Section 2.1 Describing Qualitative data

1) The data below show the types of medals won by athletes representing the United States in 2006 Winter Olympics

<table>
<thead>
<tr>
<th>Type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td></td>
</tr>
<tr>
<td>Bronze</td>
<td></td>
</tr>
<tr>
<td>Silver</td>
<td></td>
</tr>
</tbody>
</table>

Complete the table

<table>
<thead>
<tr>
<th>Type</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bronze</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silver</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2) A physical therapist wants to get a sense of the types of rehabilitation required by her patients. To do so, she obtains a simple random sample of 30 of her patients and records the body part requiring rehabilitation. See table 1.

Table 1

<table>
<thead>
<tr>
<th>Body Part</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back</td>
<td>Back</td>
<td>hand</td>
<td>Neck</td>
<td>Knee</td>
<td>Knee</td>
<td>Knee</td>
</tr>
<tr>
<td>Wrist</td>
<td>Back</td>
<td>Groin</td>
<td>Shoulder</td>
<td>Shoulder</td>
<td>Back</td>
<td></td>
</tr>
<tr>
<td>Elbow</td>
<td>Back</td>
<td>Back</td>
<td>Back</td>
<td>Back</td>
<td>Back</td>
<td>Back</td>
</tr>
<tr>
<td>Back</td>
<td>Shoulder</td>
<td>Shoulder</td>
<td>Knee</td>
<td>Knee</td>
<td>Back</td>
<td></td>
</tr>
<tr>
<td>Hip</td>
<td>Knee</td>
<td>Hip</td>
<td>Hand</td>
<td>Back</td>
<td>Wrist</td>
<td></td>
</tr>
</tbody>
</table>
a) Construct a frequency distribution and a relative frequency distribution of location of injury.

<table>
<thead>
<tr>
<th>Body Part</th>
<th>Tally</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elbow</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knee</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neck</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Construct a Frequency bar graph of location injury.

c) Construct a relative frequency Pareto chart for the data in Table 1.

d) Construct a pie chart for the data set. The following data give the distribution of the types of houses in a town containing 44,000 houses.

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capes</td>
<td>11,000</td>
</tr>
<tr>
<td>Garrisons</td>
<td>15,400</td>
</tr>
<tr>
<td>Splits</td>
<td>17,600</td>
</tr>
</tbody>
</table>

2) Complete the following table:

<table>
<thead>
<tr>
<th>Grade on Statistics Exam</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: 90-100</td>
<td></td>
<td>.08</td>
</tr>
<tr>
<td>B: 80 – 89</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>C: 65 – 79</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>D: 50 – 64</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>F: Below 50</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>1.00</td>
</tr>
</tbody>
</table>
A qualitative variable with three classes (X, Y, and Z) is measured for each of 20 units randomly sampled from a target population. The data (observed for each unit) are as follows:

<table>
<thead>
<tr>
<th>Y</th>
<th>X</th>
<th>X</th>
<th>Z</th>
<th>X</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>X</th>
<th>X</th>
<th>Z</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td>Z</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Compute the frequency for each of the three classes.
b) Compute the relative frequency for each of the three classes.
c) Display the results from part a in frequency bar graph.
d) Display the results from part b in a pie chart.
Section 2.2 Organizing Quantitative data

Activity 2

Dr Paul randomly selects 40 of his 20- to 29-year-old patients and obtains the following data regarding their serum HDL cholesterol:

<table>
<thead>
<tr>
<th>Class</th>
<th>Tally</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Complete the following table

<table>
<thead>
<tr>
<th>Class</th>
<th>Tally</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-79</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Draw the Frequency Histogram.
Activity 3

Construct the dotplot for the given data.

A manufacturer records the number of errors each work station makes during the week. The data are as follows.

6 3 2 3 5 2 0 2 5 4 2 0 1

Activity 4

Following are the mechanical aptitude scores of 30 applicants for production line jobs in a factory.

<table>
<thead>
<tr>
<th>47</th>
<th>89</th>
<th>50</th>
<th>61</th>
<th>98</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>89</td>
<td>52</td>
<td>47</td>
<td>61</td>
<td>61</td>
<td>50</td>
</tr>
<tr>
<td>44</td>
<td>98</td>
<td>90</td>
<td>98</td>
<td>98</td>
<td>61</td>
</tr>
<tr>
<td>52</td>
<td>52</td>
<td>56</td>
<td>47</td>
<td>47</td>
<td>47</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>50</td>
<td>47</td>
<td>44</td>
<td>44</td>
</tr>
</tbody>
</table>

a) Construct a dot diagram for the data.

b) Construct a stem-and-leaf display with the stem labels 4, 5, 6, 8, 9.

Activity 5

Consider the stem-and-leaf display shown here:

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>457</td>
</tr>
<tr>
<td>3</td>
<td>00036</td>
</tr>
<tr>
<td>2</td>
<td>1134599</td>
</tr>
<tr>
<td>1</td>
<td>2248</td>
</tr>
<tr>
<td>0</td>
<td>0012</td>
</tr>
</tbody>
</table>

a) How many observations were in the original data set?

b) In the bottom row of the stem-and-leaf display, identify the stem, the leaves, and the numbers in the original data set represented by this stem and its leaves.

c) Re-create all the numbers in the data set, and construct a dot plot.
Activity 6

Use the data to create a stem-and-leaf display.

Twenty-four workers were surveyed about how long it takes them to travel to work each day. The data below are given in minutes.

20  35  42  52  65  20  60  49  24  37  23  24
22  20  41  25  28  27  50  47  58  30  32  48
SECTION 2.4 Measures of Central Tendency

1) Calculate the mode, mean, and median of the following data:
   18  10  15  13  17  12  15  18  16  11

2) For the given data set, find the mean, the median, and the mode.
   Ages of teachers in the mathematics department of a certain high school:
   30, 55, 33, 58, 30, 55, 30, 48, 48, 43.

2) The following table gives the total amount charged in 2007 for the top ranked banks.

<table>
<thead>
<tr>
<th>U.S. Bank/ Card Issuer</th>
<th>Amount Charged ($billions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of America</td>
<td>462.20</td>
</tr>
<tr>
<td>JPMorgan Chase</td>
<td>458.84</td>
</tr>
<tr>
<td>American Express</td>
<td>445.32</td>
</tr>
<tr>
<td>Citigroup</td>
<td>240.58</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>192.68</td>
</tr>
</tbody>
</table>

   a) Find the mean amount charged for the top five banks and practically interpret this value.
   b) Find the median amount charged for the top five banks and practically interpret this value.

2) Describe how the mean compares to the median for a distribution as follows:
   a) Skewed to the left
   b) Skewed to the right
   c) Symmetric
Section 2.5: Numerical Measures of Variability

1) Calculate the range, variance, and standard deviation for the following samples:
   a) 4, 2, 1, 0, 1.
   b) 1, 6, 2, 2, 3, 0, 3.

2) The following table gives the total amount charged in 2007 for the top ranked banks.

<table>
<thead>
<tr>
<th>U.S. Bank/ Card Issuer</th>
<th>Amount Charged ($billions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of America</td>
<td>462.20</td>
</tr>
<tr>
<td>JPMorgan Chase</td>
<td>458.84</td>
</tr>
<tr>
<td>American Express</td>
<td>445.32</td>
</tr>
<tr>
<td>Citigroup</td>
<td>240.58</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>192.68</td>
</tr>
</tbody>
</table>

a) Find the range of the data for the five top-ranked banks. Give the units of measurement for the range.
b) Find the variance of the data for five the top-ranked banks. Give the units of measurement for variance.
c) Find the standard deviation of the data for five the top-ranked banks. Give the units of measurement for standard deviation.
Section 2.7 Numerical Measures of Relative Standing

1) Give the percentage of measurements in a data set that are above and below each of the following percentiles:
   a) 75th percentile
   b) 50th percentile
   c) 84th percentile

2) Scores on a test have a mean of 68 and standard deviation of 11. Michelle has a score of 90. Convert Michelle's score to a z-score.

3) Suppose 200 steelworkers are selected, and the annual income of each is determined. The mean and standard deviation are $\bar{x} = $54,000 and $s = $2,000. Suppose Joe Smith's annual income is $52,000. What is his sample z-score?

4) Which is better: a score of 82 on a test with mean of 70 and a standard deviation of 8, or a score of 82 on a test with a mean of 75 and a standard deviation of 4?

5) Three students take equivalent tests of a sense of humor and, after the laughter dies down, their score are calculated. Which is higher relative score?
   a) A score of 144 on a test with a mean of 128 and a standard deviation of 34.
   b) A score of 90 on a test with a mean of 86 and a standard deviation of 18.
   c) A score of 25 on a test with a mean of 15 and a standard deviation of 5.

6) In an English class the final examination grade average 60 with a standard deviation of 16, and in mathematics class the final examination grades average 58 with a standard deviation 10. If a student gets a 72 on the English examination and a 68 on the mathematics examination, how many standard deviations is each of her grades above the average of the respective class? What does this tell us about her performance in the two subjects?
Section 3.1 Events, Sample Space, and Probability

1) An experiment results in one of the following points: $E_1, E_2, E_3, E_4,$ or $E_5$.
   a) Find $P(E_3)$ if $P(E_1) = .1$, $P(E_2) = .2$, $P(E_4) = .1$, and $P(E_5) = .1$.
   b) Find $P(E_3)$ if $P(E_1) = P(E_3), P(E_2) = .1, P(E_4) = .2,$ and $P(E_5) = .1$.
   c) Find $P(E_3)$ if $P(E_1) = P(E_2) = P(E_4) = P(E_5) = .1$.

2) The sample space for an experiment contains five sample points with probabilities as shown in the table. Find the probability of each of the following events:
   
   A : { Either 1, 2, or 3 occurs}   
   B : { Either 1, 3, or 5 occurs}   
   C : { 4 does not occur}   

<table>
<thead>
<tr>
<th>Sample Points</th>
<th>Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.05</td>
</tr>
<tr>
<td>2</td>
<td>.20</td>
</tr>
<tr>
<td>3</td>
<td>.30</td>
</tr>
<tr>
<td>4</td>
<td>.30</td>
</tr>
<tr>
<td>5</td>
<td>.15</td>
</tr>
</tbody>
</table>

Section 3.2 Unions and Intersections

1) Consider the die-toss experiment. Define the following events:
   
   A: { Toss an event number}   
   B: { Toss a number less than or equal to 3}   

   a) Describe $A \cup B$ for this experiment.
   b) Describe $A \cap B$ for this experiment.
   c) Calculate $P(A \cup B)$ and $P(A \cap B)$ assuming the die is fair.
2) In the 2002 General Social Survey, respondents were asked their political leaning and whether they favored the legalization of marijuana. Of the 826 respondents who answered both questions, the results are listed in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Liberal</th>
<th>Moderate</th>
<th>Conservative</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>In favor of Legalization</td>
<td>120</td>
<td>102</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Not in favor Legalization</td>
<td>98</td>
<td>222</td>
<td>206</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Find the probability that a randomly chosen respondent is a liberal and favors legalization.
b) Find the probability that a randomly chosen respondent is moderate and favors legalization.
c) Find the probability that a randomly chosen respondent is conservative and favors legalization.

Section 3.3 Complementary Events

1) Suppose \( P(A) = .4 \), \( P(B) = .7 \), and \( P(A \cap B) = .3 \). Find the following probabilities:
   a) \( P(B^c) \)
   b) \( P(A^c) \)
Section 3.4 The additive Rule and Mutually Exclusive Events

1) Suppose \( P(A) = .4, P(B) = .7 \) and \( P(A \cap B) = .3 \).
   Find the following probabilities
   a) \( P(A \cup B) \)
   b) \( P(A^c) \)
   c) \( P(B^c) \)

2) The probability that Doug will read a book on a Saturday night is .6, and the probability that he will listen to records is .4, and the probability that he will do both is .3. What is the probability that Doug will read a book or listen to records?

3) A fair coin is tossed three times, and the events A and B are defined as follows:
   A: {At least one head is observed.}
   B: {The number of heads observed is odd.}
   a) Identify the sample points in the events A, B, \( A \cup B \), \( A^c \), and \( A \cap B \).
   b) Find \( P(A) \), \( P(B) \), \( P(A \cup B) \), \( P(A^c) \), and \( P(A \cap B) \) by summing the probabilities of the appropriate sample points.
   c) Find \( P(A \cup B) \) using the addition rule. Compare your answer to the one you obtained in part b.
   d) Are the events A and B mutually exclusive? Why?

4) The outcomes of two variables are \((\text{Low}, \text{Medium}, \text{High})\) and \((\text{On}, \text{Off})\), respectively. An experiment is conducted in which the outcomes of each of the two variables are observed. The probabilities associated with each of the six possible outcome pairs are given in the accompanying two-way table.

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>.50</td>
<td>.10</td>
<td>.05</td>
</tr>
<tr>
<td>Off</td>
<td>.25</td>
<td>.07</td>
<td>.03</td>
</tr>
</tbody>
</table>

Consider the following events:
A : {On}
B : {Medium or Low}
C : {Off and Low}
D : {High}

a) Find \( P(A) \).
b) Find \( P(B) \).
c) Find \( P(C) \).
d) Find \( P(D) \).
e) Find \( P(A^c) \).
f) Find \( P(A \cup B) \).
g) Find \( P(A \cap B) \).
h) Consider each pair of events (A and B, A and C, A and D, B and C, C and D).
   List the pairs of events that are mutually exclusive. Justify your choice.
Section 3.5 Conditional Probability

1) For two events, A and B, P(A) = .4, P(B) = .2, and P(A ∩ B) = .1.
   a) Find P(A | B)
   b) Find P(B | A)

2) For two events, A and B, P(A) = .4, P(B) = .2, and P(A | B) = .6:
   a) Find P(A ∩ B).
   b) Find P(B | A)

3) A fast-food restaurant chain with 700 outlets in the United States describes the geographic location of its restaurants with the accompanying table of percentages. A restaurant is to be chosen at random from the 700 to test market a new style of chicken. Given that the restaurant is located in the eastern United States, what is the probability it is located in a city with a population of at least 10,000?

<table>
<thead>
<tr>
<th>Region</th>
<th>Population of City</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NE</td>
</tr>
<tr>
<td>&lt; 10,000</td>
<td>1%</td>
</tr>
<tr>
<td>10,000-100,000</td>
<td>15%</td>
</tr>
<tr>
<td>&gt;100,000</td>
<td>20%</td>
</tr>
</tbody>
</table>

4) After completing an inventory of three warehouses, a golf club shaft manufacturer described its stock of 12,246 shafts with the percentages given in the table. Suppose a shaft is selected at random from the 12,246 currently in stock, and the warehouse number and type of shaft are observed. Given that the shaft is produced in warehouse 2, find the probability it has an extra stiff shaft.

<table>
<thead>
<tr>
<th>Warehouse</th>
<th>Regular</th>
<th>Stiff</th>
<th>Extra Stiff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warehouse 1</td>
<td>19%</td>
<td>8%</td>
<td>18%</td>
</tr>
<tr>
<td>Warehouse 2</td>
<td>14%</td>
<td>5%</td>
<td>8%</td>
</tr>
<tr>
<td>Warehouse 3</td>
<td>10%</td>
<td>18%</td>
<td>0%</td>
</tr>
</tbody>
</table>
5) The breakdown of workers in a particular state according to their political affiliation and type of job held is shown here. Suppose a worker is selected at random within the state and the worker’s political affiliation and type of job are noted. Given the worker is a Democrat, what is the probability that the worker is in a white collar job.

<table>
<thead>
<tr>
<th>Type of job</th>
<th>Political Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Republican</td>
</tr>
<tr>
<td>White collar</td>
<td>15%</td>
</tr>
<tr>
<td>Blue collar</td>
<td>18%</td>
</tr>
</tbody>
</table>

6) The following table, based on data from the Centers for Disease Control, gives the estimated number of people with HIV/AIDS for men and women in 32 states with confidential, name-based reporting of HIV infection.

<table>
<thead>
<tr>
<th>Method</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homosexual Contact (HC)</td>
<td>54,424</td>
<td>0</td>
<td>54,424</td>
</tr>
<tr>
<td>Intravenous Drug Use (IDU)</td>
<td>12,998</td>
<td>6,723</td>
<td>19,721</td>
</tr>
<tr>
<td>HC and IDU</td>
<td>5,022</td>
<td>0</td>
<td>5,022</td>
</tr>
<tr>
<td>Heterosexual Contact</td>
<td>15,413</td>
<td>26,882</td>
<td>42,295</td>
</tr>
<tr>
<td>Other</td>
<td>1,009</td>
<td>987</td>
<td>1,996</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>88,866</strong></td>
<td><strong>34,592</strong></td>
<td><strong>123,458</strong></td>
</tr>
</tbody>
</table>

a) Given that the person living with HIV/AIDS is male, find the probability that he contracted the disease via heterosexual contact.

b) Given that the person living with HIV/AIDS is female, find the probability that she contracted the disease via intravenous drug use.

c) Find the probability that a person living with HIV/AIDS is female.
Section 3.6 The Multiplication Rule and Independent Events

1) For two events A and B, \( P(A) = .4, P(B) = .2, \) and \( P(A \cap B) = .1. \)
   a) Find \( P(A|B) \)
   b) Find \( P(B|A) \)
   c) Are A and B independent?

2) Consider the experiment of tossing a fair die and let
   A: {Observe an even number.}
   B: {Observe a number less than or equal to 4.}
Are events A and B independent?

3) An experiment results in one of five sample points with the following probabilities:
   \( P(E_1) = .22, P(E_2) = .31, P(E_3) = .15, P(E_4) = .22, P(E_5) = .1. \)

The following events have been defined:
   A: \( \{E_1, E_2\} \)
   B: \( \{E_2, E_3, E_4\} \)
   C: \( \{E_1, E_5\} \)

Find each of the following probabilities:
   d) \( P(A) \)
   e) \( P(B) \)
   f) \( P(A \cap B) \)
   g) \( P(A|B) \)
   h) \( P(B \cap C) \)
   i) \( P(C|B) \)
   j) Consider each pair of events: A and B, A and C, and B and C. Are any of the pairs of events independent? Why?
Part 2

Section 4.1 Two Types of Random Variables

1) Which of the following describe continuous random variables? Which describe discrete random variables?
   a) The number of newspaper sold by the New York Times each month
   b) The amount of ink used in printing a Sunday edition of the New York Times
   c) The actual number of ounces in a one-gallon bottle of laundry detergent
   d) The number of detective parts in a shipment of nuts and bolts
   e) The number of people collecting unemployment insurance each month

Section 4.2 Probability Distributions for Discrete Random Variables

1) A die is tossed. Let x be the number of spots observed on the upturned of the die.
   a) Find the probability distribution of x and display it in tabular form.
   b) Display the probability distribution of x in graphical form.

2) The random variable x has the following discrete probability distribution:

<table>
<thead>
<tr>
<th>x</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(x)</td>
<td>.1</td>
<td>.2</td>
<td>.4</td>
<td>.2</td>
<td>.1</td>
</tr>
</tbody>
</table>

   a) List the values x may assume.
   b) What value of x is most probable?
   c) Display the probability distribution as a graph.
   d) Find P(x = 7).
   e) Find P(x ≥ 5).
   f) Find P(x > 2).
3) Suppose you invest a fixed sum of money in each of five Internet Business ventures. Assume you know that 70% of such ventures are successful, the outcomes of the ventures are independent of one another, and the probability distribution for the number, $x$, of successful ventures out of five is

<table>
<thead>
<tr>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(x)</td>
<td>.002</td>
<td>.029</td>
<td>.132</td>
<td>.309</td>
<td>.360</td>
<td>.168</td>
</tr>
</tbody>
</table>

a) Find $\mu = E(x)$.

b) Find $\sigma = \sqrt{E[(x - \mu)^2]}$.

c) Graph $p(x)$. Locate $\mu$ and the interval $\mu \pm 2 \sigma$ on the graph. Use either Chebyshev’s Rule or the Empirical Rule to approximate the probability that $x$ falls in this interval. Compare this result with the actual probability.

4) Consider the probability distribution shown here:

<table>
<thead>
<tr>
<th>$x$</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(x)</td>
<td>.02</td>
<td>.07</td>
<td>.10</td>
<td>.15</td>
<td>.30</td>
<td>.18</td>
<td>.10</td>
<td>.06</td>
<td>.02</td>
</tr>
</tbody>
</table>

a) Calculate $\mu, \sigma^2,$ and $\sigma$.

b) Graph $p(x)$. Locate $\mu, \mu - 2 \sigma,$ and $\mu + 2 \sigma$ on the graph.

c) What is the probability that $x$ is in the interval $\mu \pm 2 \sigma$
Section 4.3 The Binomial Distribution

1) If \( x \) is a binomial random variable, use the binomial table to find the following probabilities:
   a) \( P(x = 2) \) for \( n = 10 \), \( p = .4 \)
   b) \( P(x \leq 5) \) for \( n = 15 \), \( p = .6 \)
   c) \( P(x < 10) \) for \( n = 20 \), \( p = .7 \)

2) If \( x \) is a binomial random variable, use the binomial formula to find the following probabilities:
   a) \( P(x=0) \), \( n = 3 \) and \( p = .3 \)
   b) \( P(x=1) \), \( n = 3 \) and \( p = .3 \)
   c) \( P(x=1) \), \( n = 5 \) and \( p = .1 \)

3) A machine that produces stampings for automobile engines is malfunctioning and producing 10% defectives. The defective and nondefective stampings proceed from the machine in a random manner. If the next five stampings are tested, find the probability that three of them are defective.

4) A study has shown that 50% of the families in a certain large area have at least two cars. Find the probabilities that among 15 families randomly selected in this area
   a) 7 have at least two cars.
   b) More than 8 have at least two cars

5) If \( x \) is a binomial random variable, calculate \( \mu, \sigma^2 \), and \( \sigma \) for each of the following:
   a) \( n = 25 \), \( p = .5 \)
   b) \( n = 80 \), \( p = .2 \)
   c) \( n = 100 \), \( p = .6 \)
   d) \( n = 70 \), \( p = .9 \)
   e) \( n = 1000 \), \( p = .04 \)

6) Suppose a poll of 20 employees is taken in a large company. The purpose is to determine \( x \), the number who favor unionization. Suppose 60% of all the company’s employees favor unionization.
   a) Find the mean and the standard deviation of \( x \).
   b) Use the Table to find the probability that \( x > 12 \).
   c) Use the Binomial Table to find the probability that \( x = 11 \).
   d) Graph the probability distribution of \( x \) and locate the interval \( \mu \pm 2 \sigma \) on the graph.
Sections 5.1 & 5.2 The Normal Distribution

1) Find the following probabilities for the normal random variable:
   a) \( P(z > 1.46) \)  
   b) \( P(z < -1.56) \)  
   c) \( P(0.67 < z < 2.41) \)  
   d) \( P(-1.96 < z < -0.33) \)  
   e) \( P(z \geq 0) \)  
   f) \( P(-2.33 < z \leq 1.50) \).

2) Find the probability that the standard normal random variable \( z \) falls between \(-1.33\) and 1.33.

3) Find the probability that a standard normal random variable exceeds 1.64; that is find \( P(z > 1.64) \).

4) Find the area under the standard normal probability distribution between the following pairs of \( z \)-scores:
   a) \( z = 0 \) and \( z = 2 \)  
   b) \( z = 0 \) and \( z = 1.5 \)

5) Find the probability that a standard normal random variable lie to the left of 0.67.

6) Find the probability that a standard normal random variable exceeds 1.96 in absolute value; that is \( P(|z| > 1.96) = P(z < -1.96 \text{ or } z > -1.96) \).

7) Find a value of the standard normal random variable \( z \), call it \( z_0 \), such that
   a) \( P(z \geq z_0) = 0.05 \)  
   b) \( P(z \geq z_0) = 0.025 \)  
   c) \( P(z \leq z_0) = 0.025 \)  
   d) \( P(z \leq z_0) = 0.2090 \)  
   e) \( P(-z_0 \leq z < z_0) = 0.8472 \)  
   f) \( P(-z_0 \leq z < z_0) = 0.1664 \).

8) For a standard normal curve, find the z-score that separates the bottom 90% from the top 10%.

9) For a standard normal curve, find the z-score that separates the bottom 30% from the top 70%.

10) Find the z-score that is less than the mean and for which 70% of the distribution’s area lies to its right.

11) Assume that the length of time \( x \), between charges of a cellular phone is normally distributed with a mean of 10 hours and a standard deviation of 1.5 hours. Find the probability that the cell phone will last between 8 and 12 hours between charges.

12) Suppose the height of college students is normally distributed with a mean of 70.2 inches and a standard deviation of 5.3 inches. What percent of the students have a height of less than 59.6 inches?
13) The mean weight of newborn babies is 7 pounds and the standard deviation is 0.8 pounds. The characteristic of weight of newborn infants is normally distributed.
   a) What percent of newborn babies weight more than 8 pounds?
   b) What percent of newborn babies weight between 6 and 8 pounds?

14) A bank’s loan officer rates applicants for credit. The ratings are normally distributed with a mean of 200 and a standard deviation of 50. If an applicant is randomly selected, find the probability of a rating that is between 170 and 220.

15) Assume that the weights of quarters are normally distributed with a mean of 5.67 g and a standard deviation 0.070 g. A vending machine will only accept coins weighing between 5.48g and 5.82 g. What percentage of legal quarters will be rejected?

Section 6.3 The Sampling Distribution of a Sample Mean and the Central Limit Theorem

1) Suppose a random sample of n = 25 measurements is selected from a population with mean \( \mu \) and standard deviation \( \sigma \). For each of the following values of \( \mu \) and \( \sigma \), give the values of \( \mu_\bar{x} \) and \( \sigma_\bar{x} \)
   a) \( \mu = 10, \sigma = 3 \)
   b) \( \mu = 20, \sigma = 40 \)
   c) \( \mu = 100, \sigma = 25 \)
   d) \( \mu = 10, \sigma = 100 \)
Sections 7.1 & 7.2 Confidence Interval for a Population Mean: Normal (z) Statistic (Large Samples)

1) If $z_\alpha$ denotes the value of $z$ for which the area under the standard normal curve to its right is equal to $\alpha$ (Greek lowercase alpha), find
   1) $z_{0.01}$
   2) $z_{0.05}$
   3) $z_{0.025}$

2) Find $z_{\alpha/2}$ for each of the following:
   1) $\alpha = .10$
   2) $\alpha = .20$

1) A random sample of 100 observations form a normally distributed population possesses a mean equal to 83.2 and a standard deviation equal to 6.4.
   a) Find a 90% confidence interval for $\mu$. Write a statement that interprets the confidence interval.
   b) Find a 95% confidence interval for $\mu$. Write a statement that interprets the confidence interval.
   c) Find a 99% confidence interval for $\mu$.
   d) What happens to the width of a confidence interval as the value of the confidence coefficient is increased while the sample size is held fixed?

2) The mean and standard deviation of a random sample of $n$ measurements are equal to 33.9 and 3.3, respectively.
   a) Find a 95% confidence interval for $\mu$ if $n = 100$.
   b) Find a 95% confidence interval for $\mu$ if $n = 400$.
   c) Find the widths of the confidence intervals you calculated in parts a and b. What is the effect on the width of a confidence interval of quadrupling the sample size while holding the confidence coefficient fixed?

3) A random sample of 90 observations produced a mean $\bar{x} = 25.9$ and a standard deviation $s = 2.7$.
   A) Find an approximate 90% confidence interval for the population mean $\mu$.
   B) Find an approximate 95% confidence interval for the population mean $\mu$.
   C) Find an approximate 99% confidence interval for the population mean $\mu$. 
4) A random sample of 105 light bulbs had a mean life of $\bar{x} = 441$ hours with a standard deviation of $\sigma = 40$ hours. Construct a 90% confidence interval for the mean life, $\mu$, of all light bulbs of this type.

5) A laboratory tested 83 chicken eggs and found that the mean amount of cholesterol was 233 milligrams with $\sigma = 12.9$ milligrams. Construct a 95% confidence interval for the true mean cholesterol content, $\mu$, of all such eggs.

Section 7.3 Confidence Interval for a Population Mean Student’s $t$-Statistic (Small-Sample)

1) In six attempts it took a locksmith 9, 14, 7, 8, 11, and 5 seconds to open a certain kind of lock. Verify that $\bar{x} = 9$ and $s = 3.16$ for these data, and construct a 95% confidence interval for the average time it takes the locksmith to open this kind of lock.

2) To test the durability of a new paint for white center lines, a highway department painted test strips across heavily traveled roads eight different locations, and electronic counter showed that deteriorated after having been crossed by 14.26, 16.78, 13.65, 10.83, 12.64, 13.37, 16.20, and 14.94 million cars. Construct a 95% confidence interval for the average amount of traffic (car crossing) this paint can withstand before it deteriorates. (Verify that $\bar{x} = 14.08$ and $s = 1.92$)

3) The following random sample was selected from a normal distribution: 4, 6, 3, 4, 9, 3. Verify that $\bar{x} = 4.83$ and $s = 2.32$
   a) Construct a 90% confidence interval for the population mean $\mu$.
   b) Construct a 95% confidence interval for the population mean $\mu$.
   c) Construct a 99% confidence interval for the population mean $\mu$.

4) A laboratory tested twelve chicken eggs and found that the mean amount of cholesterol was 225 milligrams with $s = 15.7$ milligrams. Construct a 95% confidence interval for the true mean cholesterol content of all such eggs.

5) A savings and loan association needs information concerning the checking account balances of its local customers. A random sample of 14 accounts was checked and
yielded a mean balance of $664.14 and a standard deviation of $297.29. Find a 98% confidence interval for the true mean checking account balance for local customers.

Section 7.4: Confidence Interval for a Population Proportion (Large-Sample)

1) A random sample of size n= 121 yielded \( \hat{p} = 0.78 \).
   a) Is the sample size large enough to use the methods of this chapter to construct a confidence interval for \( p \)?
   b) Construct a 90% confidence interval for \( p \).
   c) Construct a 99% confidence interval for \( p \).

2) A random sample of size n =225 yielded \( \hat{p} = 0.46 \).
   a) Construct a 95% confidence interval for \( p \).
   b) Interpret the 95% confidence interval.

3) In a Gallup poll, 491 randomly selected adults were asked whether they are in favor of the death penalty for a person convicted of murder, and 65% of them said that they were in favor.
   a) Find the point estimate of the percentage of adults who are in favor of this death penalty.
   b) Find a 95% confidence interval estimate of the percentage of adults who are in favor of this penalty.

4) An inspector examined a random sample of 400 bottles of shampoo and found that 152 of the bottles had loose bottle-caps that required tightening. Construct a 95% confidence interval for the true proportion of bottle-caps that needed tightening, in the population sample.
5) For over 20 years, the Occupational Safety & Health Administration has required companies that handle hazardous chemicals to complete material safety data sheets (MSDSs). These MSDSs have been criticized for being too hard to understand and complete by workers. Although improvements were implemented in 1990, a more recent study of 150 MSDSs revealed that only 11% were satisfactorily completed (Chemical & Engineering News, Feb. 7, 2005).

   a) Give a point estimate of \( p \), the true proportion of MSDSs that are satisfactorily completed.
   b) Find a 95% confidence interval for \( p \).
   c) Give a practical interpretation of the interval, part b.

Section 7.5 Determining the Sample Size

1) If you wish to estimate a population mean with a sampling error of \( \text{SE} = .3 \) using a 95% confidence interval, and you know from prior sampling that \( \sigma^2 \) is approximately equal to 7.2, how many observations would have to be included in your sample?

2) How many women must be randomly selected to estimate the mean weight of women in age group? We want 90% confidence that the sample mean is within 3.7 lb (\( \text{SE} = 3.7\text{lb} \)) of the population mean, and the population standard deviation is known to be 28lb.

3) How many students must be randomly selected to estimate the mean weekly earnings of students at one college? We want 95% confidence that the sample mean is within $5 of the population mean, and the population standard deviation is known to be $63.

Sections 8.1 The Elements of a Test of Hypothesis
1) Example problem: Test that the population mean is not 3.

Steps:
- State the question statistically \( \mu \neq 3 \)
- State the opposite statistically \( \mu = 3 \)
  — Must be mutually exclusive & exhaustive
- Select the alternative hypothesis \( \mu \neq 3 \)
  — Has the \( \neq, <, \text{ or } > \) sign
- State the null hypothesis \( \mu = 3 \)

Alternative hypothesis: \( H_a : \mu \neq 3 \)
Null hypothesis: \( H_0 : \mu = 3 \)

What are the Hypotheses?

2) Is the population average amount of TV viewing 12 hours?
3) Is the population average amount of TV viewing different from 12 hours?
4) Is the average cost per hat less than or equal to $20?
5) Is the average amount spent in the bookstore greater than $25?

Section 8.2 Formulating Hypothesis and Setting Up the Rejection Region

1) Which hypothesis, the null or the alternative, is the status quo hypothesis? Which is the research hypothesis?
2) Consider a test of \( H_0 : \mu = 4 \). In each of the following cases, give the rejection region for the test in terms of the \( z \)-statistic:
   a) \( H_a : \mu > 4 \), \( \alpha = .05 \)
   b) \( H_a : \mu > 4 \), \( \alpha = .10 \)
   c) \( H_a : \mu < 4 \), \( \alpha = .05 \)
   d) \( H_a : \mu \neq 4 \), \( \alpha = .05 \)

Section 8.3 Test of Hypothesis about a Population Mean: Normal (z) Statistic
1) A random sample of 100 observations from a population with standard deviation 60 yielded a sample mean of 110.
   a) Test the null hypothesis that \( \mu = 100 \) against the alternative hypothesis that \( \mu > 100 \) using \( \alpha = .05 \)
   b) Test the null hypothesis that \( \mu = 100 \) against the alternative hypothesis that \( \mu \neq 100 \) using \( \alpha = .05 \)

2) According to CTIA- The wireless Association, the mean monthly cell phone bill in 2004 was $50.64. A market researcher believes that the mean monthly cell phone bill is different today, but is not sure whether bills have declined because of technological advances or increased due to additional use. The researcher phones a simple random sample of 12 cell phone subscribers and finds that their mean monthly bill is $65.014. Assuming \( \sigma = 18.49 \), determine whether the mean monthly cell phone bills is different from $50.64 at the \( \alpha = 0.05 \) level of significance.

3) A taxicab firm suspects that the average lifetime of 25,000 miles claimed for certain tires is too high. To test the claim, the firm puts a random sample of 40 of these tires on its taxicabs and later finds that their mean lifetime is 24,421 miles and the standard deviation is 1,349 miles. What can the firm conclude at the 0.01 level of significance?

4) According to the U.S. Federal Highway Administration, the mean number of miles driven annually is 12,200. Patricia believes that residents of the state of Montana drive more than the national average. She obtains a simple random sample of 35 drivers from a list of registered drivers in the state of Montana. The mean number of miles driven for the 35 drivers is 12,895.9. Assuming \( \sigma = 3800 \) miles, does the sample provide sufficient evidence that residents of the state of Montana drive more than the national average at the level \( \alpha = 0.1 \) level of significance?

5) A random sample of 64 observations produced the following summary statistics: \( \bar{x} = .323 \) and \( s^2 = .034 \).
   a) Test the null hypothesis that \( \mu = .36 \) against the alternative that \( \mu < .36 \) using \( \alpha = .10 \).
   b) Test the null hypothesis that \( \mu = .36 \) against the alternative that \( \mu \neq .36 \) using \( \alpha = .10 \).

Section 8.4 Observed Significance Levels- p-values
1) If a test were conducted using $\alpha = 0.05$, for which of the following $p$-values would the null hypothesis be rejected?
   a) 0.06  
   b) 0.10  
   c) 0.01  
   d) 0.001 
   e) 0.251 
   f) 0.042

2) For each $\alpha$ and observed significance level (p-value) pair, indicate whether the null hypothesis would be rejected.
   a) $\alpha = 0.05$, p-value = 0.10  
   b) $\alpha = 0.10$, p-value = 0.05  
   c) $\alpha = 0.025$, p-value = 0.05

3) In test of the hypothesis $H_0 : \mu = 50$ versus $H_a : \mu > 50$, a sample of $n = 100$ observations possessed mean $\bar{x} = 49.4$ and standard deviation $s = 4.1$. Find and interpret the $p$-value for this test.

4) In a test of the hypothesis $H_0 : \mu = 100$ against $H_a : \mu > 100$, the sample data yielded the test statistic $z = 2.17$. Find and interpret the $p$-value for the test.

5) An analyst tested the null hypothesis $\mu \geq 20$ against the alternative hypothesis that $\mu < 20$. The analyst reported a $p$-value of 0.06. What is the smallest value of $\alpha$ for which the null hypothesis would be rejected.

---

Section 8.5 Test of Hypothesis about a Population Mean: Student’s $t$-Statistic

1) A random sample of $n$ observations is selected from a normal population to test the null hypothesis that $\mu = 10$.
   Specify the rejection region for each of the following combinations of $H_a$, $\alpha$, $n$:
   a) $H_a : \mu \neq 10$; $\alpha = 0.05$; $n = 14$
   b) $H_a : \mu > 10$; $\alpha = 0.01$; $n = 24$
   c) $H_a : \mu > 10$; $\alpha = 0.10$; $n = 9$
   d) $H_a : \mu < 10$; $\alpha = 0.01$; $n = 12$
2) A sample of five measurements, randomly selected from a normally distributed population, resulted in the following summary statistics: $\bar{x} = 4.8$, $s = 1.3$.

   a) Test the null hypothesis that the mean of the population is $6$ against the alternative, $\mu < 6$. Use $\alpha = 0.05$.

   b) Test the null hypothesis that the mean of the population is $6$ against the alternative hypothesis, $\mu \neq 6$. Use $\alpha = 0.05$.

3) When bonding teeth, orthodontists must maintain a dry field. A new bonding adhesive (called Smartbond) has been developed to eliminate the necessity of a dry field. However, there is concern that the new bonding adhesive is not as strong enough as the current standard, a composite adhesive. Tests on a sample of $10$ extracted teeth bonded with the new adhesive resulted in a mean breaking strength (after 24 hours) of $\bar{x} = 5.07$ Mpa and a standard deviation of $s = 0.46$ Mpa. Orthodontists want to know if the true mean breaking strength of the new bonding adhesive is less than $5.70$ Mpa, the mean breaking strength of the composite adhesive.

   a) Set up the null and alternative hypotheses for the test.
   b) Find the rejection region for the test using $\alpha = 0.01$.
   c) Compute the test statistic.
   d) Give the appropriate conclusion for the test.
Section 8.6 Large-Sample Test of Hypothesis about a Population Proportion

1) Suppose a random sample of 100 observations from a binomial population gives a value of \( \hat{p} = .63 \) and you wish to test the null hypothesis that the population parameter \( p \) is equal to .70 against the null hypothesis that the population is less than .70.
   a) Set up the null and alternative hypotheses for the test.
   b) Find the rejection region for the test using \( \alpha = 0.05 \).
   c) Compute the test statistic.
   d) Give the appropriate conclusion for the test.

2) Suppose a random sample of 100 observations from a binomial population gives a value of \( \hat{p} = .83 \) and you wish to test the null hypothesis that the population parameter \( p \) is equal to .9 against the null hypothesis that the population is less than .9.
   a) Set up the null and alternative hypotheses for the test.
   b) Find the rejection region for the test using \( \alpha = 0.05 \).
   c) Compute the test statistic.
   d) Give the appropriate conclusion for the test.

3) According to Hewitt Association, 36% of the companies surveyed in 1999 paid holiday bonuses to their employees (The Wall Street Journal, December 22, 1999). Assume that this result holds true for all U.S. companies in 1999. A recent random sample of 400 companies showed that 33% of them pay holiday bonuses to their employees. Using the 1% significance level, can you conclude that the current percentage of companies that pay holiday bonuses to their employees is different from that for 1999?
   a) Set up the null and alternative hypotheses for the test.
   b) Find the rejection region for the test using \( \alpha = 0.05 \).
   c) Compute the test statistic.
   d) Give the appropriate conclusion for the test.

4) Managers often tolerate poor performance by employees because the termination process is so complicated. In a survey of federal government managers who had not taken action against unsatisfactory employees, 66% stated that they failed to act because of the long process required (USA TODAY, July 29, 1998). Suppose that 80 such managers are randomly selected and 48 of them failed to act because of the long process required. Can you conclude that the current percentage of such employees who fail to act because of the long process involved is less than 66%. Use \( \alpha = .025 \)