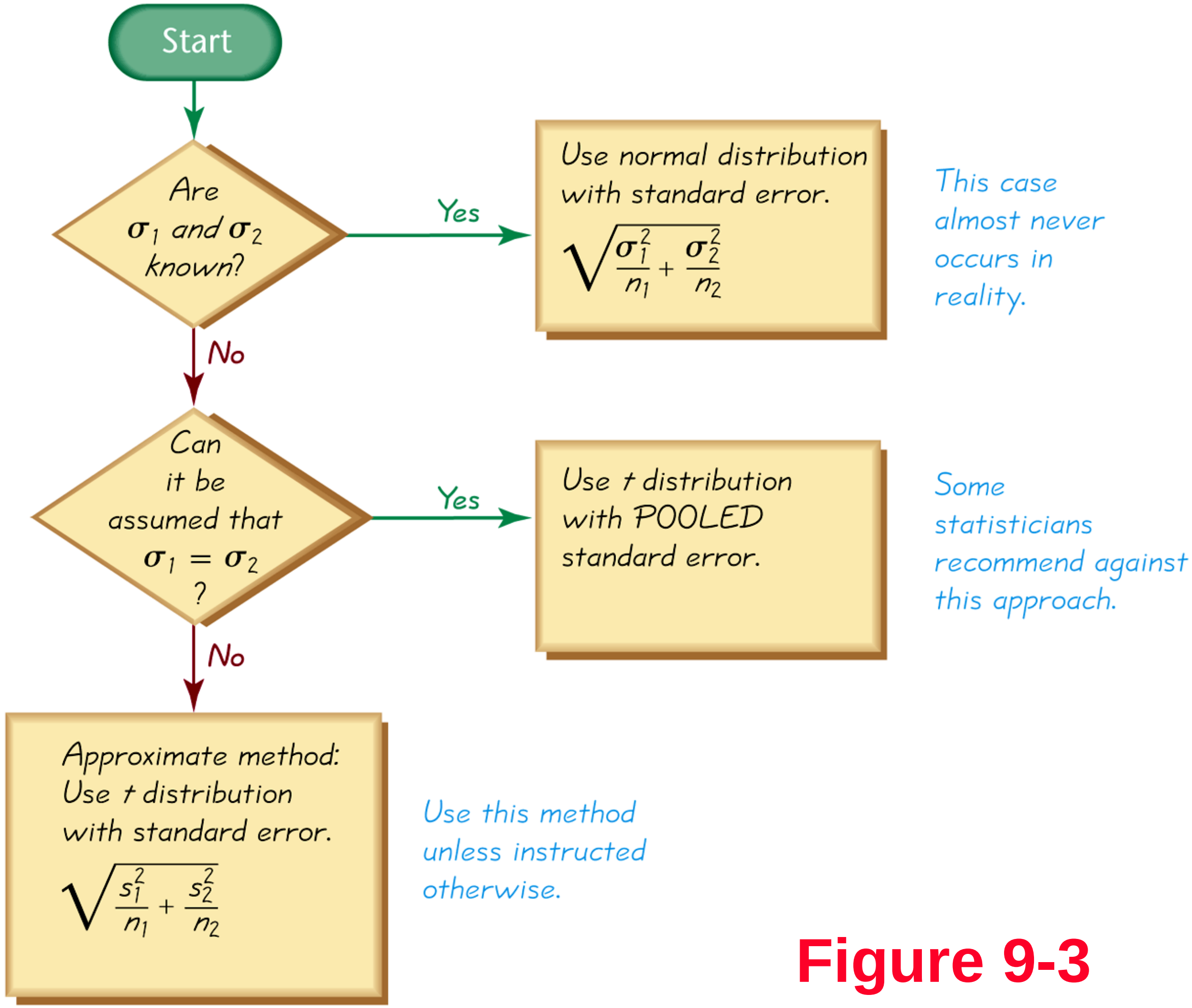


Figure 9-3

Assume that $\sigma_1 = \sigma_2$ and Pool the Sample Variances.

Requirements

1. The two population standard deviations are not known, but they are assumed to be equal. That is $\sigma_1 = \sigma_2$.

2. The two samples are independent.

3. Both samples are simple random samples.

4. Either or both of these conditions are satisfied: The two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

Hypothesis Test Statistic for Two Means: Independent Samples and $\sigma_1 = \sigma_2$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$$

Where

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

and the number of degrees of freedom is df = $n_1 + n_2 - 2$

Confidence Interval Estimate of $\mu_1 - \mu_2$: Independent Samples with $\sigma_1 = \sigma_2$

$$(\bar{x}_1 - \bar{x}_2) - E \; < \; (\mu_1 - \mu_2) \; < \; (\bar{x}_1 - \bar{x}_2) + E$$

where $E = t_{\alpha/2} \sqrt{\dfrac{s_p^{\,2}}{n_1} + \dfrac{s_p^{\,2}}{n_2}}$

and number of degrees of freedom is df = $n_1 + n_2 - 2$

Strategy

Unless instructed otherwise, use the following strategy:

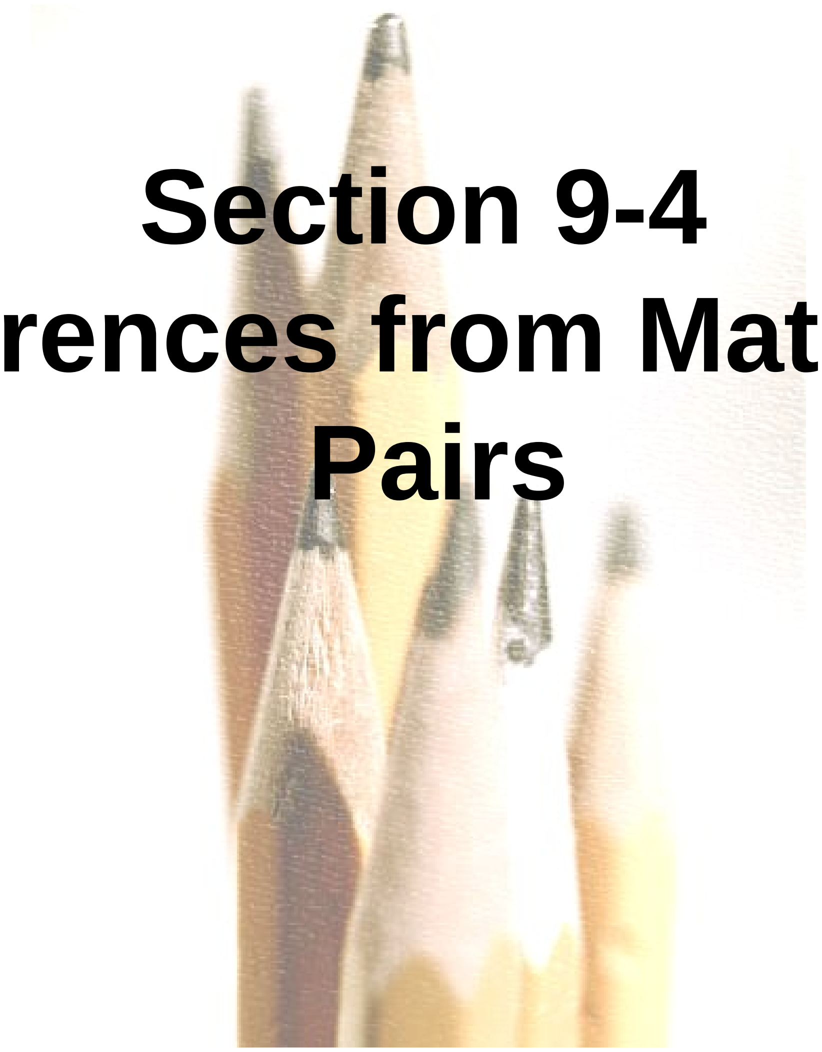
Assume that σ_1 and σ_2 are unknown, do not assume that $\sigma_1 = \sigma_2$, and use the test statistic and confidence interval given in Part 1 of this section. (See Figure 9-3.)

Recap

In this section we have discussed:

- **Independent samples with the standard deviations unknown and not assumed equal.**
- **Alternative method where standard deviations are known**
- **Alternative method where standard deviations are assumed equal and sample variances are pooled.**

Section 9-4
Inferences from Matched Pairs



Key Concept

In this section we develop methods for testing hypotheses and constructing confidence intervals involving the mean of the differences of the values from two dependent populations.

With dependent samples, there is some relationship whereby each value in one sample is paired with a corresponding value in the other sample.

Key Concept

Because the hypothesis test and confidence interval use the same distribution and standard error, they are equivalent in the sense that they result in the same conclusions. Consequently, the null hypothesis that the mean difference equals 0 can be tested by determining whether the confidence interval includes 0. There are no exact procedures for dealing with dependent samples, but the *t* distribution serves as a reasonably good approximation, so the following methods are commonly used.

Pearson Education

Notation for Dependent Samples

d = individual difference between the two values of a single matched pair

μ_d = mean value of the differences d for the **population** of paired data

$\bar{d}$ = mean value of the differences d for the paired **sample** data (equal to the mean of the $x - y$ values)

s_d = standard deviation of the differences d for the paired **sample** data

n = number of **pairs** of data.

Requirements

1. The sample data are dependent.

2. The samples are simple random samples.

3. Either or both of these conditions is satisfied: The number of pairs of sample data is large ($n > 30$) or the pairs of values have differences that are from a population having a distribution that is approximately normal.

Hypothesis Test Statistic for Matched Pairs

$$t = \frac{\bar{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}}$$

where degrees of freedom = n **– 1**

Pearson Education

Use Table A-3 (t-distribution).

Pearson Education

Confidence Intervals for Matched Pairs

$$\bar{d} - E < \mu_d < \bar{d} + E$$

where $E = t_{\alpha/2} \dfrac{s_d}{\sqrt{n}}$

Critical values of $t_{\alpha/2}$: Use Table A-3 with $n - 1$ degrees of freedom.

Example:

Data Set 3 in Appendix B includes measured weights of college students in September and April of their freshman year. Table 9-1 lists a small portion of those sample values. (Here we use only a small portion of the available data so that we can better illustrate the method of hypothesis testing.) Use the sample data in Table 9-1 with a 0.05 significance level to test the claim that for the population of students, the mean change in weight from September to April is equal to 0 kg.

Example:

Table 9-1 Weight (kg) Measurements of Students in Their Freshman Year

April weight	66	52	68	69	71
September weight	67	53	64	71	70
Difference d = (April weight) − (September weight)	−1	−1	4	−2	1

Requirements are satisfied: samples are dependent, values paired from each student; although a volunteer study, we'll proceed as if simple random sample and deal with this in the interpretation; STATDISK displays a histogram that is approximately normal

Example:

Weight gained = April weight – Sept. weight

μ_d denotes the mean of the "April – Sept." differences in weight; the claim is $\mu_d = 0$ kg

Step 1: claim is $\mu_d = 0$ kg

Step 2: If original claim is not true, we have $\mu_d \neq 0$ kg

Step 3: H_0: $\mu_d = 0$ kg original claim

H_1: $\mu_d \neq 0$ kg

Step 4: significance level is $\alpha = 0.05$

Step 5: use the student *t* distribution

Example:

Step 6: find values of *d* and s_d

differences are: –1, –1, 4, –2, 1

***d* = 0.2 and s_d = 2.4**

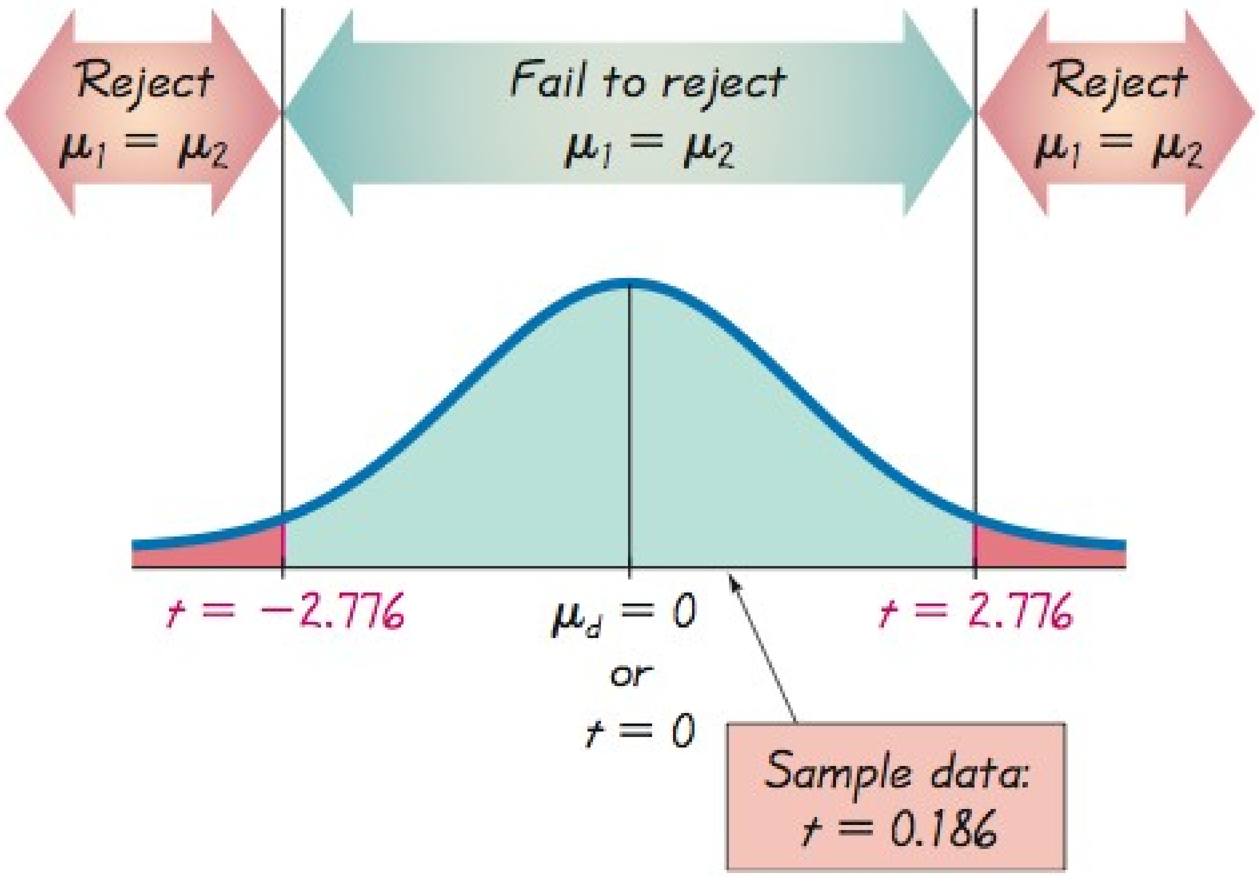
now find the test statistic

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{0.2 - 0}{\frac{2.4}{\sqrt{5}}} = 0.186$$

Table A-3: df = *n* – 1, area in two tails is 0.05, yields a critical value *t* = ±2.776

Example:

Step 7: Because the test statistic does not fall in the critical region, we fail to reject the null hypothesis.



Example:

We conclude that there is not sufficient evidence to warrant rejection of the claim that for the population of students, the mean change in weight from September to April is equal to 0 kg. Based on the sample results listed in Table 9-1, there does not appear to be a significant weight gain from September to April.

Example:

The conclusion should be qualified with the limitations noted in the article about the study. The requirement of a simple random sample is not satisfied, because only Rutgers students were used. Also, the study subjects are volunteers, so there is a potential for a self-selection bias. In the article describing the study, the authors cited these limitations and stated that “Researchers should conduct additional studies to better characterize dietary or activity patterns that predict weight gain among young adults who enter college or enter the workforce during this critical period in their lives.”

Pearson Education

Example:

The *P*-value method:

Using technology, we can find the *P*-value of 0.8605. (Using Table A-3 with the test statistic of $t = 0.186$ and 4 degrees of freedom, we can determine that the P-value is greater than 0.20.) We again fail to reject the null hypothesis, because the *P*-value is greater than the significance level of $\alpha = 0.05$.

Pearson Education

Example:

Confidence Interval method:

Construct a 95% confidence interval estimate of μ_d , which is the mean of the "April–September" weight differences of college students in their freshman year.

$\overline{d}$ **= 0.2, s_d = 2.4, n = 5, $t_{\alpha/2}$ = 2.776**

Find the margin of error, *E*

$$E = t_{\alpha/2} \frac{s_d}{\sqrt{n}} = 2.776 \cdot \frac{2.4}{\sqrt{5}} = 3.0$$

Example:

Construct the confidence interval:

$$\bar{d} - E < \mu_d < \bar{d} + E$$

$$0.2 - 3.0 < \mu_d < 0.2 + 3.0$$

$$-2.8 < \mu_d < 3.2$$

We have 95% confidence that the limits of 2.8 kg and 3.2 kg contain the true value of the mean weight change from September to April. In the long run, 95% of such samples will lead to confidence interval limits that actually do contain the true population mean of the differences.

Recap

In this section we have discussed:

- Requirements for inferences from matched pairs.
- Notation.
- Hypothesis test.
- Confidence intervals.

Section 9-5
Comparing Variation in Two Samples

Key Concept

This section presents the *F* test for comparing two population variances (or standard deviations). We introduce the *F* distribution that is used for the *F* test.

Note that the *F* test is very sensitive to departures from normal distributions.

Part 1

F test for Comparing Variances

Notation for Hypothesis Tests with Two Variances or Standard Deviations

s_1^2 = *larger* of two sample variances

n_1 = size of the sample with the *larger* variance

σ_1^2 = variance of the population from which the sample with the *larger* variance is drawn

$s_2^2, n_2, \text{and } \sigma_2^2$ are used for the other sample and population

Requirements

1. The two populations are **independent**.
2. The two samples are simple random samples.
3. The two populations are each **normally distributed.**

Test Statistic for Hypothesis Tests with Two Variances

$$F = \frac{s_1^2}{s_2^2}$$

Where s_1^2 is the larger of the two sample variances

Critical Values: Using Table A-5, we obtain critical *F* values that are determined by the following three values:

1. The significance level α
2. Numerator degrees of freedom = $n_1 - 1$
3. Denominator degrees of freedom = $n_2 - 1$

Properties of the *F* Distribution

- The *F* distribution is not symmetric.
- Values of the *F* distribution cannot be negative.
- The exact shape of the *F* distribution depends on the two different degrees of freedom.

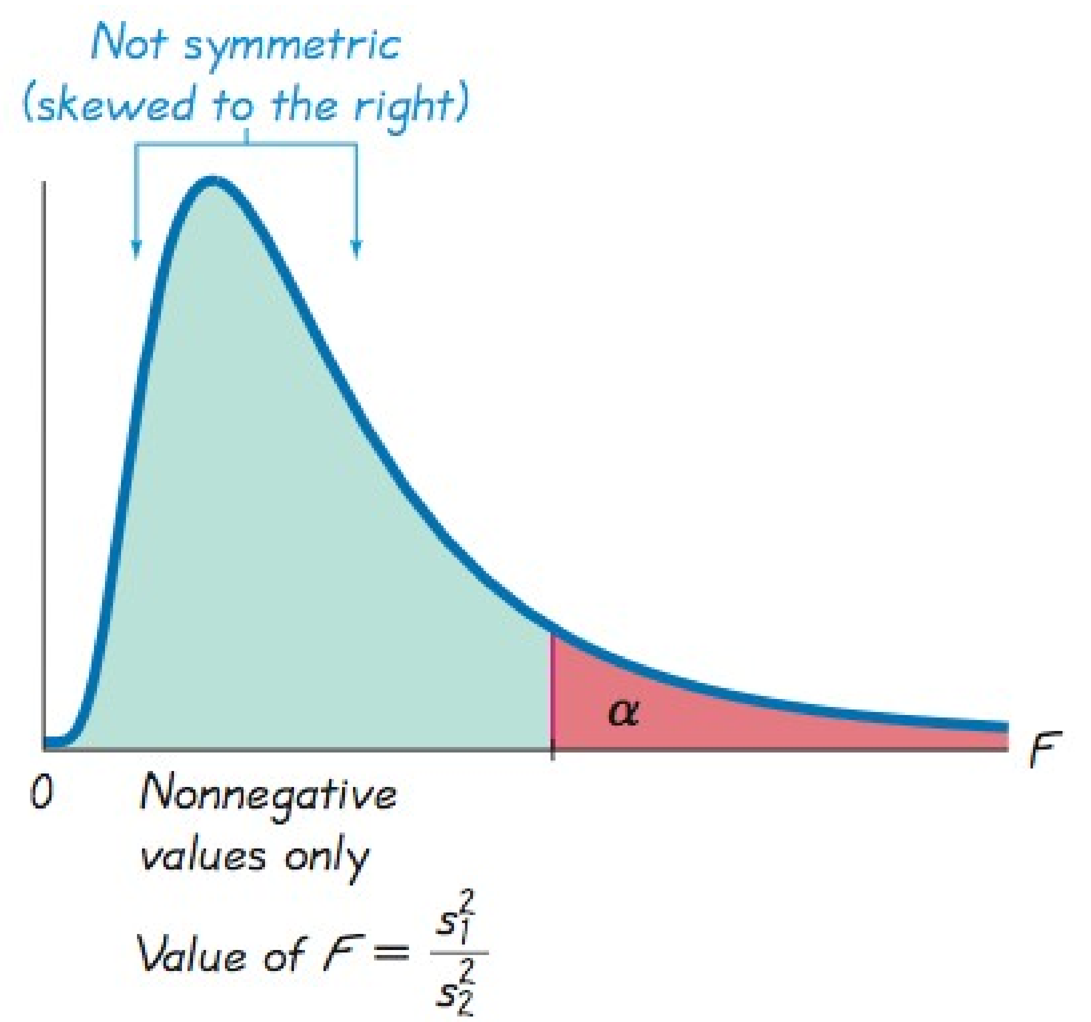
Finding Critical *F* Values

To find a critical *F* value corresponding to a 0.05 significance level, refer to Table A-5 and use the right-tail are of 0.025 or 0.05, depending on the type of test:

Two-tailed test: use 0.025 in right tail

One-tailed test: use 0.05 in right tail

Finding Critical *F* Values



Properties of the *F* Distribution - continued

If the two populations do have equal variances, then $F = \dfrac{s_1^2}{s_2^2}$ will be close to 1 because s_1^2 and s_2^2 are close in value.

Properties of the *F* Distribution - continued

If the two populations have radically different variances, then *F* will be a large number.

Remember, the larger sample variance will be s_1^2.

Conclusions from the *F* Distribution

Consequently, a value of *F* near 1 will be evidence in favor of the conclusion that $\sigma_1^2 = \sigma_2^2$.

But a large value of *F* will be evidence against the conclusion of equality of the population variances.

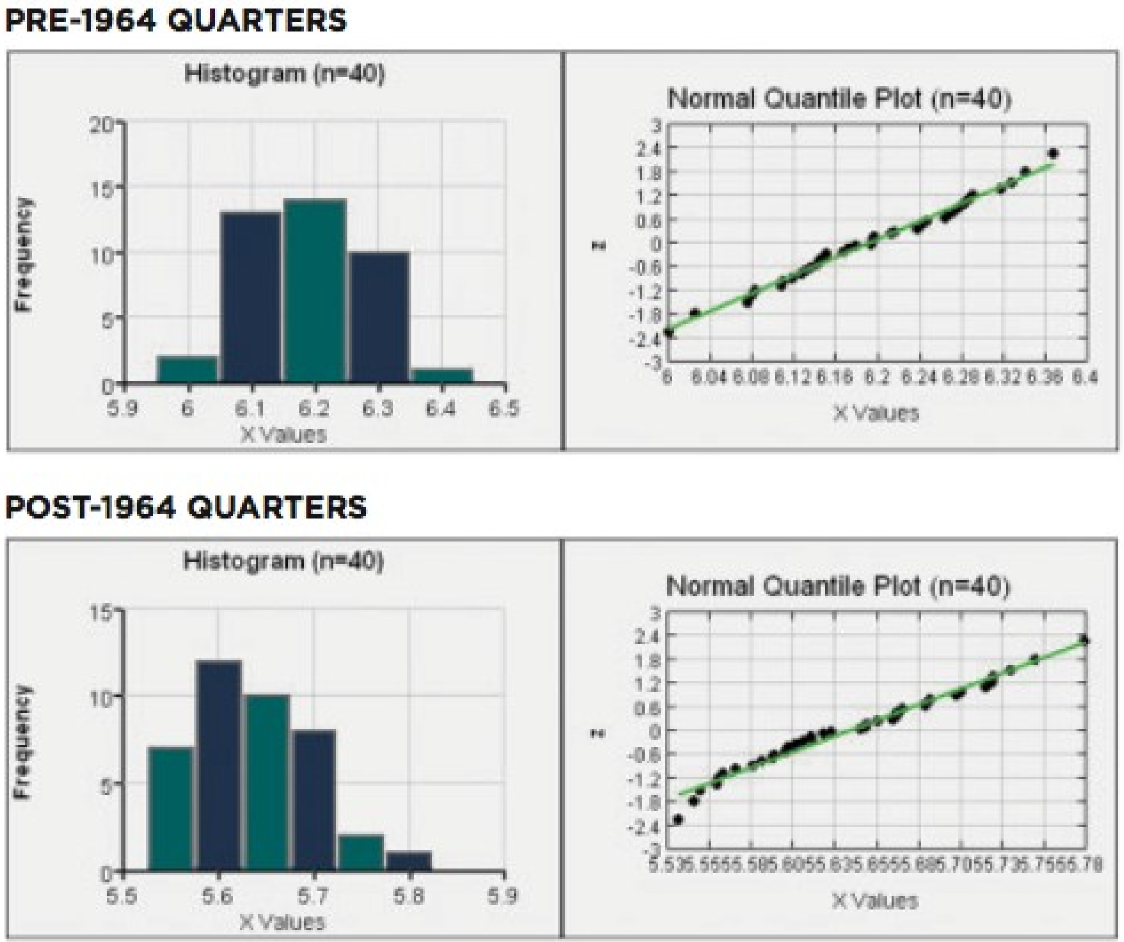
Example:

Data Set 20 in Appendix B includes weights (in g) of quarters made before 1964 and weights of quarters made after 1964. Sample statistics are listed below. When designing coin vending machines, we must consider the standard deviations of pre-1964 quarters and post-1964 quarters. Use a 0.05 significance level to test the claim that the weights of pre-1964 quarters and the weights of post-1964 quarters are from populations with the same standard deviation.

Pre-1964 Quarters	Post-1964 Quarters
$n = 40$	$n = 40$
$s = 0.08700$ g	$s = 0.06194$ g

Example:

Requirements are satisfied: populations are independent; simple random samples; from populations with normal distributions



Example:

Use sample variances to test claim of equal population variances, still state conclusion in terms of standard deviations.

Step 1: claim of equal standard deviations is equivalent to claim of equal variances

$$\sigma_1^2 = \sigma_2^2$$

Step 2: if the original claim is false, then

$$\sigma_1^2 \neq \sigma_2^2$$

Step 3: $H_0:\ \sigma_1^2 = \sigma_2^2$ original claim

$$H_1:\ \sigma_1^2 \neq \sigma_2^2$$

Example:

Step 4: significance level is 0.05

Step 5: involves two population variances, use *F* distribution variances

Step 6: calculate the test statistic

$$F = \frac{s_1^2}{s_2^2} = \frac{0.08700^2}{0.016194^2} = 1.9729$$

For the critical values in this two-tailed test, refer to Table A-5 for the area of 0.025 in the right tail. Because we stipulate that the larger variance is placed in the numerator of the F test statistic, we need to find only the right-tailed critical value.

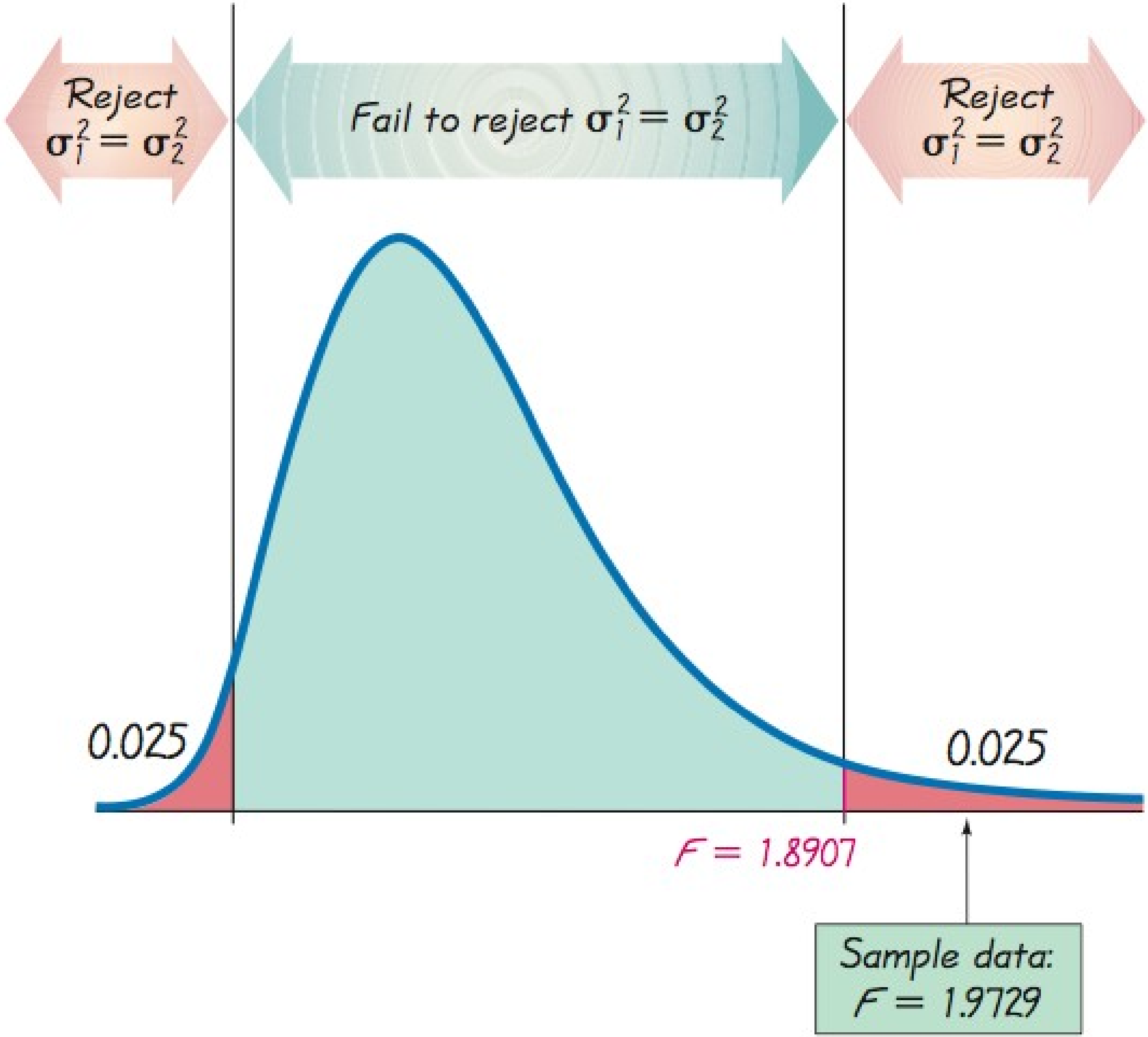
Pearson Education

Example:

From Table A-5 we see that the critical value of F is between 1.8752 and 2.0739, but it is much closer to 1.8752. Interpolation provides a critical value of 1.8951, but STATDISK, Excel, and Minitab provide the accurate critical value of 1.8907.

Step 7: The test statistic F = 1.9729 does fall within the critical region, so we reject the null hypothesis of equal variances. There is sufficient evidence to warrant rejection of the claim of equal standard deviations.

Example:


Reject
$\sigma_1^2 = \sigma_2^2$
Fail to reject $\sigma_1^2 = \sigma_2^2$
Reject
$\sigma_1^2 = \sigma_2^2$
0.025
0.025
$F = 1.8907$
Sample data:
$F = 1.9729$

Example:

There is sufficient evidence to warrant rejection of the claim that the two standard deviations are equal. The variation among weights of quarters made after 1964 is significantly different from the variation among weights of quarters made before 1964.

Recap

In this section we have discussed:

- Requirements for comparing variation in two samples
- Notation.
- Hypothesis test.
- Confidence intervals.
- *F* test and distribution.

Pearson Education