PRINTED TEST BANK

JUSTINE C. BAKER
Peirce College, Philadelphia, PA

to accompany

THE TRIOLA STATISTICS SERIES:

*Elementary Statistics*, Tenth Edition

Mario F. Triola
Dutchess Community College

PEARSON
Addison Wesley

Boston San Francisco New York
London Toronto Sydney Tokyo Singapore Madrid
Mexico City Munich Paris Cape Town Hong Kong Montreal
CONTENTS

Each test is immediately followed by its Answer Key.

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>16</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>49</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>67</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>82</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>100</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>118</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>133</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>152</td>
</tr>
<tr>
<td>Chapter 10</td>
<td>173</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>204</td>
</tr>
<tr>
<td>Chapter 12</td>
<td>222</td>
</tr>
<tr>
<td>Chapter 13</td>
<td>252</td>
</tr>
<tr>
<td>Chapter 14</td>
<td>275</td>
</tr>
</tbody>
</table>
CHAPTER 1 FORM A

Name: ______________________________ Course Number: ________ Section Number: ______

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Define random sample. Explain why this is important in design of experiments.

Determine whether the given value is a statistic or a parameter.

2) A sample of 120 employees of a company is selected, and the average age is found to be 37 years.
   A) Statistic  B) Parameter

Identify the number as either continuous or discrete.

3) The number of freshmen entering college in a certain year is 621.
   A) Continuous  B) Discrete

Determine which of the four levels of measurement (nominal, ordinal, interval, ratio) is most appropriate.

4) Survey responses of "good, better, best".
   A) Nominal  B) Interval  C) Ratio  D) Ordinal

5) Salaries of college professors.
   A) Interval  B) Ordinal  C) Ratio  D) Nominal

Identify the sample and population. Also, determine whether the sample is likely to be representative of the population.

6) An employee at the local ice cream parlor asks three customers if they like chocolate ice cream.
CHAPTER 1 FORM A

Use critical thinking to develop an alternative conclusion.

7) In a study of headache patients, every one of the study subjects with a headache was found to be improved after taking a week off of work. Conclusion: Taking time off work cures headaches.

Use critical thinking to address the key issue.

8) An airline company advertises that 100% of their flights are on time after checking 5 randomly selected flights and finding that these 5 were on time.

9) “38% of adults in the United States regularly visit a doctor”. This conclusion was reached by a college student after she had questioned 520 randomly selected members of her college. What is wrong with her survey?

Perform the requested conversions. Round decimals to the nearest thousandth and percents to the nearest tenth of a percent, if necessary.

10) Convert 0.4 to an equivalent fraction and percentage.
    A) $\frac{3}{10}$, 40%  
    B) $\frac{2}{5}$, 40%  
    C) $\frac{3}{10}$, 4%  
    D) $\frac{2}{5}$, 4%

11) Convert 90% to an equivalent fraction and decimal.
    A) $\frac{4}{5}$, 9  
    B) $\frac{9}{10}$, 9  
    C) $\frac{9}{10}$, 0.9  
    D) $\frac{4}{5}$, 0.9
CHAPTER 1 FORM A

Solve the problem.

12) On a test, if 125 questions are answered and 68% of them are correct, what is the number of correct answers?

   A) 54    B) 62    C) 85    D) 90

Is the description an observational study or an experiment?

13) A stock analyst compares the relationship between stock prices and earnings per share to help him select a stock for investment.

   A) Observational study    B) Experiment

14) A T.V. show’s executives raised the fee for commercials following a report that the show received a “No. 1” rating in a survey of viewers.

   A) Experiment    B) Observational study

Identify the type of observational study.

15) A town obtains current employment data by polling 10,000 of its citizens this month.

   A) Retrospective    B) Cross-sectional
   C) Prospective    D) None of these

Identify which of these types of sampling is used: random, stratified, systematic, cluster, convenience.

16) 49, 34, and 48 students are selected from the Sophomore, Junior, and Senior classes with 496, 348, and 481 students respectively.

   A) Convenience    B) Cluster
   C) Random    D) Systematic
   E) Stratified

17) A sample consists of every 49th student from a group of 496 students.

   A) Convenience    B) Random
   C) Cluster    D) Stratified
   E) Systematic
CHAPTER 1 FORM A

18) The name of each contestant is written on a separate card, the cards are placed in a bag, and three names are picked from the bag.
   A) Convenience
   B) Stratified
   C) Cluster
   D) Random
   E) Systematic

Provide an appropriate response.

19) Explain what is meant by the term “confounding” and give an example of an experiment in which confounding is likely to be a problem.

20) A researcher wants to obtain a sample of 100 school teachers from the 800 school teachers in a school district. Describe procedures for obtaining a sample of each type: random, systematic, convenience, stratified, cluster.
Answer Key
Testname: CHAPTER 1 FORM A

1) In random sampling, each member of the population has an equal chance of being selected. Random sampling provides us with the best representative sample in which all groups of the population are approximately proportionately represented. Careless sampling can easily result in a biased sample which may be useless.

2) A
3) B
4) D
5) C

6) Sample: the 3 selected customers; population: all customers; not representative
7) Headaches generally last for only a few hours, so anything would seem like a cure. There is no evidence to suggest that taking time off work will cure a headache.

8) The sample was too small.
9) The sample is biased. College students are not representative of the U.S. population as a whole.
10) B
11) C
12) C
13) A
14) B
15) B
16) E
17) E
18) D
19) Confounding occurs in an experiment when the effects of two or more variables cannot be distinguished from each other. Examples will vary.

One example is that of a school district that conducts a study regarding whether the science laboratory approach or the computer simulation approach is better for learning chemistry among seniors. One school is randomly selected to conduct only science labs; the other, only computer simulations. A standardized achievement test is used to measure learning, and the results of the two schools are compared. Unless controlled in the study, two confounding variables are teaching expertise and student motivation.

20) Answers will vary.

One answer is as follows. (1) Random: List the names of the teachers in alphabetical order from 1 through 800. Select 100 teachers by a random number computer program.
(2) Systematic: Blindly select from a box one of eight index cards, each of which has a number from 1 to 8 written on it. Sample from the alphabetized list, beginning with that number followed by all its integral multiples until 100 teachers are selected.
(3) Convenience: Offer an incentive to the teachers, and select the first 100 volunteers. (4) Stratified: Prepare an alphabetized list of teachers by school (i.e., strata) and randomly select teachers in proportion to school size until 100 teachers are selected.
(5) Cluster: Form 8 clusters from 8 consecutive blocks of 100 teachers in the alphabetized list. Blindly draw an index card from the box, and whichever card is drawn, all 100 teachers in that cluster will be the sample. Making clusters from the individual schools might not work, since the school or schools randomly selected might not have 100 teachers in total.
CHAPTER 1 FORM B

Name:____________________________ Course Number:_______ Section Number:_______

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Define sampling error and nonsampling error. Give examples of nonsampling error.

---

Determine whether the given value is a statistic or a parameter.

2) After inspecting all of 55,000 kg of meat stored at the Wurst Sausage Company, it was found that 45,000 kg of the meat was spoiled.
   A) Statistic       B) Parameter

Identify the number as either continuous or discrete.

3) The number of stories in a Manhattan building is 22.
   A) Continuous       B) Discrete

Determine which of the four levels of measurement (nominal, ordinal, interval, ratio) is most appropriate.

4) Student’s grades, A, B, or C, on a test.
   A) Ordinal       B) Ratio       C) Nominal       D) Interval

5) Ages of survey respondents.
   A) Ordinal       B) Ratio       C) Interval       D) Nominal

Identify the sample and population. Also, determine whether the sample is likely to be representative of the population.

6) In a poll of 50,000 randomly selected college students, 74% answered "yes" when asked "Do you have a television in your dorm room?".
CHAPTER 1 FORM B

Use critical thinking to develop an alternative conclusion.

7) A study shows that adults who work at their desk all day weigh more than those who do not. Conclusion: Desk jobs cause people to gain weight.

Use critical thinking to address the key issue.

8) You plan to make a survey of 200 people. The plan is to talk to every 10th person coming out of the school library. Is there a problem with your plan?

9) A questionnaire is sent to 10,000 persons. 5,000 responded to the questionnaire. 3,000 of the respondents say that they "love chocolate ice cream". We conclude that 60% of people love chocolate ice cream. What is wrong with this survey?

Perform the requested conversions. Round decimals to the nearest thousandth and percents to the nearest tenth of a percent, if necessary.

10) Convert 90% to an equivalent fraction and decimal.
   A) $\frac{4}{9}$  9  
   B) $\frac{4}{5}$  0.9  
   C) $\frac{9}{10}$  0.9  
   D) $\frac{9}{10}$  9

11) Convert $\frac{17}{150}$ to an equivalent decimal and percent.
   A) 0.113, 11.3%  
   B) 0.113, 1.13%  
   C) 0.233, 233%  
   D) 0.233, 23.3%
CHAPTER 1 FORM B

Solve the problem.

12) On a test, if 80 questions are answered and 36% of them are correct, what is the number of correct answers?

A) 50          B) 45          C) 32          D) 29

Is the description an observational study or an experiment?

13) A quality control specialist compares the output from a machine with a new lubricant to the output of machines with the old lubricant.

A) Experiment          B) Observational study

14) A stock analyst selects a stock from a group of twenty for investment by choosing the stock with the greatest earnings per share reported for the last quarter.

A) Experiment          B) Observational study

Identify the type of observational study.

15) A statistical analyst obtains data about ankle injuries by examining a hospital's records from the past 3 years.

A) Cross-sectional          B) Retrospective
C) Prospective          D) None of these

Identify which of these types of sampling is used: random, stratified, systematic, cluster, convenience.

16) A market researcher selects 500 drivers under 30 years of age and 500 drivers over 30 years of age.

A) Random
B) Convenience
C) Systematic
D) Cluster
E) Stratified

17) A market researcher selects 500 people from each of 10 cities.

A) Systematic
B) Stratified
C) Convenience
D) Cluster
E) Random
CHAPTER 1 FORM B

18) An education researcher randomly selects 48 middle schools and interviews all the teachers at each school.
    A) Systematic
    B) Convenience
    C) Cluster
    D) Stratified
    E) Random

Provide an appropriate response.

19) A researcher obtains a sample of high school teachers in his school district by randomly selecting 10 high schools and interviewing all the teachers at each of these 10 schools. What kind of sampling is being used here? Will the resulting sample be a simple random sample of the population of teachers in the school district? Explain your thinking.

20) Why do you think that cluster sampling is frequently used in practice?
Answer Key
Testname: CHAPTER 1 FORM B

1) Sampling error is the difference between a sample result and the true population result. Such an error results from chance sample fluctuations. A nonsampling error occurs when the sample data are incorrectly collected, recorded, or analyzed. Examples include nonrandom samples, defective measuring instruments, biased survey questions, a large number of refusals, copying sample data incorrectly.

2) B
3) B
4) A
5) B
6) Sample: the 50,000 selected college students; population: all college students; representative
7) Desk job workers are confined to their chairs for most of their work day. Other jobs require standing or walking around which burns calories. It is probably the lack of exercise that causes higher weights, not the desk job itself. Avoid causality altogether by saying lack of walking and exercise is associated with higher weights.
8) People who don’t go to the library are excluded.
9) This is not a random sample. The survey is based on voluntary, self-selected responses and therefore has serious potential for bias.
10) C
11) A
12) D
13) A
14) B
15) B
16) E
17) D
18) C
19) This is cluster sampling. The sample obtained will not be a simple random sample of all high school teachers in the district because different samples have different chances of being selected.
20) Answers will vary. Possible answer: Cluster sampling can save time and money and be more efficient, especially when the clusters are geographically far apart from each other. For example, if a researcher wishes to interview a sample of high school teachers in a school district, it will be easier to interview all the teachers at a few schools than to interview a few teachers from many different schools.
CHAPTER 1 FORM C

Name:________________________ Course Number:____________ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Define observational study and experiment. Define the terms “treatment group” and “control group” as part of your answer.

Determine whether the given value is a statistic or a parameter.

2) After taking the first exam, 15 of the students dropped the class.
   A) Statistic             B) Parameter

Identify the number as either continuous or discrete.

3) The average height of all freshmen entering college in a certain year is 68.4 inches.
   A) Continuous           B) Discrete

Determine which of the four levels of measurement (nominal, ordinal, interval, ratio) is most appropriate.

4) Nationalities of survey respondents.
   A) Ratio                B) Interval         C) Nominal       D) Ordinal

5) Temperatures of the ocean at various depths.
   A) Ratio                B) Nominal          C) Interval      D) Ordinal

Identify the sample and population. Also, determine whether the sample is likely to be representative of the population.

6) 100,000 randomly selected adults were asked whether they drink at least 48 oz of water each day and only 45% said yes.
CHAPTER 1 FORM C

7) A study of achievement scores by sixth-grade students on a standardized math test showed the three top scorers were all gifted piano players. Conclusion: Playing the piano leads to mathematical achievement.

Use critical thinking to address the key issue.

8) A researcher published this survey result: “74% of people would be willing to spend 10 percent more for energy from a non-polluting source”. The survey question was announced on a national radio show and 1,200 listeners responded by calling in. What is wrong with this survey?

9) A researcher wished to gauge public opinion on gun control. He randomly selected 1000 people from among registered voters and asked them the following question: “Do you believe that gun control laws which restrict the ability of Americans to protect their families should be eliminated?”. Identify the abuse of statistics and suggest a way the researcher’s methods could be improved.

Perform the requested conversions. Round decimals to the nearest thousandth and percents to the nearest tenth of a percent, if necessary.

10) Convert 0.328 to an equivalent fraction and percent.
   A) \(\frac{41}{125}\) 32.8%  B) \(\frac{8}{25}\) 3.28%  C) \(\frac{41}{125}\), 3.28%  D) \(\frac{8}{25}\) 32.8%

11) Convert 2.75 to an equivalent fraction and percent.
   A) \(\frac{3}{4}\) 27.5%  B) \(\frac{1}{2}\) 27.5%  C) \(\frac{3}{4}\) 275%  D) \(\frac{1}{2}\) 275%
CHAPTER 1 FORM C

Solve the problem.

12) Alex and Juana went on a 95-mile canoe trip with their class. On the first day they traveled 19 miles. What percent of the total distance did they canoe?
   A) 0.2%  B) 20%  C) 500%  D) 5%

Is the description an observational study or an experiment?

13) A T.V. show’s executives commissioned a study to gauge the impact of the show’s ratings on the sales of its advertisers.
   A) Experiment  B) Observational study

14) A doctor performs several diagnostic tests to determine the reason for a patient's illness.
   A) Experiment  B) Observational study

Identify the type of observational study.

15) Researchers collect data by interviewing athletes who have won olympic gold medals from 1980 to 1992.
   A) Cross-sectional  B) Retrospective
   C) Prospective  D) None of these

Identify which of these types of sampling is used: random, stratified, systematic, cluster, convenience.

16) A tax auditor selects every 1000th income tax return that is received.
   A) Random
   B) Systematic
   C) Stratified
   D) Convenience
   E) Cluster

17) A pollster uses a computer to generate 500 random numbers, then interviews the voters corresponding to those numbers.
   A) Cluster
   B) Convenience
   C) Systematic
   D) Random
   E) Stratified
CHAPTER 1 FORM C

18) To avoid working late, a quality control analyst simply inspects the first 100 items produced in a day.
   A) Random
   B) Systematic
   C) Stratified
   D) Cluster
   E) Convenience

Provide an appropriate response.

19) A market researcher obtains a sample of 50 people by standing outside a store and asking every 20th person who enters the store to fill out a survey until she has 50 people. What sampling method is being used here? Will the resulting sample be a random sample? Will it be a simple random sample? Explain your thinking.

20) A teacher at a school obtains a sample of students by selecting a random sample of 20 students from each grade. What kind of sampling is being used here? Will the resulting sample be a simple random sample of the population of students at the school? Explain your thinking.
Answer Key
Testname: CHAPTER 1 FORM C

1) In an observational study, we observe and measure specific characteristics, but we don’t attempt to manipulate or modify the subjects being studied. In an experiment we apply some treatment and then proceed to observe its effects on the subjects. In the experiment, the group receiving the treatment is called the treatment group. The control group is the group that is not given the treatment.

2) B
3) A
4) C
5) C
6) Sample: the 100,000 selected adults; population: all adults; representative
7) A sample of 3 among many students is not sufficient to conclude that playing the piano is conducive to math achievement. Student motivation and interest in math should be considered as factors.
8) This is not a random sample. The survey is based on voluntary, self-selected responses and therefore has serious potential for bias.
9) The question is loaded. A more neutral way to phrase the question would be, for example, "Do you believe that gun control laws should be strengthened, weakened, or left in their current form?".

10) A
11) C
12) B
13) B
14) A
15) B
16) B
17) D
18) E
19) This is systematic sampling. The sample obtained will be a random sample because everyone has the same chance of being chosen but will not be a simple random sample as different samples of 50 people have different chances of being chosen.
   Note: That the sample is random depends on the market researcher randomly selecting 20 as the starting point prior to research.
20) This is stratified sampling. The sample obtained will not be a simple random sample because different samples of students have different chances of being selected.
CHAPTER 2 FORM A

Name: _________________________ Course Number: _______ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Histograms and Pareto charts are both bar charts. What is the significant difference between the two?

2) Suppose that a data set has a minimum value of 25 and a max of 75 and that you want 5 classes. Explain how to find the class width for this frequency table. What happens if you mistakenly use a class width of 10 instead of 11?

Solve the problem.

3) Using the employment information in the table on Alpha Corporation, determine the width of each class.

<table>
<thead>
<tr>
<th>Years employed at Alpha Corporation</th>
<th>Frequency (No. of employees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Limits</td>
<td></td>
</tr>
<tr>
<td>(years of service)</td>
<td></td>
</tr>
<tr>
<td>1 - 5</td>
<td>5</td>
</tr>
<tr>
<td>6 - 10</td>
<td>20</td>
</tr>
<tr>
<td>11 - 15</td>
<td>25</td>
</tr>
<tr>
<td>16 - 20</td>
<td>10</td>
</tr>
<tr>
<td>21 - 25</td>
<td>5</td>
</tr>
<tr>
<td>26 - 30</td>
<td>3</td>
</tr>
</tbody>
</table>

A) 10  B) 4  C) 6  D) 5
4) Using the information in the table on home sale prices in the city of Summerhill for the month of June, find the class boundaries for class 80.0–110.9.

<table>
<thead>
<tr>
<th>Class Limits (Sale price in thousands)</th>
<th>Frequency (No. of homes sold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.0 - 110.9</td>
<td>2</td>
</tr>
<tr>
<td>111.0 - 141.9</td>
<td>5</td>
</tr>
<tr>
<td>142.0 - 172.9</td>
<td>7</td>
</tr>
<tr>
<td>173.0 - 203.9</td>
<td>10</td>
</tr>
<tr>
<td>204.0 - 234.9</td>
<td>3</td>
</tr>
<tr>
<td>235.0 - 265.9</td>
<td>1</td>
</tr>
</tbody>
</table>

A) 79.90, 111.0  B) 79.95, 110.95  C) 79.90, 110.95  D) 80.00, 110.95

Construct the relative frequency distribution that corresponds to the given frequency distribution.

5)

<table>
<thead>
<tr>
<th>Incomes</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>200–300</td>
<td>69</td>
</tr>
<tr>
<td>301–400</td>
<td>73</td>
</tr>
<tr>
<td>401–500</td>
<td>81</td>
</tr>
<tr>
<td>501–600</td>
<td>65</td>
</tr>
<tr>
<td>&gt;600</td>
<td>16</td>
</tr>
</tbody>
</table>

A) Incomes | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200–300</td>
<td>22.70%</td>
</tr>
<tr>
<td>301–400</td>
<td>24.01%</td>
</tr>
<tr>
<td>401–500</td>
<td>26.64%</td>
</tr>
<tr>
<td>501–600</td>
<td>21.38%</td>
</tr>
<tr>
<td>&gt;600</td>
<td>5.26%</td>
</tr>
</tbody>
</table>

B) Incomes | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200–300</td>
<td>12.5%</td>
</tr>
<tr>
<td>301–400</td>
<td>20.1%</td>
</tr>
<tr>
<td>401–500</td>
<td>37.3%</td>
</tr>
<tr>
<td>501–600</td>
<td>15.2%</td>
</tr>
<tr>
<td>&gt;600</td>
<td>14.9%</td>
</tr>
</tbody>
</table>

C) Incomes | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>201–300</td>
<td>15.5%</td>
</tr>
<tr>
<td>301–400</td>
<td>22.1%</td>
</tr>
<tr>
<td>401–500</td>
<td>31.3%</td>
</tr>
<tr>
<td>501–600</td>
<td>16.2%</td>
</tr>
<tr>
<td>&gt;600</td>
<td>14.9%</td>
</tr>
</tbody>
</table>

D) Incomes | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200–300</td>
<td>26.21%</td>
</tr>
<tr>
<td>301–400</td>
<td>21.59%</td>
</tr>
<tr>
<td>401–500</td>
<td>5.30%</td>
</tr>
<tr>
<td>501–600</td>
<td>22.40%</td>
</tr>
<tr>
<td>&gt;600</td>
<td>26.30%</td>
</tr>
</tbody>
</table>
CHAPTER 2 FORM A

Solve the problem.

6) Use the high closing values of Naristar Inc. stock from the years 1990 – 2001 to construct a time-series graph. (Let x = 0 stand for 1990 and so on...) Identify a trend.

<table>
<thead>
<tr>
<th>Year</th>
<th>High</th>
<th>Year</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>42</td>
<td>1996</td>
<td>47</td>
</tr>
<tr>
<td>1991</td>
<td>40</td>
<td>1997</td>
<td>60</td>
</tr>
<tr>
<td>1992</td>
<td>31</td>
<td>1998</td>
<td>61</td>
</tr>
<tr>
<td>1993</td>
<td>42</td>
<td>1999</td>
<td>57</td>
</tr>
<tr>
<td>1994</td>
<td>44</td>
<td>2000</td>
<td>54</td>
</tr>
<tr>
<td>1995</td>
<td>47</td>
<td>2001</td>
<td>30</td>
</tr>
</tbody>
</table>

Construct the cumulative frequency distribution that corresponds to the given frequency distribution.

7) 

<table>
<thead>
<tr>
<th>Height (inches)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>69.0 - 71.9</td>
<td>15</td>
</tr>
<tr>
<td>72.0 - 74.9</td>
<td>17</td>
</tr>
<tr>
<td>75.0 - 77.9</td>
<td>16</td>
</tr>
<tr>
<td>78.0 - 80.9</td>
<td>15</td>
</tr>
<tr>
<td>81.0 - 83.9</td>
<td>17</td>
</tr>
</tbody>
</table>

A) 

<table>
<thead>
<tr>
<th>Height (inches)</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>69.0 - 71.9</td>
<td>0.188</td>
</tr>
<tr>
<td>72.0 - 74.9</td>
<td>0.212</td>
</tr>
<tr>
<td>75.0 - 77.9</td>
<td>0.200</td>
</tr>
<tr>
<td>78.0 - 80.9</td>
<td>0.188</td>
</tr>
<tr>
<td>81.0 - 83.9</td>
<td>0.212</td>
</tr>
</tbody>
</table>

B) 

<table>
<thead>
<tr>
<th>Height (inches)</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>69.0 - 71.9</td>
<td>15</td>
</tr>
<tr>
<td>72.0 - 74.9</td>
<td>32</td>
</tr>
<tr>
<td>75.0 - 77.9</td>
<td>48</td>
</tr>
<tr>
<td>78.0 - 80.9</td>
<td>63</td>
</tr>
<tr>
<td>81.0 - 83.9</td>
<td>80</td>
</tr>
</tbody>
</table>

C) 

<table>
<thead>
<tr>
<th>Height (inches)</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>69.0 - 71.9</td>
<td>15</td>
</tr>
<tr>
<td>72.0 - 74.9</td>
<td>32</td>
</tr>
<tr>
<td>75.0 - 77.9</td>
<td>48</td>
</tr>
<tr>
<td>78.0 - 80.9</td>
<td>61</td>
</tr>
<tr>
<td>81.0 - 83.9</td>
<td>80</td>
</tr>
</tbody>
</table>

D) 

<table>
<thead>
<tr>
<th>Height (inches)</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>69.0 - 71.9</td>
<td>32</td>
</tr>
<tr>
<td>72.0 - 74.9</td>
<td>48</td>
</tr>
<tr>
<td>75.0 - 77.9</td>
<td>63</td>
</tr>
<tr>
<td>78.0 - 80.9</td>
<td>80</td>
</tr>
<tr>
<td>81.0 - 83.9</td>
<td>97</td>
</tr>
</tbody>
</table>
CHAPTER 2 FORM A

Provide an appropriate response.

8) Sturges' guideline suggests that when constructing a frequency distribution, the ideal number of classes can be approximated by \(1 + \frac{\log n}{\log 2}\), where \(n\) is the number of data values. Use this guideline to find the ideal number of classes when the number of data values is 50. Round your answer to the nearest whole number.

A) 9  
B) 6  
C) 8  
D) 7

A nurse measured the blood pressure of each person who visited her clinic. Following is a relative-frequency histogram for the systolic blood pressure readings for those people aged between 25 and 40. Use the histogram to answer the question. The blood pressure readings were given to the nearest whole number.

9) Approximately what percentage of the people aged 25-40 had a systolic blood pressure reading between 110 and 119 inclusive?

A) 30%  
B) 3.5%  
C) 35%  
D) 0.35%
CHAPTER 2 FORM A

Provide an appropriate response.

10) Suppose that a histogram is constructed for the frequency distribution shown below:

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>30–39</td>
<td>11</td>
</tr>
<tr>
<td>40–49</td>
<td>23</td>
</tr>
<tr>
<td>50–59</td>
<td>17</td>
</tr>
<tr>
<td>60–69</td>
<td>12</td>
</tr>
<tr>
<td>70–89</td>
<td>6</td>
</tr>
</tbody>
</table>

The class 60–69 has twice the frequency of the class 70–89. In the histogram, will the area of the bar for the class 60–69 be twice the area of the bar for the class 70–89? In other words, will areas be proportional to frequencies in this histogram? Explain your thinking. Are there any conditions under which areas are proportional to frequencies in histograms?

Use the pie chart to solve the problem.

11) A survey of the 9225 vehicles on the campus of State University yielded the following circle graph.

Motorcycles 11%
Convertibles 14%
Vans 9%
Sedans 6%
Hatchbacks 31%
Pickups 29%

Find the number of hatchbacks. Round your result to the nearest whole number.
A) 268  B) 286  C) 2860  D) 2675
CHAPTER 2 FORM A

12) The pie chart below gives the inventory of the men's department of a store.

What is the total inventory?
A) $121,380  B) $180,336  C) $175,712  D) $171,088

Find the original data from the stem-and-leaf plot.

13)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>58</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>55</td>
</tr>
</tbody>
</table>

A) 85, 88, 91, 98, 105, 105
B) 85, 88, 91, 91, 105, 105
C) 81, 85, 81, 98, 108, 105
D) 85, 81, 88, 91, 101, 105

14)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>258</td>
</tr>
<tr>
<td>74</td>
<td>239</td>
</tr>
<tr>
<td>75</td>
<td>18</td>
</tr>
</tbody>
</table>

A) 73258, 74239, 7518
B) 75, 78, 81, 76, 77, 83, 76, 83
C) 732, 735, 738, 742, 743, 749, 751, 758
D) 732, 735, 748, 742, 743, 749, 751, 768
CHAPTER 2 FORM A

Construct the dot plot for the given data.

15) 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

   A) 
   B) 
   C) 
   D) 

Use the data to create a stemplot.

16) The midterm test scores for the seventh-period typing class are listed below.
   85 77 93 91 74 65 68 97 88 59 74 83 85 72 63 79

   A)  
   B)
CHAPTER 2 FORM A

Use the given paired data to construct a scatterplot.

17) x  2  -1  -1  -6  3  3  1  8  -5  -1
   y  3  -3  -1  -9  2  -1  -4  2  -3  -2
18) A car dealer is deciding what kinds of vehicles he should order from the factory. He looks at his sales report for the preceding period. Choose the vertical scale so that the relative frequencies are represented.

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>38</td>
</tr>
<tr>
<td>Sports</td>
<td>9.5</td>
</tr>
<tr>
<td>Family</td>
<td>66.5</td>
</tr>
<tr>
<td>Luxury</td>
<td>19</td>
</tr>
<tr>
<td>Truck</td>
<td>57</td>
</tr>
</tbody>
</table>

Construct a Pareto chart to help him decide.
CHAPTER 2 FORM A

Construct a pie chart representing the given data set.

19) The following data give the distribution of the types of houses in a town containing 18,000 houses.

<table>
<thead>
<tr>
<th>Capes</th>
<th>Garrisons</th>
<th>Splits</th>
</tr>
</thead>
<tbody>
<tr>
<td>4500</td>
<td>6300</td>
<td>7200</td>
</tr>
</tbody>
</table>

A) 

B) 

Provide an appropriate response.

20) One purpose of displaying data graphically is to provide clues about trends. The given values are weights (ounces) of steaks listed on a restaurant menu as "20 ounce porterhouse" steaks. The weights are supposed to be 21 ounces because they supposedly lose an ounce when cooked. Create a frequency distribution with 5 classes. Based on your distribution, comment on the advertised "20 ounce" steaks.

17 20 21 18 20 20 18 19 19 20 19 21 20 18 20 19 18 19
1) Answers will vary. Possible answer: Histograms convey quantitative information about shapes of distributions. Pareto charts convey comparative information about relative standing of categorical data.

2) Since the range is $75 - 25 = 50$, and 50 divided by 5 equals 10, a whole number, the class width has to be widened from 10 to 11. In that way, data values equal to 75 will not be omitted from the frequency table.

3) D
4) B
5) A

6) 


7) B
8) D
9) C

10) The areas of the bars for the two classes will actually be the same. This is because the bar for the class 60-69, while it is twice as tall as the bar for the class 70-89, is also only half the width because the class widths are not the same. Heights, not areas, are proportional to frequencies. For classes of equal width, areas will also be proportional to frequencies.

11) C
12) C
13) A
14) C
15) C
16) B
17) B
18) A
19) A
20) Answers will vary. Possible answer: The frequency distribution shows that half of the cooked steaks are less than their advertised weights.
CHAPTER 2 FORM B

Name:_________________________ Course Number:_______ Section Number:______

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Describe at least two advantages to using stemplots rather than frequency distributions.

2) Suppose that a data set has a minimum value of 18 and a max of 83 and that you want 5 classes. Explain how to find the class width for this frequency table. What happens if you mistakenly use a class width of 13 instead of 14?

Solve the problem.

3) The following frequency distribution analyzes the scores on a math test. Find the indicated class boundaries.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>40–59</td>
<td>2</td>
</tr>
<tr>
<td>60–75</td>
<td>4</td>
</tr>
<tr>
<td>76–82</td>
<td>6</td>
</tr>
<tr>
<td>83–94</td>
<td>15</td>
</tr>
<tr>
<td>95–99</td>
<td>5</td>
</tr>
</tbody>
</table>

The class boundaries of scores interval 95–99

A) 94.5, 99.5  B) 95.5, 100.5  C) 94.5, 100.5  D) 95.5, 99.5
CHAPTER 2 FORM B

4) Using the information in the table on home sale prices in the city of Summerhill for the month of June, determine the width of each class.

<table>
<thead>
<tr>
<th>Class Limits (Sale price in thousands)</th>
<th>Frequency (No. of homes sold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.0 - 110.9</td>
<td>2</td>
</tr>
<tr>
<td>111.0 - 141.9</td>
<td>5</td>
</tr>
<tr>
<td>142.0 - 172.9</td>
<td>7</td>
</tr>
<tr>
<td>173.0 - 203.9</td>
<td>10</td>
</tr>
<tr>
<td>204.0 - 234.9</td>
<td>3</td>
</tr>
<tr>
<td>235.0 - 265.9</td>
<td>1</td>
</tr>
</tbody>
</table>

A) 31  B) 61  C) 28  D) 30

Construct the relative frequency distribution that corresponds to the given frequency distribution.

5)

<table>
<thead>
<tr>
<th>Scores</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>91-100</td>
<td>3</td>
</tr>
<tr>
<td>81-90</td>
<td>5</td>
</tr>
<tr>
<td>71-80</td>
<td>12</td>
</tr>
<tr>
<td>61-70</td>
<td>5</td>
</tr>
<tr>
<td>&lt;61</td>
<td>2</td>
</tr>
</tbody>
</table>

A) Scores | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>91-100</td>
<td>12.5%</td>
</tr>
<tr>
<td>81-90</td>
<td>20.1%</td>
</tr>
<tr>
<td>71-80</td>
<td>37.3%</td>
</tr>
<tr>
<td>61-70</td>
<td>15.2%</td>
</tr>
<tr>
<td>&lt;61</td>
<td>14.9%</td>
</tr>
</tbody>
</table>

B) Scores | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>91-100</td>
<td>11.11%</td>
</tr>
<tr>
<td>81-90</td>
<td>18.52%</td>
</tr>
<tr>
<td>71-80</td>
<td>44.44%</td>
</tr>
<tr>
<td>61-70</td>
<td>18.52%</td>
</tr>
<tr>
<td>&lt;61</td>
<td>7.41%</td>
</tr>
</tbody>
</table>

C) Scores | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>91-100</td>
<td>0.22%</td>
</tr>
<tr>
<td>81-90</td>
<td>0.07%</td>
</tr>
<tr>
<td>71-80</td>
<td>0.63%</td>
</tr>
<tr>
<td>61-70</td>
<td>0.07%</td>
</tr>
<tr>
<td>&lt;61</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

D) Scores | Relative Frequency |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>91-100</td>
<td>15.5%</td>
</tr>
<tr>
<td>81-90</td>
<td>22.1%</td>
</tr>
<tr>
<td>71-80</td>
<td>31.3%</td>
</tr>
<tr>
<td>61-70</td>
<td>16.2%</td>
</tr>
<tr>
<td>&lt;61</td>
<td>14.9%</td>
</tr>
</tbody>
</table>
CHAPTER 2 FORM B

Solve the problem.

6) Use the high closing values of Naristar Inc. stock from the years 1992 – 2003 to construct a time-series graph. (Let \( x = 0 \) stand for 1992 and so on...) Identify a trend.

<table>
<thead>
<tr>
<th>Year</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>48</td>
</tr>
<tr>
<td>1993</td>
<td>53</td>
</tr>
<tr>
<td>1994</td>
<td>47</td>
</tr>
<tr>
<td>1995</td>
<td>55</td>
</tr>
<tr>
<td>1996</td>
<td>58</td>
</tr>
<tr>
<td>1997</td>
<td>61</td>
</tr>
<tr>
<td>1998</td>
<td>62</td>
</tr>
<tr>
<td>1999</td>
<td>60</td>
</tr>
<tr>
<td>2000</td>
<td>68</td>
</tr>
<tr>
<td>2001</td>
<td>42</td>
</tr>
<tr>
<td>2002</td>
<td>51</td>
</tr>
<tr>
<td>2003</td>
<td>78</td>
</tr>
</tbody>
</table>

Construct the cumulative frequency distribution that corresponds to the given frequency distribution.

7)

<table>
<thead>
<tr>
<th>Days of vacation</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 3</td>
<td>16</td>
</tr>
<tr>
<td>4 - 7</td>
<td>20</td>
</tr>
<tr>
<td>8 - 11</td>
<td>14</td>
</tr>
<tr>
<td>12 - 15</td>
<td>24</td>
</tr>
<tr>
<td>16 - 19</td>
<td>26</td>
</tr>
</tbody>
</table>

A) Days of vacation | Cumulative Frequency | B) Days of vacation | Cumulative Frequency
|-------------------|----------------------|---------------------|----------------------
| 0 - 3             | 16                   | 0 - 3               | 36                   |
| 4 - 7             | 36                   | 4 - 7               | 50                   |
| 8 - 11            | 51                   | 8 - 11              | 74                   |
| 12 - 15           | 75                   | 12 - 15             | 100                  |
| 16 - 19           | 100                  | 16 - 19             | 126                  |

C) Days of vacation | Cumulative Frequency | D) Days of vacation | Cumulative Frequency
|-------------------|----------------------|---------------------|----------------------
| 0 - 3             | 16                   | 0 - 3               | 0.16                 |
| 4 - 7             | 36                   | 4 - 7               | 0.2                  |
| 8 - 11            | 50                   | 8 - 11              | 0.14                 |
| 12 - 15           | 74                   | 12 - 15             | 0.24                 |
| 16 - 19           | 100                  | 16 - 19             | 0.26                 |

Provide an appropriate response.

8) Sturges’ guideline suggests that when constructing a frequency distribution, the ideal number of classes can be approximated by \( 1 + (\log n)/(\log 2) \), where \( n \) is the number of data values. Use this guideline to find the ideal number of classes when the number of data values is 100. Round your answer to the nearest whole number.

A) 6  B) 9  C) 7  D) 8

29
CHAPTER 2 FORM B

A nurse measured the blood pressure of each person who visited her clinic. Following is a relative-frequency histogram for the systolic blood pressure readings for those people aged between 25 and 40. Use the histogram to answer the question. The blood pressure readings were given to the nearest whole number.

![Relative Frequency](image)

9) Approximately what percentage of the people aged 25–40 had a systolic blood pressure reading between 110 and 139 inclusive?

A) 39%   B) 74%   C) 59%   D) 89%

Provide an appropriate response.

10) Construct a frequency distribution and the corresponding histogram in which the following conditions are satisfied:
- The frequency for the second class is twice the frequency of the first class.
- In the histogram, the area of the bar corresponding to the second class is four times the area of the bar corresponding to the first class.

What do you know about the width of the second class in relation to that of the first class?
CHAPTER 2 FORM B

Use the pie chart to solve the problem.

11) A survey of the 6054 vehicles on the campus of State University yielded the following circle graph.

<table>
<thead>
<tr>
<th>Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motorcycles</td>
<td>11%</td>
</tr>
<tr>
<td>Convertibles</td>
<td>14%</td>
</tr>
<tr>
<td>Vans</td>
<td>6%</td>
</tr>
<tr>
<td>Sedans</td>
<td>4%</td>
</tr>
<tr>
<td>Hatchbacks</td>
<td>36%</td>
</tr>
<tr>
<td>Pickups</td>
<td>29%</td>
</tr>
</tbody>
</table>

Find the number of vans. Round your result to the nearest whole number.

A) 304  B) 363  C) 3632  D) 3625

12) The pie chart below gives the inventory of the men’s department of a store.

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suits</td>
<td>$20,650</td>
</tr>
<tr>
<td>Trousers</td>
<td>$18,585</td>
</tr>
<tr>
<td>Underwear</td>
<td>$2478</td>
</tr>
<tr>
<td>Shirts</td>
<td>$3717</td>
</tr>
<tr>
<td>Socks</td>
<td>$1239</td>
</tr>
<tr>
<td>Sweaters</td>
<td>$7847</td>
</tr>
<tr>
<td>Ties</td>
<td>$3304</td>
</tr>
</tbody>
</table>

What is the total inventory?

A) $57,720  B) $58,820  C) $58,720  D) $57,820

Find the original data from the stem-and-leaf plot.

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>2 5 8</td>
</tr>
<tr>
<td>74</td>
<td>2 3 9</td>
</tr>
<tr>
<td>75</td>
<td>1 8</td>
</tr>
</tbody>
</table>

A) 75, 78, 81, 76, 77, 83, 76, 83  B) 732, 735, 738, 742, 743, 749, 751, 758

C) 732, 735, 748, 742, 743, 749, 751, 768  D) 73258, 74239, 7518
CHAPTER 2 FORM B

Find the original data from the stem-and-leaf plot.

14)  

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3</td>
<td>17</td>
</tr>
<tr>
<td>5.4</td>
<td>79</td>
</tr>
<tr>
<td>5.5</td>
<td>199</td>
</tr>
</tbody>
</table>

A) 5.31, 5.32, 5.47, 5.49, 5.53, 5.59, 5.59  
B) 0.63, 1.23, 1.24, 1.44, 0.65, 1.45, 1.45  
C) 5.31, 5.37, 5.47, 5.49, 5.51, 5.59, 5.59  
D) 0.63, 0.63, 1.24, 1.24, 1.44, 0.65, 1.45, 1.46

Construct the dot plot for the given data.

15) A store manager counts the number of customers who make a purchase in his store each day. The data are as follows.

10 11 8 14 7 10 11 8 7

A)  
B)  
C)  
D)  

5 10 15
CHAPTER 2 FORM B

Use the data to create a stemplot.

16) The attendance counts for this season’s basketball games are listed below.
   227 239 215 219
   221 233 229 233
   235 228 245 231

   A)  
   21 5 9  
   22 1 7 8 9  
   23 1 3 3 5 9  
   24 5  

   B)  
   21 5 7 9  
   22 1 8 9  
   23 1 3 3 5 9  
   24 5

17) The normal monthly precipitation (in inches) for August is listed for 39 different U.S. cities. Construct an expanded stemplot with about 9 rows.
   3.5 1.6 2.4 3.7 4.1 3.9 1.0 3.6 1.7 0.4 3.2 4.2 4.1 4.2 3.4 3.7 2.2 1.5 4.2 3.4 2.7 4.0 2.0 0.8 3.6 3.7 0.4 3.7 2.0 3.6 3.8 1.2 4.0 3.1 0.5 3.9 0.1 3.5 3.4

   A)  
   0. 0 1 4 4  
   0. 5 8  
   1. 0 2  
   1. 5 6 7  
   2. 0 0 2 4  
   2. 7 7 7  
   3. 1 2 4 4 4  
   3. 5 5 6 6 6 7 7 8 9  
   4. 0 0 1 1 2 2 2

   B)  
   0. 1 4 4  
   0. 5 8  
   1. 0 2  
   1. 5 6 7  
   2. 0 0 2 4  
   2. 7  
   3. 1 2 4 4 4  
   3. 5 5 6 6 6 7 7 7 8 9 9  
   4. 0 0 1 1 2 2 2
Construct a pie chart representing the given data set.

18) After reviewing a movie, 900 people rated the movie as excellent, good, or fair. The following data give the rating distribution.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>180</td>
</tr>
<tr>
<td>Good</td>
<td>450</td>
</tr>
<tr>
<td>Fair</td>
<td>270</td>
</tr>
</tbody>
</table>

A) ![Pie chart A]

B) ![Pie chart B]
Solve the problem.

19) At the National Criminologists Association’s annual convention, participants filled out a questionnaire asking what they thought was the most important cause for criminal behavior. The tally was as follows.

<table>
<thead>
<tr>
<th>Cause</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>education</td>
<td>15.4</td>
</tr>
<tr>
<td>drugs</td>
<td>46.2</td>
</tr>
<tr>
<td>family</td>
<td>30.8</td>
</tr>
<tr>
<td>poverty</td>
<td>53.9</td>
</tr>
<tr>
<td>other</td>
<td>7.7</td>
</tr>
</tbody>
</table>

Make a Pareto chart to display these findings.

A) ![Pareto Chart A]

B) ![Pareto Chart B]

C) ![Pareto Chart C]

D) ![Pareto Chart D]
CHAPTER 2 FORM B

Use the given paired data to construct a scatterplot.

20) x 0.31 0.88 0.49 0.23 0.14 0.72 -0.02 0.2
    y 0.33 0.12 0.76 0.65 -0.37 0.47 0.72 0.21
Answer Key
Testname: CHAPTER 2 FORM B

1) Answers will vary. Possible answer: The shape of a distribution can readily be seen. The plot can be drawn quicker, since class width need not be calculated.
2) Data values equal to 83 would not be included in the frequency distribution.
3) A
4) A
5) B
6) Trend: Answers will vary. Possible answer: Except for a drop in high closing value in 1994, there was a steady rise through 2000, after which there was a sharp drop in 2001 followed by increases through 2003.

7) C
8) D
9) B
10) Answers will vary. The class width of the second class should be twice the class width of the first class.
11) B
12) D
13) B
14) C
15) B
16) A
17) B
18) A
19) B
20) C
CHAPTER 2 FORM C

Name: __________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Suppose you are comparing frequency data for two different groups, 25 managers and 150 blue collar workers. Why would a relative frequency distribution be better than a frequency distribution?

2) Create an example displaying data in a pie chart. Display the same data in a Pareto chart. Which graph is more effective? List at least two reasons in support of your choice.

Solve the problem.

3) Using the employment information in the table on Alpha Corporation, find the class boundaries for class 26–30.

<table>
<thead>
<tr>
<th>Years employed at Alpha Corporation</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Limits</td>
<td>(No. of employees)</td>
</tr>
<tr>
<td>(years of service)</td>
<td></td>
</tr>
<tr>
<td>1 - 5</td>
<td>5</td>
</tr>
<tr>
<td>6 - 10</td>
<td>20</td>
</tr>
<tr>
<td>11 - 15</td>
<td>25</td>
</tr>
<tr>
<td>16 - 20</td>
<td>10</td>
</tr>
<tr>
<td>21 - 25</td>
<td>5</td>
</tr>
<tr>
<td>26 - 30</td>
<td>3</td>
</tr>
</tbody>
</table>

A) 25.5, 20.5  B) 26.5, 29.5  C) 26.5, 30.5  D) 25.5, 30.5
4) Using the information in the table on home sale prices in the city of Summerhill for the month of June, find the class midpoint for class 235.0–265.9.

<table>
<thead>
<tr>
<th>Class Limits (Sale price in thousands)</th>
<th>Frequency (No. of homes sold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.0 – 110.9</td>
<td>2</td>
</tr>
<tr>
<td>111.0 – 141.9</td>
<td>5</td>
</tr>
<tr>
<td>142.0 – 172.9</td>
<td>7</td>
</tr>
<tr>
<td>173.0 – 203.9</td>
<td>10</td>
</tr>
<tr>
<td>204.0 – 234.9</td>
<td>3</td>
</tr>
<tr>
<td>235.0 – 265.9</td>
<td>1</td>
</tr>
</tbody>
</table>

A) 250.50    B) 250.45    C) 250.40    D) 250.55

Construct the cumulative frequency distribution that corresponds to the given frequency distribution.

5)

<table>
<thead>
<tr>
<th>Days of vacation</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 3</td>
<td>16</td>
</tr>
<tr>
<td>4 – 7</td>
<td>20</td>
</tr>
<tr>
<td>8 – 11</td>
<td>14</td>
</tr>
<tr>
<td>12 – 15</td>
<td>24</td>
</tr>
<tr>
<td>16 – 19</td>
<td>26</td>
</tr>
</tbody>
</table>

A) Days of vacation  Cumulative Frequency  B) Days of vacation  Cumulative Frequency
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 3</td>
<td>36</td>
<td>0 – 3</td>
<td>0.16</td>
</tr>
<tr>
<td>4 – 7</td>
<td>50</td>
<td>4 – 7</td>
<td>0.2</td>
</tr>
<tr>
<td>8 – 11</td>
<td>74</td>
<td>8 – 11</td>
<td>0.14</td>
</tr>
<tr>
<td>12 – 15</td>
<td>100</td>
<td>12 – 15</td>
<td>0.24</td>
</tr>
<tr>
<td>16 – 19</td>
<td>126</td>
<td>16 – 19</td>
<td>0.26</td>
</tr>
</tbody>
</table>

C) Days of vacation  Cumulative Frequency  D) Days of vacation  Cumulative Frequency
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 3</td>
<td>16</td>
<td>0 – 3</td>
<td>16</td>
</tr>
<tr>
<td>4 – 7</td>
<td>36</td>
<td>4 – 7</td>
<td>36</td>
</tr>
<tr>
<td>8 – 11</td>
<td>50</td>
<td>8 – 11</td>
<td>51</td>
</tr>
<tr>
<td>12 – 15</td>
<td>74</td>
<td>12 – 15</td>
<td>75</td>
</tr>
<tr>
<td>16 – 19</td>
<td>100</td>
<td>16 – 19</td>
<td>100</td>
</tr>
</tbody>
</table>
CHAPTER 2 FORM C

Use the given data to construct a frequency distribution.

6) On a math test, the scores of 24 students were

\[
\begin{align*}
98 & \quad 72 & \quad 71 & \quad 64 & \quad 71 & \quad 98 & \quad 84 & \quad 71 & \quad 68 & \quad 81 & \quad 72 \\
72 & \quad 81 & \quad 71 & \quad 72 & \quad 81 & \quad 71 & \quad 72 & \quad 84 & \quad 72 & \quad 81 & \quad 64 \\
\end{align*}
\]

Construct a frequency table. Use 4 classes beginning with a lower class limit of 60.

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Provide an appropriate response.

7) Sturges' guideline suggests that when constructing a frequency distribution, the ideal number of classes can be approximated by \(1 + (\log n)/(\log 2)\), where \(n\) is the number of data values. Use this guideline to find the ideal number of classes when the number of data values is 150. Round your answer to the nearest whole number.

A) 9  
B) 8  
C) 7  
D) 6

Solve the problem.
A nurse measured the blood pressure of each person who visited her clinic. Following is a relative-frequency histogram for the systolic blood pressure readings for those people aged between 25 and 40. Use the histogram to answer the question. The blood pressure readings were given to the nearest whole number.

8) What common class width was used to construct the frequency distribution?

A) 100  
B) 11  
C) 9  
D) 10
CHAPTER 2 FORM C

Provide an appropriate response.

9) Construct a frequency distribution that includes an outlier. Construct the corresponding histogram. Then, construct the corresponding histogram without including the outlier. How much does the outlier affect the shape of the histogram?

Use the pie chart to solve the problem.

10) A survey of the 9189 vehicles on the campus of State University yielded the following circle graph.

![Pie Chart]

Motorcycles 10%
Convertibles 15%
Vans 9%
Sedans 8%
Hatchbacks 30%
Pickups 28%

Find the number of convertibles. Round your result to the nearest whole number.
A) 7811  
B) 15  
C) 919  
D) 1378
11) The pie chart below gives the number of students in the residence halls at the state university.

Write the ratio of the number of residents at Brown to the number of students at Adams.

A) \( \frac{144}{191} \)  \quad B) \( \frac{6}{5} \)  \quad C) \( \frac{5}{6} \)  \quad D) \( \frac{24}{191} \)

Use the data to create a stemplot.

12) The weights of 22 members of the varsity football team are listed below.

144 152 142 151 160 152 131 164 141 153 140
144 175 156 147 133 172 159 135 159 148 171

A)  

| 13 | 1 3 5 |
| 14 | 1 2 2 3 6 9 9 |
| 15 | 0 1 2 4 4 7 8 |
| 16 | 0 4 |
| 17 | 1 2 5 |

B)  

| 13 | 1 3 5 |
| 14 | 0 1 2 4 4 7 8 |
| 15 | 1 2 2 3 6 9 9 |
| 16 | 0 4 |
| 17 | 1 2 5 |
CHAPTER 2 FORM C

Construct the dot plot for the given data.

13) The frequency chart shows the distribution of defects for the machines used to produce a product.

<table>
<thead>
<tr>
<th>Defects</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

A) ![Dot Plot A](image)

B) ![Dot Plot B](image)

C) ![Dot Plot C](image)

D) ![Dot Plot D](image)
CHAPTER 2 FORM C

Use the data to create a stemplot.

14) The ages of the 45 members of a track and field team are listed below. Construct an expanded stemplot with about 8 rows.
21 18 42 35 32 21 44 25 38 48 14 19 23 22 28
32 34 27 31 17 16 41 37 22 24 33 32 21 26 30
22 27 32 30 20 18 17 21 15 26 36 31 40 16 25

A) 1 4 5
    1 5 6 6 7 7 8 8 9
    2 0 1 1 1 2 2 2 3 4 5 5
    2 5 5 6 6 7 7 8
    3 0 0 1 1 2 2 2 3 4 5
    3 5 6 7 8
    4 0 1 2 4
    4 8

B) 1 4
    1 5 6 6 7 7 8 8 9
    2 0 1 1 1 2 2 2 3 4 5 5
    2 5 5 6 6 7 7 8
    3 0 0 1 1 2 2 2 3 4 5
    3 5 6 7 8
    4 0 1 2 4
    4 8

Find the original data from the stem-and-leaf plot.

15)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1 8</td>
</tr>
<tr>
<td>5</td>
<td>1 3 6</td>
</tr>
<tr>
<td>6</td>
<td>1 3 3 8 9</td>
</tr>
<tr>
<td>7</td>
<td>3 4</td>
</tr>
</tbody>
</table>

A) 41, 48, 51, 51, 53, 56, 61, 63, 63, 68, 69, 73, 74
B) 41, 48, 53, 53, 56, 63, 64, 68, 69, 73, 74
C) 5, 12, 5, 5, 7, 10, 7, 7, 9, 14, 15, 10, 11, 14
D) 43, 43, 44, 51, 51, 53, 56, 61, 61, 73, 74

16)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>47 91</td>
</tr>
<tr>
<td>39</td>
<td>13 21 78</td>
</tr>
<tr>
<td>40</td>
<td>47 91</td>
</tr>
</tbody>
</table>

A) 85, 129, 60, 60, 117, 87, 131
B) 3847, 3891, 3913, 3921, 3978, 4047, 4091
C) 3847, 3848, 3914, 3921, 3978, 4047, 4092
D) 85, 129, 52, 60, 117, 87, 131

44
Solve the problem.

17) The Kappa Iota Sigma Fraternity polled its members on the weekend party theme. The vote was as follows: six for toga, four for hayride, eight for beer bash, and two for masquerade. Display the vote count in a Pareto chart.

A) 

![Graph A]

B) 

![Graph B]

C) 

![Graph C]

D) 

![Graph D]
CHAPTER 2 FORM C

Construct a pie chart representing the given data set.

18) The following figures give the distribution of land (in acres) for a county containing 61,000 acres.

<table>
<thead>
<tr>
<th>Forest</th>
<th>Farm</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>9150</td>
<td>6100</td>
<td>45,750</td>
</tr>
</tbody>
</table>

19) A college student wants to purchase one of two stocks. She has the average annual high values for each of these stocks over the most recent ten-year period. For comparison, she decides to sketch a time-series graph. How should she prepare her graph, and what should she look for?
CHAPTER 2 FORM C

Use the given paired data to construct a scatterplot.

20) x -6  4  9  5  11  7  1  -1  -5
    y  3  9  11  7  9  11  7  2  3
Answer Key
Testname: CHAPTER 2 FORM C

1) Answers will vary. Possible answer: A relative frequency distribution is better for comparison between groups whose numbers are different, since ratios are readily comparable.

2) Answers will vary. Answer should include the fact that pie charts are better for showing categories that are parts of a whole, whereas Pareto charts are better for displaying relative importance among categories.

3) D
4) B
5) C

6)

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 – 69</td>
<td>3</td>
</tr>
<tr>
<td>70 – 79</td>
<td>12</td>
</tr>
<tr>
<td>80 – 89</td>
<td>7</td>
</tr>
<tr>
<td>90 – 99</td>
<td>2</td>
</tr>
</tbody>
</table>

7) B
8) D
9) Answers will vary. Possible answer: The outlier does not drastically change the shape of the histogram. Rather, it appears far removed from the contiguous bars of the histogram.

10) D
11) B
12) B
13) B
14) B
15) A
16) B
17) D
18) A
19) The student should plot her data on a baseline marked by year and with vertical axis marked by high values. The stock that shows less volatility and a steady rise would be the better choice.

20) B
CHAPTER 3 FORM A

Name:__________________________ Course Number:________ Section Number:_____

**Directions:** Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

**Provide an appropriate response.**

1) The mean of a data set is always/sometimes/never (select one) one of the data points in a set of data. Explain your answer by giving at least one example.

2) Without calculating the standard deviation, compare the standard deviation for the following data sets. (Note: All data sets have a mean of 30.) Which do you expect to have the largest standard deviation and which do you expect to have the smallest standard deviation? Explain your answers in terms of the formula

\[ s = \sqrt{\frac{\sum(x - \overline{x})^2}{n - 1}}. \]

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>30, 30, 30, 30, 30, 30, 30, 30</td>
<td>( \sqrt{\frac{\sum(x - \overline{x})^2}{n - 1}} )</td>
</tr>
<tr>
<td>20, 25, 25, 30, 30, 30, 35, 35</td>
<td>( \sqrt{\frac{\sum(x - \overline{x})^2}{n - 1}} )</td>
</tr>
<tr>
<td>20, 20, 25, 25, 35, 35, 40, 40</td>
<td>( \sqrt{\frac{\sum(x - \overline{x})^2}{n - 1}} )</td>
</tr>
</tbody>
</table>

3) Marla scored 85% on her last unit exam in her statistics class. When Marla took the SAT exam, she scored at the 85 percentile in mathematics. Explain the difference in these two scores.

4) The normal monthly precipitation (in inches) for August is listed for 20 different U.S. cities. Find the mean of the data.

<table>
<thead>
<tr>
<th>Precipitation</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td></td>
</tr>
</tbody>
</table>

A) 2.80 in. B) 2.94 in. C) 3.27 in. D) 3.09 in.
CHAPTER 3 FORM A

Find the median for the given sample data.

5) The distances traveled (in miles) to 7 different swim meets are given below:
   12, 18, 31, 46, 69, 71, 85
   Find the median distance traveled.
   A) 31 miles    B) 69 miles    C) 46 miles    D) 47 miles

Find the mode(s) for the given sample data.

6) The weights (in ounces) of 14 different apples are shown below.
   6.8 5.9 5.7 6.5 5.4 6.8 5.9
   4.9 4.8 6.5 6.8 4.9 6.5 4.7
   A) 6.8    B) 6.65    C) 6.8, 6.5    D) None

Find the midrange for the given sample data.

7) Bill kept track of the number of hours he spent exercising each week. The results for 15
   weeks are shown below. Find the midrange.
   7.2 6.5 7.2 7.2 7.9
   8.0 6.5 8.2 8.5 7.2
   8.7 6.5 8.0 8.9 7.9
   A) 7.9    B) 2.4    C) 7.60    D) 7.70

Find the mean of the data summarized in the given frequency distribution.

8) The heights of a group of professional basketball players are summarized in the
   frequency distribution below. Find the mean height. Round your answer to one decimal
   place.

<table>
<thead>
<tr>
<th>Height (in.)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>70 – 71</td>
<td>2</td>
</tr>
<tr>
<td>72 – 73</td>
<td>5</td>
</tr>
<tr>
<td>74 – 75</td>
<td>9</td>
</tr>
<tr>
<td>76 – 77</td>
<td>13</td>
</tr>
<tr>
<td>78 – 79</td>
<td>8</td>
</tr>
<tr>
<td>80 – 81</td>
<td>4</td>
</tr>
<tr>
<td>82 – 83</td>
<td>2</td>
</tr>
</tbody>
</table>

   A) 7.46 in.    B) 76.4 in.    C) 78.2 in.    D) 13.5 in.

Solve the problem.

9) Elaine gets quiz grades of 67, 64, and 87. She gets a 84 on her final exam. Find the
   weighted mean if the quizzes each count for 15% and the final exam counts for 55% of
   the final grade.

   A) 78.9    B) 72.1    C) 78.3    D) 75.5
CHAPTER 3 FORM A

10) The mean salary of the female employees of one company is $29,525. The mean salary of the male employees of the same company is $33,470. Can the mean salary of all employees of the company be obtained by finding the mean of $29,525 and $33,470? Explain your thinking. Under what conditions would the mean of $29,525 and $33,470 yield the mean salary of all employees of the company?

Find the range for the given data.

11) The manager of an electrical supply store measured the diameters of the rolls of wire in the inventory. The diameters of the rolls (in m) are listed below.
0.165  0.114  0.503  0.392  0.579  0.311
Compute the range.
A) 0.114   B) 0.503   C) 0.146   D) 0.465

Find the variance for the given data. Round your answer to one more decimal place than the original data.

12) 5.0, 8.0, 4.9, 6.8, and 2.8
A) 3.96   B) 3.86   C) 3.17   D) 10.26

Solve the problem.

13) A distribution of data has a maximum value of 74, a median value of 58, and a minimum of 42. Use the range rule of thumb to find the standard deviation. Round results to the nearest tenth.
A) 16.0   B) 4.3   C) 6.4   D) 8.0

14) A company performs quality control on its juice bottles. It finds that the volumes of juice in its 16 ounce bottles have a mean of 16.3 ounces and a standard deviation of 0.09 ounces. Use the range rule of thumb to estimate the minimum and maximum "usual" volumes.
A) 16.12 ounces, 16.48 ounces   B) 16.21 ounces, 16.39 ounces
C) 16.07 ounces, 16.17 ounces   D) 16.03 ounces, 16.57 ounces

Use the empirical rule to solve the problem.

15) The amount of Jen’s monthly phone bill is normally distributed with a mean of $75 and a standard deviation of $9. What percentage of her phone bills are between $48 and $102?
A) 95.44%   B) 99.99%   C) 68.26%   D) 99.74%
CHAPTER 3 FORM A

Find the z-score corresponding to the given value and use the z-score to determine whether the value is unusual. Consider a score to be unusual if its z-score is less than -2.00 or greater than 2.00. Round the z-score to the nearest tenth if necessary.

16) A time for the 100 meter sprint of 15.0 seconds at a school where the mean time for the 100 meter sprint is 17.5 seconds and the standard deviation is 2.1 seconds.
   A) -1.2; not unusual   B) 1.2; not unusual
   C) -2.5; unusual       D) -1.2; unusual

Determine which score corresponds to the higher relative position.

17) Which is better, a score of 92 on a test with a mean of 71 and a standard deviation of 15, or a score of 688 on a test with a mean of 493 and a standard deviation of 150?
   A) A score of 688
   B) A score of 92
   C) Both scores have the same relative position.

Find the percentile for the data point.

18) Data set: 51 36 48 75 75 75 49;
    data point 51
    A) 43   B) 20   C) 57   D) 50

Provide an appropriate response.

19) If all the values in a data set are converted to z-scores, the shape of the distribution of the z-scores will be the same as the distribution of the original data. True or false?
    A) True   B) False
CHAPTER 3 FORM A

Construct a boxplot for the given data. Include values of the 5-number summary in all boxplots.

20) The test scores of 40 students are listed below. Construct a boxplot for the data set.
25 35 43 44 47 48 54 55 56 57
59 62 63 65 66 68 69 71 72
72 73 74 76 77 78 79 80 81
81 82 83 85 89 92 93 94 97 98

A) 

B) 

C) 

D) 

53
Answer Key
Testname: CHAPTER 3 FORM A

1) The general answer is that the mean is sometimes in a data set. For example, the mean of the data set 1, 2, 3, 4, 5 is 3.
2) The bottom data set would have the largest standard deviation, since its values as a group are farther from 30 than those in the upper data sets. The top data set has the smallest standard deviation, namely 0, since all values are the same, and thereby equal the mean.
3) Marla’s score of 85% on her statistics exam tells us that Marla knew 85% of the content on that exam. Marla’s percentile score of 85 tells us that her score was better than 85% of the scores of examinees on that test.
4) B
5) C
6) C
7) D
8) B
9) A
10) In general, the mean salary of all employees of the company cannot be obtained by finding the mean of $29,525 and $33,470 because each of these means typically is obtained by averaging a different number of salaries for male and female employees. The mean of $29,525 and $33,470 will yield the mean salary of all employees of the company only if the number of female employees is equal to the number of male employees.
11) D
12) A
13) D
14) A
15) D
16) A
17) B
18) A
19) A
20) C
CHAPTER 3 FORM B

Name:_________________________ Course Number:_______ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Explain how two data sets could have equal means and modes but still differ greatly. Give an example with two data sets to illustrate.

2) We want to compare two different groups of students, students taking Composition 1 in a traditional lecture format and students taking Composition 1 in a distance learning format. We know that the mean score on the research paper is 85 for both groups. What additional information would be provided by knowing the standard deviation?

3) Describe how to find the percentile for a given score in a set of data. How does this process relate to the definition of a percentile score?

Find the mean for the given sample data.

4) Bill kept track of the number of hours he spent exercising each week. The results for four months are shown below. Find the mean number of hours Bill spent exercising per week. Round your answer to two decimal places.

<table>
<thead>
<tr>
<th>Hours</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7.50</td>
<td>8.20</td>
<td>7.10</td>
<td>7.90</td>
<td>8.00</td>
<td>7.50</td>
<td></td>
</tr>
<tr>
<td>7.80</td>
<td>7.10</td>
<td>7.30</td>
<td>7.50</td>
<td>7.90</td>
<td>8.90</td>
<td></td>
</tr>
<tr>
<td>7.10</td>
<td>8.20</td>
<td>8.20</td>
<td>8.00</td>
<td>7.80</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A) 7.38      B) 8.01      C) 7.79      D) 8.25
CHAPTER 3 FORM B

Find the median for the given sample data.

5) A store manager kept track of the number of newspapers sold each week over a seven-week period. The results are shown below.
   95, 38, 221, 122, 258, 237, 233
   Find the median number of newspapers sold.
   A) 122 newspapers       B) 172 newspapers
   C) 233 newspapers       D) 221 newspapers

Find the mode(s) for the given sample data.

6) 98, 25, 98, 13, 25, 29, 56, 98
   A) 25                B) 55.3
   C) 98                D) 42.5

Find the midrange for the given sample data.

7) The weights (in ounces) of 18 cookies are shown. Find the midrange.
   0.60 1.34 0.89 0.96 0.80 1.43
   1.34 1.16 0.60 1.49 1.34 1.16
   1.34 1.49 0.80 1.34 0.96 0.89
   A) 1.045             B) 1.015
   C) 1.145             D) 1.16

Find the mean of the data summarized in the given frequency distribution.

8) The highway speeds of 100 cars are summarized in the frequency distribution below. Find the mean speed.

<table>
<thead>
<tr>
<th>Speed (mph)</th>
<th>Cars</th>
</tr>
</thead>
<tbody>
<tr>
<td>30–39</td>
<td>3</td>
</tr>
<tr>
<td>40–49</td>
<td>18</td>
</tr>
<tr>
<td>50–59</td>
<td>52</td>
</tr>
<tr>
<td>60–69</td>
<td>17</td>
</tr>
<tr>
<td>70–79</td>
<td>10</td>
</tr>
</tbody>
</table>

   A) 58.6 mph        B) 61.4 mph        C) 54.5 mph        D) 55.8 mph

Solve the problem.

9) Melissa gets quiz grades of 75, 85, and 95. She gets an 80 on her final exam. Find the weighted mean if the quizzes each count for 15% and the final exam counts for 55% of the final grade.

   A) 82.1           B) 82.2
   C) 82.3           D) 82.4
CHAPTER 3 FORM B

10) The data below consists of the heights (in inches) of 20 randomly selected women. Find the 10% trimmed mean of the data set. The 10% trimmed mean is found by arranging the data in order, deleting the bottom 10% of the values and the top 10% of the values and then calculating the mean of the remaining values.

\[
\begin{array}{cccccc}
64 & 62 & 62 & 61 & 67 \\
68 & 70 & 67 & 62 & 63 \\
61 & 64 & 75 & 67 & 60 \\
59 & 64 & 68 & 65 & 71 \\
\end{array}
\]

A) 65.0 in  B) 64.8 in  C) 64.7 in  D) 51.8 in

Find the range for the given data.

11) Fred, a local mechanic, gathered the following data regarding the price, in dollars, of an oil and filter change at twelve competing service stations:

\[
\begin{array}{cccc}
32.95 & 24.95 & 26.95 & 28.95 \\
18.95 & 28.95 & 30.95 & 22.95 \\
24.95 & 26.95 & 29.95 & 28.95 \\
\end{array}
\]

Compute the range.

A) $14  B) $12  C) $8  D) $10

Find the variance for the given data. Round your answer to one more decimal place than the original data.

12) 1, 4, -5, -9, and 6

A) 31.4  B) 39.2  C) 39.3  D) 39.4

Solve the problem.

13) The maximum value of a distribution is 36.3 and the minimum value is 3.3. Use the range rule of thumb to find the standard deviation. Round results to the nearest tenth.

A) 14.9  B) 13.3  C) 5.3  D) 8.3

14) The ages of the members of a gym have a mean of 39 years and a standard deviation of 13. Use the range rule of thumb to estimate the minimum and maximum “usual” ages.

A) 13 years, 65 years  B) 12 years, 66 years

C) 15 years, 67 years  D) 26 years, 52 years

Use the empirical rule to solve the problem.

15) At one college, GPA’s are normally distributed with a mean of 2.7 and a standard deviation of 0.4. What percentage of students at the college have a GPA between 2.3 and 3.1?

A) 84.13%  B) 68.26%  C) 99.74%  D) 95.44%
CHAPTER 3 FORM B

Find the z-score corresponding to the given value and use the z-score to determine whether the value is unusual. Consider a score to be unusual if its z-score is less than -2.00 or greater than 2.00. Round the z-score to the nearest tenth if necessary.

16) A body temperature of 99.7° F given that human body temperatures have a mean of 98.20° F and a standard deviation of 0.62°.
   A) -2.4; unusual  B) 2.4; not unusual
   C) 2.4; unusual  D) 1.5; not unusual

Determine which score corresponds to the higher relative position.

17) Which is better: a score of 82 on a test with a 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?
   A) The second 82
   B) Both scores have the same relative position.
   C) The first 82

Find the percentile for the data point.

18) Data set: 3 11 8 6 3 3 11 6 3 11 2 11 15 4 9 3 12 8 6 11;
    data point 6
    A) 25  B) 62  C) 35  D) 40

Provide an appropriate response.

19) If all the values in a data set are converted to z-scores, the shape of the distribution of the z-scores will be bell-shaped regardless of the distribution of the original data. True or false?
    A) False  B) True
Construct a boxplot for the given data. Include values of the 5-number summary in all boxplots.

20) The test scores of 32 students are listed below. Construct a boxplot for the data set.

32 37 41 44 46 48 53 55
57 57 59 63 65 66 68 69
70 71 74 74 75 77 78 79
81 82 83 86 89 92 95 99

A)  

B)  

C)  

D)  

59
Answer Key

Testname: CHAPTER 3 FORM B

1) The sets would have different sizes and standard deviations. Examples will vary. A general method for constructing examples is as follows: (1) compose several values and find the sum, (2) prepare another data set with, say, twice that sum and twice that size. Examples are Set A; 5, 10, 15, 2, 5, 8; Set B: 9, 10, 11, 3, 5, 7, 19, 1, 2, 5, 6, 12. Notice that set A and Set B both have means = 7.5 and modes = 5. Yet, the n for A is 6, while the n for B is 12. Also, the s for A is 4.6, while the s for B is 5.1.

2) By knowing the standard deviation for both groups, we would have an idea about how the individual scores for each group varied about 85. The smaller standard deviation would indicate that individual scores were closer to 85 than would a larger standard deviation.

3) After the data values are arranged from lowest to highest, the number of values lower than the given score should be divided by the total number of values, and this quotient should be multiplied by 100. The percentile for the given score is the nearest whole number to the resulting product. Since a percentile score is a score at a specified percentile, it can be found by the method described above.

4) C
5) D
6) C
7) A
8) D
9) C
10) C
11) A
12) C
13) D
14) A
15) B
16) C
17) A
18) C
19) A
20) D
CHAPTER 3 FORM C

Name:_________________________ Course Number:_______ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided.
Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) The two most frequently used measures of central tendency are the mean and the median. Compare these two measures for the following characteristics: Takes every score into account? Affected by extreme scores? Advantages.

2) A company advertises an average of 42,000 miles for one of its new tires. In the manufacturing process there is some variation around that average. Would the company want a process that provides a large or a small variance? Justify your answer.

3) The textbook defines unusual values as those data points with z scores less than \( z = -2.00 \) or z scores greater than \( z = 2.00 \). Comment on this definition with respect to “Chebyshev’s Theorem”; refer specifically to the percent of scores which would be defined as unusual according to “Chebyshev’s Theorem”.

Find the mean for the given sample data.

4) Frank’s Furniture employees earned the following amounts last week:

\[
\begin{align*}
$375.51 & \quad $339.04 & \quad $180.97 & \quad $421.73 & \quad $516.56 \\
$370.13 & \quad $466.92 & \quad $165.25 & \quad $457.10
\end{align*}
\]

What was the mean amount earned by an employee last week? Round your answer to the nearest cent.

A) $365.91 \quad \quad \quad B) $470.46 \quad \quad \quad C) $359.25 \quad \quad \quad D) $411.65
CHAPTER 3 FORM C

Find the median for the given sample data.

5) The number of vehicles passing through a bank drive-up line during each 15-minute period was recorded. The results are shown below. Find the median number of vehicles going through the line in a fifteen-minute period.
28 30 28 31
31 28 33 30
38 34 34 32
27 34 28 23
18 30 30 30

A) 31  B) 29.85  C) 34  D) 30

Find the mode(s) for the given sample data.

6) The speeds (in mi/h) of the cars passing a certain checkpoint are measured by radar. The results are shown below. Find the mode(s). Speeds:
40.4 42.1 42.3 43.2 44.1
44.1 42.1 40.9 43.4 43.2
40.4 40.9 43.2 42.4 42.1
42.2 42.2 42.3 40.7 40.4

A) 42.1  B) 41.90
C) 40.4, 43.2, 42.1  D) 40.4

Find the midrange for the given sample data.

7) The speeds (in mph) of the cars passing a certain checkpoint are measured by radar. The results are shown below. Find the midrange.
44.4 41.7 43.0 40.7 43.0
40.3 44.8 42.0 44.4 42.8
43.1 42.0 40.7 43.1 41.7

A) 42.8  B) 42.55  C) 42.35  D) 4.50

Find the mean of the data summarized in the given frequency distribution.

8) A company had 80 employees whose salaries are summarized in the frequency distribution below. Find the mean salary.

<table>
<thead>
<tr>
<th>Salary ($)</th>
<th>Employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,001–10,000</td>
<td>16</td>
</tr>
<tr>
<td>10,001–15,000</td>
<td>11</td>
</tr>
<tr>
<td>15,001–20,000</td>
<td>16</td>
</tr>
<tr>
<td>20,001–25,000</td>
<td>10</td>
</tr>
<tr>
<td>25,001–30,000</td>
<td>27</td>
</tr>
</tbody>
</table>

A) $16,931.70  B) $20,694.30  C) $18,813.00  D) $17,500
CHAPTER 3 FORM C

Solve the problem.

9) Carl gets quiz grades of 70, 75, 82, and 85. He gets a 94 on his final exam. Find the weighted mean if the quizzes each count for 10% and the final exam counts for 60% of the final grade.
   A) 85.3   B) 86.7   C) 87.6   D) 88.0

10) The 10% trimmed mean of a data set is found by arranging the data in order, deleting the bottom 10% of the values and the top 10% of the values and then calculating the mean of the remaining values. What advantages do you think that the trimmed mean has as compared to the mean?

Find the range for the given data.

11) A class of sixth grade students kept accurate records on the amount of time they spent playing video games during a one-week period. The times (in hours) are listed below:
   23.0  15.5  9.6  24.4  15.2
   30.5  24.1  16.6  15.7  14.0
   Compute the range.
   A) 20.9   B) 9.6   C) 15.2   D) 7.5

Find the variance for the given data. Round your answer to one more decimal place than the original data.

12) 4, 11, 11, 2, and 8
   A) 16.7   B) 27.5   C) 13.4   D) 16.6

Solve the problem.

13) The race speeds for the top eight cars in a 200-mile race are listed below. Use the range rule of thumb to find the standard deviation. Round results to the nearest tenth.
   189.1  185.9  189.2  182.4  175.6  184.2  188.3  177.2
   A) 1.1   B) 3.4   C) 7.5   D) 6.8

14) The maximum value of a distribution is 36.3 and the minimum value is 3.3. Use the range rule of thumb to find the standard deviation. Round results to the nearest tenth.
   A) 8.3   B) 14.9   C) 5.3   D) 13.3
CHAPTER 3 FORM C

Use the empirical rule to solve the problem.

15) The systolic blood pressure of 18-year-old women is normally distributed with a mean of 120 mmHg and a standard deviation of 12 mmHg. What percentage of 18-year-old women have a systolic blood pressure between 96 mmHg and 144 mmHg?
   A) 68.26%  B) 99.99%  C) 95.44%  D) 99.74%

Find the z-score corresponding to the given value and use the z-score to determine whether the value is unusual. Consider a score to be unusual if its z-score is less than –2.00 or greater than 2.00. Round the z-score to the nearest tenth if necessary.

16) A weight of 225 pounds among a population having a mean weight of 161 pounds and a standard deviation of 23.0 pounds.
   A) –2.8; not unusual  B) 64.4; unusual
   C) 2.8; unusual  D) 2.8; not unusual

Determine which score corresponds to the higher relative position.

17) Which score has the higher relative position: a score of 55 on a test for which \(\bar{x} = 43\) and \(s = 10\), a score of 5.0 on a test for which \(\bar{x} = 4\) and \(s = 0.8\) or a score of 435.6 on a test for which \(\bar{x} = 396\) and \(s = 44\)?
   A) A score of 435.6  B) A score of 55  C) A score of 5.0

Find the percentile for the data point.

18) Data set: 10 15 35 20 10 25 50 45 55 15 15 50 30 5 50; data point 35
   A) 35  B) 70  C) 52  D) 60

Provide an appropriate response.

19) For data which are heavily skewed to the right, \(P_{10}\) is likely to be closer to the median than \(P_{90}\). True or false?
   A) True  B) False
Construct a boxplot for the given data. Include values of the 5-number summary in all boxplots.

20) The ages of the 35 members of a track and field team are listed below. Construct a boxplot for the data set.

15 16 18 18 18 19 20
20 20 21 21 22 22 23
23 24 24 24 25 25 26
27 27 28 29 29 30 31
31 33 34 35 39 42 48

A)

B)

C)

D)
1) The mean takes every score of a data set into account. The median only takes into account the middle score of a ranked odd-numbered data set or the two middle scores of a ranked even-numbered data set. The mean is sensitive to extremes and can be drawn toward very low or very high values, but the median is not affected by extremes. Since it uses all the values of a data set, the mean is the preferred average, unless there are extreme values. In the latter case, the median is preferred. An example of the latter is the comparison of salaries of occupations.

2) A small variance is preferred, since this measure denotes consistency in the lifetime of the tires. Given small variation, buyers would get useful mileage from those tires around 42,000. Large variation would indicate that some buyers could have their tires wear out many miles short of 42,000, whereas others might get good use out of many miles past 42,000.

3) Chebychev’s Theorem states that in any data set, at least 75% of the values lie within 2 standard deviations from the mean of the set. Therefore, at most 25% of the values in any data set would be considered unusual by that theorem.

4) A
5) D
6) C
7) B
8) C
9) C

10) The mean is very sensitive to extreme values. When calculating the trimmed mean, the extreme values are omitted. The trimmed mean is therefore less sensitive to extreme values and therefore more resistant.

11) A
12) A
13) B
14) A
15) C
16) C
17) C
18) D
19) A
20) C
CHAPTER 4 FORM A

Name: ______________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Probabilities are useful in the decision-making process. Suppose a random sample of 152 students was surveyed regarding an instructor's teaching. Suppose 105 students rated the instructor either excellent or above average on lecture presentations, 96 students rated the instructor as giving difficult or very difficult assignments and tests. Would you take this instructor for a class? Discuss the influence of probabilities on making this decision.

2) Give an example of events which are independent but not mutually exclusive.

3) Suppose a student is taking a 5-response multiple choice exam; that is, the choices are A, B, C, D, and E, with only one of the responses correct. Describe the complement method for determining the probability of getting at least one of the questions correct on the 15-question exam. Why would the complement method be the method of choice for this problem?

Express the indicated degree of likelihood as a probability value.

4) "It will definitely turn dark tonight."
   A) 1  B) 0.30  C) 0.67  D) 0.5

67
CHAPTER 4 FORM A

Answer the question.

5) Which of the following cannot be a probability?
   A) $\frac{1}{2}$  B) 0  C) -1  D) 1

Find the indicated probability.

6) A class consists of 13 women and 49 men. If a student is randomly selected, what is the probability that the student is a woman?
   A) $\frac{49}{62}$  B) $\frac{1}{62}$  C) $\frac{13}{62}$  D) $\frac{13}{49}$

Answer the question, considering an event to be "unusual" if its probability is less than or equal to 0.05.

7) Assume that a study of 500 randomly selected school bus routes showed that 483 arrived on time. Is it "unusual" for a school bus to arrive late?
   A) Yes  B) No

Estimate the probability of the event.

8) A polling firm, hired to estimate the likelihood of the passage of an up-coming referendum, obtained the set of survey responses to make its estimate. The encoding system for the data is: 0 = FOR, 1 = AGAINST. If the referendum were held today, estimate the probability that it would pass.

   0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0

   A) 0.65  B) 0.5  C) 0.6  D) 0.4

Answer the question.

9) In a certain town, 25% of people commute to work by bicycle. If a person is selected randomly from the town, what are the odds against selecting someone who commutes by bicycle?
   A) 3 : 4  B) 1 : 3  C) 3 : 1  D) 1 : 4

Find the indicated probability.

10) Based on meteorological records, the probability that it will snow in a certain town on January 1st is 0.230. Find the probability that in a given year it will not snow on January 1st in that town.
    A) 4.348  B) 1.230  C) 0.770  D) 0.299

11) Of the 57 people who answered "yes" to a question, 8 were male. Of the 61 people that answered "no" to the question, 5 were male. If one person is selected at random from the group, what is the probability that the person answered "yes" or was male?
    A) 0.14  B) 0.525  C) 0.11  D) 0.593
12) The table below describes the smoking habits of a group of asthma sufferers.

<table>
<thead>
<tr>
<th></th>
<th>Nonsmoker</th>
<th>Occasional smoker</th>
<th>Regular smoker</th>
<th>Heavy smoker</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>339</td>
<td>33</td>
<td>61</td>
<td>34</td>
<td>467</td>
</tr>
<tr>
<td>Women</td>
<td>377</td>
<td>32</td>
<td>84</td>
<td>36</td>
<td>529</td>
</tr>
<tr>
<td>Total</td>
<td>716</td>
<td>65</td>
<td>145</td>
<td>70</td>
<td>996</td>
</tr>
</tbody>
</table>

If one of the 996 people is randomly selected, find the probability of getting a regular or heavy smoker.

A) 0.442      B) 0.216      C) 0.095      D) 0.146

13) In one town, 76% of adults have health insurance. What is the probability that 8 adults selected at random from the town all have health insurance?

A) 0.111      B) 0.76       C) 6.08       D) 0.105

14) A IRS auditor randomly selects 3 tax returns from 58 returns of which 8 contain errors. What is the probability that she selects none of those containing errors?

A) 0.0026     B) 0.0018     C) 0.6352     D) 0.6407

15) A sample of 4 different calculators is randomly selected from a group containing 17 that are defective and 36 that have no defects. What is the probability that at least one of the calculators is defective?

A) 0.787      B) 0.170      C) 0.799      D) 0.201

16) The table below describes the smoking habits of a group of asthma sufferers.

<table>
<thead>
<tr>
<th></th>
<th>Nonsmoker</th>
<th>Light smoker</th>
<th>Heavy smoker</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>313</td>
<td>84</td>
<td>69</td>
<td>466</td>
</tr>
<tr>
<td>Women</td>
<td>319</td>
<td>78</td>
<td>67</td>
<td>464</td>
</tr>
<tr>
<td>Total</td>
<td>632</td>
<td>162</td>
<td>136</td>
<td>930</td>
</tr>
</tbody>
</table>

If one of the 930 subjects is randomly selected, find the probability that the person chosen is a nonsmoker given that it is a woman. Round to the nearest thousandth.

A) 0.505      B) 0.399      C) 0.688      D) 0.343
CHAPTER 4 FORM A

Solve the problem.

17) Swinging Sammy Skor's batting prowess was simulated to get an estimate of the probability that Sammy will get a hit. Let 1 = HIT and 2 = OUT. The output from the simulation was as follows.

```
1 2 2 2 1 2 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2 2 1 2 1 2 2 2 2 2 1 1
```

Estimate the probability that he makes an out.
A) 0.782        B) 0.667        C) 0.621        D) 0.810

18) The library is to be given 8 books as a gift. The books will be selected from a list of 20 titles. If each book selected must have a different title, how many possible selections are there?

A) 2.432902008e+18        B) 5.0791104e+09
C) 6.033983155e+13        D) 125,970

19) A musician plans to perform 4 selections. In how many ways can she arrange the musical selections?

A) 4        B) 16        C) 120        D) 24

20) A tourist in France wants to visit 10 different cities. If the route is randomly selected, what is the probability that she will visit the cities in alphabetical order?

A) 3,628,800        B) \( \frac{1}{100} \)        C) \( \frac{1}{10} \)        D) \( \frac{1}{3,628,800} \)
Answer Key
Testname: CHAPTER 4 FORM A

1) Answers will vary. Ratios could be considered as relative frequency probabilities, such as $P(I$ take this instructor). Answer could address $P($good lecturer$) = 0.691; P($difficult assessments$) = 0.632$.

2) Answers will vary, but might include something like eating steak for supper and losing your car keys.

3) $P($at least one correct$) = 1 - P($none are correct$)$. The alternative to the complement method is to find $P(1), P(2), ... P(15)$ and take this sum. This method is too time consuming and too difficult.

4) A
5) C
6) C
7) A
8) C
9) C
10) C
11) B
12) B
13) A
14) C
15) C
16) C
17) B
18) D
19) D
20) D
CHAPTER 4 FORM B

Name: __________________________ Course Number: _______ Section Number: _______

Directions: Write your answers to the short-answer items in the spaces provided.
Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Describe an event whose probability of occurring is 1 and explain what that probability means. Describe an event whose probability of occurring is 0 and explain what that probability means.

2) Under what circumstances can you sample without replacement and still use the multiplication rule for independence? Discuss population and sample size as you answer this question.

3) Consider the following formulas: \( nP_r = \frac{n!}{(n-r)!} \) and \( nC_r = \frac{n!}{(n-r)!r!} \).

Given the same values for \( n \) and \( r \) in each formula, which is the smaller value, \( P \) or \( C \)? How does this relate to the concept of counting the number of outcomes based on whether or not order is a criterion?

Express the indicated degree of likelihood as a probability value.

4) "You have a 50-50 chance of choosing the correct answer."
   A) 0.9          B) 0.50          C) 0.25          D) 50

Answer the question.

5) Which of the following cannot be a probability?
   A) \( \frac{5}{3} \)          B) \( \frac{1}{2} \)          C) \( \frac{2}{3} \)          D) \( \frac{3}{5} \)
CHAPTER 4 FORM B

Find the indicated probability.

6) In a poll, respondents were asked whether they had ever been in a car accident. 471 respondents indicated that they had been in a car accident and 420 respondents said that they had not been in a car accident. If one of these respondents is randomly selected, what is the probability of getting someone who has been in a car accident? Round to the nearest thousandth, if necessary.

A) 0.471  
B) 0.002  
C) 0.529  
D) 1.121

Answer the question, considering an event to be "unusual" if its probability is less than or equal to 0.05.

7) Is it "unusual" to get a 12 when a pair of dice is rolled?

A) No

B) Yes

Estimate the probability of the event.

8) Of 1275 people who came into a blood bank to give blood, 278 people had high blood pressure. Estimate the probability that the next person who comes in to give blood will have high blood pressure.

A) 0.269  
B) 0.218  
C) 0.186  
D) 0.137

Answer the question.

9) Find the odds against correctly guessing the answer to a multiple choice question with 6 possible answers.

A) 5 : 1  
B) 5 : 6  
C) 6 : 1  
D) 6 : 5

Find the indicated probability.

10) The probability that Luis will pass his statistics test is 0.65. Find the probability that he will fail his statistics test.

A) 0.35  
B) 1.54  
C) 1.86  
D) 0.33

11) A study of consumer smoking habits includes 161 people in the 18–22 age bracket (50 of whom smoke), 144 people in the 23–30 age bracket (35 of whom smoke), and 80 people in the 31–40 age bracket (28 of whom smoke). If one person is randomly selected from this sample, find the probability of getting someone who is age 23–30 or smokes.

A) 0.091  
B) 0.577  
C) 0.668  
D) 0.243
12) 100 employees of a company are asked how they get to work and whether they work full time or part time. The figure below shows the results. If one of the 100 employees is randomly selected, find the probability that the person drives alone or cycles to work.

![Diagram](image)

1. Public transportation: 8 full time, 9 part time
2. Bicycle: 3 full time, 3 part time
3. Drive alone: 28 full time, 31 part time
4. Carpool: 9 full time, 9 part time

A) 0.59  B) 0.31  C) 0.36  D) 0.65

13) A manufacturing process has a 70% yield, meaning that 70% of the products are acceptable and 30% are defective. If three of the products are randomly selected, find the probability that all of them are acceptable.

A) 0.027  B) 2.1  C) 0.343  D) 0.429

14) A sample of 4 different calculators is randomly selected from a group containing 47 that are defective and 27 that have no defects. What is the probability that all four of the calculators selected are defective?

A) 0.1089  B) 0.1550  C) 10.1632  D) 0.1627

15) In a batch of 8,000 clock radios 6% are defective. A sample of 5 clock radios is randomly selected without replacement from the 8,000 and tested. The entire batch will be rejected if at least one of those tested is defective. What is the probability that the entire batch will be rejected?

A) 0.200  B) 0.266  C) 0.0600  D) 0.734
CHAPTER 4 FORM B

16) The table below describes the smoking habits of a group of asthma sufferers.

<table>
<thead>
<tr>
<th></th>
<th>Light Nonsmoker</th>
<th>Heavy smoker</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>339</td>
<td>76</td>
<td>66</td>
</tr>
<tr>
<td>Women</td>
<td>327</td>
<td>89</td>
<td>65</td>
</tr>
<tr>
<td>Total</td>
<td>666</td>
<td>165</td>
<td>131</td>
</tr>
</tbody>
</table>

If one of the 962 subjects is randomly selected, find the probability that the person chosen is a woman given that the person is a light smoker.

A) 0.272  B) 0.539  C) 0.185  D) 0.093

Solve the problem.

17) Swinging Sammy Skor's batting prowess was simulated to get an estimate of the probability that Sammy will get a hit. Let 1 = HIT and 0 = OUT. The output from the simulation was as follows.

1 0 0 0 1 0 0 1 0 0 1 1 0 0 0 0 1 1 1 1 0 0 1 1 0 0 1 1 1 1 1 0 0 0 0 0 1 1 1 1

Estimate the probability that he gets a hit.

A) 0.286  B) 0.301  C) 0.476  D) 0.452

18) How many ways can an IRS auditor select 4 of 13 tax returns for an audit?

A) 17,160  B) 24  C) 715  D) 28,561

19) A tourist in France wants to visit 8 different cities. How many different routes are possible?

A) 64  B) 5040  C) 8  D) 40,320

20) In a certain lottery, five different numbers between 1 and 36 inclusive are drawn. These are the winning numbers. To win the lottery, a person must select the correct 5 numbers in the same order in which they were drawn. What is the probability of winning?

A) \( \frac{1}{120} \)  B) \( \frac{1}{45,239,040} \)  C) \( \frac{1}{36!} \)  D) \( \frac{120}{45,239,040} \)
Answer Key
Testname: CHAPTER 4 FORM B

1) Answers will vary. Answers should include: Probability = 1, anything that is a sure thing, must occur. Probability = 0, anything that will never occur.
2) The sample size must be no more than 5% of the population size. So, taking something very small out of something very large will effectively result in essentially the same probability value for sampling without replacement as that for sampling with replacement, which is an easier formula to use.
3) The combination value will be smaller, since order is not important. For example, ABC is equivalent to ACB and would not be counted twice. If, however, r is 0 or 1 then nPr = nCr.
4) B
5) A
6) C
7) B
8) B
9) A
10) A
11) B
12) D
13) C
14) B
15) B
16) B
17) C
18) C
19) D
20) B
CHAPTER 4 FORM C

Name: __________________________ Course Number: ______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) When asked about the probability that he would become a sumo wrestler, Sam replied "slim to none." Relate that phrase to a numeric probability and interpret his meaning.

2) If a game were “fair,” the payoff on a bet would be the same as the odds for the event. In one game, the odds for winning are 1:13. If the game were "fair," what would the payoff be for a $5 bet? Of course, games in casinos are designed to make a profit for the casino investors. Supposing the casino makes the payoff at 1:11 odds, what profit does the casino make on your winning bet?

3) Suppose that a class of 30 students is assigned to write an essay.
   1) Suppose 4 essays are randomly chosen to appear on the class bulletin board. How many different groups of 4 are possible?
   2) Suppose 4 essays are randomly chosen for awards of $10, $7, $5, and $3. How many different groups of 4 are possible?
   Explain the significant differences between problems 1 and 2.

Express the indicated degree of likelihood as a probability value.

4) "There is a 40% chance of rain tomorrow."
   A) 40  B) 0.40  C) 4  D) 0.60
CHAPTER 4 FORM C

Answer the question.

5) On a multiple choice test with four possible answers for each question, what is the probability of answering a question correctly if you make a random guess?
   A) 1  B) \( \frac{3}{4} \)  C) \( \frac{1}{4} \)  D) \( \frac{1}{2} \)

Find the indicated probability.

6) Two 6--sided dice are rolled. What is the probability that the sum of the two numbers on the dice will be 3?
   A) 2  B) \( \frac{17}{18} \)  C) \( \frac{1}{18} \)  D) \( \frac{1}{2} \)

Answer the question, considering an event to be "unusual" if its probability is less than or equal to 0.05.

7) Assume that one student in your class of 28 students is randomly selected to win a prize. Would it be "unusual" for you to win?
   A) Yes  B) No

Estimate the probability of the event.

8) The data set represents the income levels of the members of a country club. Estimate the probability that a randomly selected member earns at least $83,000. Round your answers to the nearest tenth.

89,000 95,000 75,000 98,000 79,000 89,000 83,000 67,000 104,000 119,000 71,000 86,000
101,000 79,000 95,000 92,000 83,000 107,000 63,000 92,000
   A) 0.4  B) 0.8  C) 0.6  D) 0.7

Answer the question.

9) Suppose you are playing a game of chance. If you bet $5 on a certain event, you will collect $200 (including your $5 bet) if you win. Find the odds used for determining the payoff.
   A) 39 : 1  B) 1 : 39  C) 40 : 1  D) 200 : 205

Find the indicated probability.

10) If a person is randomly selected, find the probability that his or her birthday is not in May. Ignore leap years.
   A) \( \frac{334}{365} \)  B) \( \frac{11}{12} \)  C) \( \frac{31}{365} \)  D) \( \frac{31}{334} \)
CHAPTER 4 FORM C

11) The table below describes the smoking habits of a group of asthma sufferers.

<table>
<thead>
<tr>
<th></th>
<th>Nonsmoker</th>
<th>Occasional smoker</th>
<th>Regular smoker</th>
<th>Heavy smoker</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>384</td>
<td>33</td>
<td>64</td>
<td>49</td>
<td>530</td>
</tr>
<tr>
<td>Women</td>
<td>349</td>
<td>44</td>
<td>72</td>
<td>38</td>
<td>503</td>
</tr>
<tr>
<td>Total</td>
<td>733</td>
<td>77</td>
<td>136</td>
<td>87</td>
<td>1033</td>
</tr>
</tbody>
</table>

If one of the 1033 people is randomly selected, find the probability that the person is a man or a heavy smoker.

A) 0.502      B) 0.563      C) 0.550      D) 0.597

12) A bag contains 8 red marbles, 3 blue marbles, and 1 green marble. Find P(not blue).

A) $\frac{3}{4}$      B) $\frac{1}{4}$      C) 9      D) $\frac{4}{3}$

13) In a homicide case 7 different witnesses picked the same man from a line up. The line up contained 5 men. If the identifications were made by random guesses, find the probability that all 7 witnesses would pick the same person.

A) 0.0000128      B) 0.0000595      C) 0.000064      D) 1.4

14) Among the contestants in a competition are 49 women and 29 men. If 5 winners are randomly selected, what is the probability that they are all men?

A) 0.09784      B) 0.07261      C) 0.06228      D) 0.00563

15) A study conducted at a certain college shows that 54% of the school’s graduates find a job in their chosen field within a year after graduation. Find the probability that among 9 randomly selected graduates, at least one finds a job in his or her chosen field within a year of graduating.

A) 0.540      B) 0.111      C) 0.996      D) 0.999

16) The following table contains data from a study of two airlines which fly to Small Town, USA.

<table>
<thead>
<tr>
<th></th>
<th>Number of flights which were on time</th>
<th>Number of flights which were late</th>
</tr>
</thead>
<tbody>
<tr>
<td>Podunk Airlines</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>Upstate Airlines</td>
<td>43</td>
<td>5</td>
</tr>
</tbody>
</table>

If one of the 87 flights is randomly selected, find the probability that the flight selected arrived on time given that it was an Upstate Airlines flight.

A) $\frac{11}{76}$      B) $\frac{43}{48}$

C) $\frac{43}{87}$      D) None of the above is correct.
CHAPTER 4 FORM C

Solve the problem.

17) A firm uses trend projection and seasonal factors to simulate sales for a given time period. It assigns "0" if sales fall, "1" if sales are steady, "2" if sales rise moderately, and "3" if sales rise a lot. The simulator generates the following output.

0 1 0 2 2 0 0 1 2 0 2 0 2 2 1 2 0 1 2 2 0 3 0 2 1 2 1

Estimate the probability that sales will remain steady.
A) 0.194  
B) 0.412  
C) 0.258  
D) 0.125

18) There are 9 members on a board of directors. If they must form a subcommittee of 5 members, how many different subcommittees are possible?
A) 126  
B) 59,049  
C) 15,120  
D) 120

19) A pollster wants to minimize the effect the order of the questions has on a person's response to a survey. How many different surveys are required to cover all possible arrangements if there are 8 questions on the survey?
A) 5040  
B) 8  
C) 40,320  
D) 64

20) A class has 8 students who are to be assigned seating by lot. What is the probability that the students will be arranged in order from shortest to tallest? (Assume that no two students are the same height.)
A) 0.00019841  
B) 0.0000248  
C) 0.1000  
D) 0.00024802
Answer Key
Testname: CHAPTER 4 FORM C

1) Sam is essentially saying: $P(\text{sumo wrestler}) = 0$, will not ever occur.

2) In a "fair" game the payoff would be $\$70 = 13 \cdot 5 + 5$ original bet. The casino payoff would be $\$60 = 11 \cdot 5 + 5$ original bet. The profit would be $\$10$.

3) Problem 1 is a combination, not dependent on order, while problem 2 is a permutation and is dependent on order. 27,405 different groups of 4 are possible for problem 1; 657,720 different groups of 4 are possible for problem 2.

4) B
5) C
6) C
7) A
8) D
9) A
10) A
11) C
12) A
13) C
14) D
15) D
16) B
17) A
18) A
19) C
20) B
CHAPTER 5 FORM A

Name:___________________________ Course Number:_______ Section Number:_______

**Directions:** Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

**Provide an appropriate response.**

1) Suppose a mathematician computed the expected value of winnings for a person playing each of seven different games in a casino. What would you expect to be true for all expected values for these seven games?

2) Previously, you learned to find the three important characteristics of data: the measure of central tendency, the measure of variation, and the nature of the distribution. We can find the same three characteristics for a binomial distribution. Given a binomial distribution with \( p = 0.4 \) and \( n = 8 \), find the three characteristics.

**Identify the given random variable as being discrete or continuous.**

3) The pH level in a shampoo
   
   A) Discrete  
   B) Continuous

4) The number of field goals kicked in a football game

   A) Discrete  
   B) Continuous
CHAPTER 5 FORM A

Determine whether the following is a probability distribution. If not, identify the requirement that is not satisfied.

5) 

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.146</td>
</tr>
<tr>
<td>1</td>
<td>0.069</td>
</tr>
<tr>
<td>2</td>
<td>-0.042</td>
</tr>
<tr>
<td>3</td>
<td>0.072</td>
</tr>
<tr>
<td>4</td>
<td>0.154</td>
</tr>
<tr>
<td>5</td>
<td>0.601</td>
</tr>
</tbody>
</table>

Determine whether the following is a probability distribution. If not, identify the requirement that is not satisfied.

6) If a person is randomly selected from a certain town, the probability distribution for the number, x, of siblings is as described in the accompanying table.

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>1</td>
<td>0.24</td>
</tr>
<tr>
<td>2</td>
<td>0.24</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>0.07</td>
</tr>
<tr>
<td>5</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Find the mean of the given probability distribution.

7) The number of golf balls ordered by customers of a pro shop has the following probability distribution.

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(x)</td>
<td>0.14</td>
<td>0.36</td>
<td>0.36</td>
<td>0.04</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>

A) 9  B) 7.8  C) 5.13  D) 9.72
CHAPTER 5 FORM A

Solve the problem.

8) In a certain town, 70% of adults have a college degree. The accompanying table describes the probability distribution for the number of adults (among 4 randomly selected adults) who have a college degree. Find the variance for the probability distribution.

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0081</td>
</tr>
<tr>
<td>1</td>
<td>0.0756</td>
</tr>
<tr>
<td>2</td>
<td>0.2646</td>
</tr>
<tr>
<td>3</td>
<td>0.4116</td>
</tr>
<tr>
<td>4</td>
<td>0.2401</td>
</tr>
</tbody>
</table>

A) 0.72  B) 0.84  C) 0.92  D) 8.68

9) The prizes that can be won in a sweepstakes are listed below together with the chances of winning each one:
$5900 (1 \text{ chance in 8600}); $2700 (1 \text{ chance in 5000}); $600 (1 \text{ chance in 4800});$200 (1 \text{ chance in 2500}).
Find the expected value of the amount won for one entry if the cost of entering is 52 cents.

A) $0.91  B) $0.67  C) $1.35  D) $200

Assume that a researcher randomly selects 14 newborn babies and counts the number of girls selected, x. The probabilities corresponding to the 14 possible values of x are summarized in the given table. Answer the question using the table.

<table>
<thead>
<tr>
<th>Probabilities of Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>x(girls)</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

10) Find the probability of selecting 9 or more girls.

A) 0.061  B) 0.212  C) 0.001  D) 0.122
CHAPTER 5 FORM A

Answer the question.

11) Focus groups of 12 people are randomly selected to discuss products of the Yummy Company. It is determined that the mean number (per group) who recognize the Yummy brand name is 9.5, and the standard deviation is 0.69. Would it be unusual to randomly select 12 people and find that fewer than 6 recognize the Yummy brand name?

A) Yes  
B) No

Determine whether the given procedure results in a binomial distribution. If not, state the reason why.

12) Rolling a single die 34 times, keeping track of the numbers that are rolled.

A) Not binomial: the trials are not independent.  
B) Not binomial: there are too many trials.  
C) Procedure results in a binomial distribution.  
D) Not binomial: there are more than two outcomes for each trial.

Assume that a procedure yields a binomial distribution with a trial repeated n times. Use the binomial probability formula to find the probability of x successes given the probability p of success on a single trial.

13) n = 10, x = 2, p = \frac{1}{3}

A) 0.1951  
B) 0.1929  
C) 0.0028  
D) 0.2156

Find the indicated probability.

14) A machine has 7 identical components which function independently. The probability that a component will fail is 0.2. The machine will stop working if more than three components fail. Find the probability that the machine will be working.

A) 0.996  
B) 0.029  
C) 0.033  
D) 0.967

Find the standard deviation, \( \sigma \), for the binomial distribution which has the stated values of n and p. Round your answer to the nearest hundredth.

15) n = 22; p = .2

A) \( \sigma = 6.00 \)  
B) \( \sigma = 1.88 \)  
C) \( \sigma = 5.15 \)  
D) \( \sigma = -0.53 \)
CHAPTER 5 FORM A
Use the given values of $n$ and $p$ to find the minimum usual value $\mu - 2\sigma$ and the maximum usual value $\mu + 2\sigma$.

16) $n = 660, p = \frac{4}{7}$

A) Minimum: 351.72; maximum: 402.57
B) Minimum: 328.45; maximum: 425.83
C) Minimum: 402.57; maximum: 351.72
D) Minimum: 364.43; maximum: 389.86

Solve the problem.

17) The probability that a person has immunity to a particular disease is 0.1. Find the mean number who have immunity in samples of size 13.

A) 11.7  B) 0.1  C) 6.5  D) 1.3

18) The probability of winning a certain lottery is 1/59,127. For people who play 514 times, find the standard deviation for the number of wins.

A) 0.0087  B) 2.1138  C) 0.1021  D) 0.0932

Blank note

Determine if the outcome is unusual. Consider as unusual any result that differs from the mean by more than 2 standard deviations. That is, unusual values are either less than $\mu - 2\sigma$ or greater than $\mu + 2\sigma$.

19) According to AccuData Media Research, 36% of televisions within the Chicago city limits are tuned to "Eyewitness News" at 5:00 pm on Sunday nights. At 5:00 pm on a given Sunday, 2500 such televisions are randomly selected and checked to determine what is being watched. Would it be unusual to find that 925 of the 2500 televisions are tuned to "Eyewitness News"?

A) No  B) Yes

Use the Poisson Distribution to find the indicated probability.

20) A computer salesman averages 0.9 sales per week. Use the Poisson distribution to find the probability that in a randomly selected week the number of computers sold is 1.

A) 0.4025  B) 0.4574  C) 0.3659  D) 0.3293
Answer Key  
Testname: CHAPTER 5 FORM A

1) The expected values would be negative for the person playing, as the casinos are designed to make money.
2) Measure of central tendency: mean = 3.2, that is 0.4 \cdot 8.
   Measure of variation: SD = 1.386, that is \sqrt[8]{0.4 \cdot 0.6}
   Nature of the Distribution

3) B
4) A
5) Not a probability distribution. One of the P(x)'s is negative.
6) Not a probability distribution. The sum of the P(x)'s is not 1, since 0.98 \neq 1.00.
7) B
8) B
9) A
10) B
11) A
12) D
13) A
14) D
15) B
16) A
17) D
18) D
19) A
20) C
CHAPTER 5 FORM B

Name: __________________________ Course Number: _______ Section Number: _______

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) A game is said to be “fair” if the expected value for winnings is 0, that is, in the long run, the player can expect to win 0. Consider the following game. The game costs $1 to play and the winnings are $5 for red, $3 for blue, $2 for yellow, and nothing for white. The following probabilities apply. What are your expected winnings? Does the game favor the player or the owner?

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>.02</td>
</tr>
<tr>
<td>Blue</td>
<td>.04</td>
</tr>
<tr>
<td>Yellow</td>
<td>.16</td>
</tr>
<tr>
<td>White</td>
<td>.78</td>
</tr>
</tbody>
</table>

2) Describe the Poisson distribution and give an example of a random variable with a Poisson distribution.

Identify the given random variable as being discrete or continuous.

3) The cost of a randomly selected orange

A) Continuous          B) Discrete

4) The height of a randomly selected student

A) Continuous          B) Discrete
CHAPTER 5 FORM B

Determine whether the following is a probability distribution. If not, identify the requirement that is not satisfied.

5) 

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.08</td>
</tr>
<tr>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>0.06</td>
</tr>
<tr>
<td>4</td>
<td>0.55</td>
</tr>
</tbody>
</table>

6) A police department reports that the probabilities that 0, 1, 2, 3, and 4 car thefts will be reported in a given day are 0.150, 0.284, 0.270, 0.171, and 0.081, respectively.

Find the mean of the given probability distribution.

7) The accompanying table shows the probability distribution for x, the number that shows up when a loaded die is rolled.

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.15</td>
</tr>
<tr>
<td>2</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
</tr>
<tr>
<td>5</td>
<td>0.10</td>
</tr>
<tr>
<td>6</td>
<td>0.37</td>
</tr>
</tbody>
</table>

A) 3.98 B) 0.17 C) 3.85 D) 3.50

Solve the problem.

8) The random variable x is the number of houses sold by a realtor in a single month at the SENDSOM's Real Estate Office. Its probability distribution is as follows. Find the standard deviation for the probability distribution.

<table>
<thead>
<tr>
<th>Houses Sold (x)</th>
<th>Probability P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.24</td>
</tr>
<tr>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>0.14</td>
</tr>
<tr>
<td>6</td>
<td>0.11</td>
</tr>
<tr>
<td>7</td>
<td>0.21</td>
</tr>
</tbody>
</table>

A) 2.62 B) 6.86 C) 2.25 D) 4.45
CHAPTER 5 FORM B

9) A contractor is considering a sale that promises a profit of $31,000 with a probability of 0.7 or a loss (due to bad weather, strikes, and such) of $13,000 with a probability of 0.3. What is the expected profit?

A) $17,800  B) $18,000  C) $21,700  D) $30,800

Assume that a researcher randomly selects 14 newborn babies and counts the number of girls selected, x. The probabilities corresponding to the 14 possible values of x are summarized in the given table. Answer the question using the table.

<table>
<thead>
<tr>
<th>x(girls)</th>
<th>P(x)</th>
<th>x(girls)</th>
<th>P(x)</th>
<th>x(girls)</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000</td>
<td>5</td>
<td>0.122</td>
<td>10</td>
<td>0.061</td>
</tr>
<tr>
<td>1</td>
<td>0.001</td>
<td>6</td>
<td>0.183</td>
<td>11</td>
<td>0.022</td>
</tr>
<tr>
<td>2</td>
<td>0.006</td>
<td>7</td>
<td>0.209</td>
<td>12</td>
<td>0.006</td>
</tr>
<tr>
<td>3</td>
<td>0.022</td>
<td>8</td>
<td>0.183</td>
<td>13</td>
<td>0.001</td>
</tr>
<tr>
<td>4</td>
<td>0.061</td>
<td>9</td>
<td>0.122</td>
<td>14</td>
<td>0.000</td>
</tr>
</tbody>
</table>

10) Find the probability of selecting 2 or more girls.

A) 0.994  B) 0.006  C) 0.001  D) 0.999

Answer the question.

11) Focus groups of 14 people are randomly selected to discuss products of the Famous Company. It is determined that the mean number (per group) who recognize the Famous brand name is 8.9, and the standard deviation is 0.69. Would it be unusual to randomly select 14 people and find that greater than 12 recognize the Famous brand name?

A) Yes  B) No

Determine whether the given procedure results in a binomial distribution. If not, state the reason why.

12) Rolling a single "loaded" die 61 times, keeping track of the numbers that are rolled.

A) Procedure results in a binomial distribution.
B) Not binomial: the trials are not independent.
C) Not binomial: there are more than two outcomes for each trial.
D) Not binomial: there are too many trials.
CHAPTER 5 FORM B

Assume that a procedure yields a binomial distribution with a trial repeated n times. Use the binomial probability formula to find the probability of x successes given the probability p of success on a single trial.

13) \( n = 6, \ x = 3, \ p = \frac{1}{6} \)

A) 0.0286 B) 0.0536 C) 0.0154 D) 0.0322

Find the indicated probability.

14) A company purchases shipments of machine components and uses this acceptance sampling plan: Randomly select and test 24 components and accept the whole batch if there are fewer than 3 defectives. If a particular shipment of thousands of components actually has a 4% rate of defects, what is the probability that this whole shipment will be accepted?

A) 0.9307 B) 0.5553 C) 0.0550 D) 0.1799

Find the standard deviation, \( \sigma \), for the binomial distribution which has the stated values of n and p. Round your answer to the nearest hundredth.

15) \( n = 2815; \ p = .63 \)

A) \( \sigma = 28.89 \) B) \( \sigma = 25.62 \) C) \( \sigma = 23.21 \) D) \( \sigma = 29.74 \)

Use the given values of n and p to find the minimum usual value \( \mu - 2\sigma \) and the maximum usual value \( \mu + 2\sigma \).

16) \( n = 274, \ p = 0.273 \)

A) Minimum: 89.551; maximum: 60.053
B) Minimum: 64.373; maximum: 85.231
C) Minimum: 60.053; maximum: 89.551
D) Minimum: 67.428; maximum: 82.176

Solve the problem.

17) The probability of winning a certain lottery is 1/77,822. For people who play 961 times, find the mean number of wins.

A) 0.0010 B) 0.000013 C) 0.0123 D) 81.0

18) On a multiple choice test with 22 questions, each question has four possible answers, one of which is correct. For students who guess at all answers, find the standard deviation for the number of correct answers.

A) 1.99 B) 1.944 C) 1.984 D) 2.031

91
CHAPTER 5 FORM B

Determine if the outcome is unusual. Consider as unusual any result that differs from the mean by more than 2 standard deviations. That is, unusual values are either less than $\mu - 2\sigma$ or greater than $\mu + 2\sigma$.

19) The Acme Candy Company claims that 60% of the jawbreakers it produces weigh more than .4 ounces. Suppose that 800 jawbreakers are selected at random from the production lines. Would it be unusual for this sample of 800 to contain 470 jawbreakers that weigh more than .4 ounces?
   A) Yes  B) No

Use the Poisson Distribution to find the indicated probability.

20) The Columbia Power Company experiences power failures with a mean of $\mu = 0.210$ per day. Find the probability that there are exactly two power failures in a particular day.
   A) 0.036  B) 0.085  C) 0.027  D) 0.018
Answer Key
Testname: CHAPTER 5 FORM B

1) The expected winnings are -$0.46. the game is not fair, and it favors the owner of the game.
2) The Poisson distribution is a discrete probability distribution that applies to occurrences of some event over a specified interval. Examples will vary.
3) B
4) A
5) Probability distribution.
6) Not a probability distribution. The sum of the P(x)'s is not 1, since 0.956 ≠ 1.000.
7) A
8) A
9) A
10) D
11) A
12) C
13) B
14) A
15) B
16) C
17) C
18) D
19) B
20) D
CHAPTER 5 FORM C

Name:_________________________ Course Number:_______ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Compare the probability histogram for the expected sum with the actual results. What do you conclude about the dice results displayed in the Actual Sum of Two Dice histogram?

![Expected Sum of Two Dice](image1)

![Actual Sum of Two Dice](image2)

2) Describe the differences in the Poisson and the binomial distribution.

Identify the given random variable as being discrete or continuous.

3) The number of oil spills occurring off the Alaskan coast

   A) Discrete                     B) Continuous
CHAPTER 5 FORM C

4) The braking time of a car
   A) Discrete  B) Continuous

Determine whether the following is a probability distribution. If not, identify the requirement that is not satisfied.

5) 
   \[
   \begin{array}{|c|c|}
   \hline
   x & P(x) \\ 
   \hline
   0 & 0.187 \\
   1 & 0.137 \\
   2 & 0.145 \\
   3 & 0.181 \\
   4 & 0.235 \\
   5 & 0.203 \\
   \hline
   \end{array}
   \]

6) In a certain town, 20% of adults have a college degree. The accompanying table describes the probability distribution for the number of adults (among 4 randomly selected adults) who have a college degree.

   \[
   \begin{array}{|c|c|}
   \hline
   x & P(x) \\ 
   \hline
   0 & 0.4096 \\
   1 & 0.4096 \\
   2 & 0.1536 \\
   3 & 0.0256 \\
   4 & 0.0016 \\
   \hline
   \end{array}
   \]

Find the mean of the given probability distribution.

7) In a certain town, 60% of adults have a college degree. The accompanying table describes the probability distribution for the number of adults (among 4 randomly selected adults) who have a college degree.

   \[
   \begin{array}{|c|c|}
   \hline
   x & P(x) \\ 
   \hline
   0 & 0.0256 \\
   1 & 0.1536 \\
   2 & 0.3456 \\
   3 & 0.3456 \\
   4 & 0.1296 \\
   \hline
   \end{array}
   \]

A) 2.30  B) 2.00  C) 2.40  D) 2.43
CHAPTER 5 FORM C

Solve the problem.

8) In a certain town, 30% of adults have a college degree. The accompanying table describes the probability distribution for the number of adults (among 4 randomly selected adults) who have a college degree. Find the standard deviation for the probability distribution.

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2401</td>
</tr>
<tr>
<td>1</td>
<td>0.4116</td>
</tr>
<tr>
<td>2</td>
<td>0.2646</td>
</tr>
<tr>
<td>3</td>
<td>0.0756</td>
</tr>
<tr>
<td>4</td>
<td>0.0081</td>
</tr>
</tbody>
</table>

A) 0.92  B) 0.84  C) 1.06  D) 1.51

9) The prizes that can be won in a sweepstakes are listed below together with the chances of winning each one:

- $3800 (1 chance in 8300);
- $2600 (1 chance in 6900);
- $800 (1 chance in 3000);
- $300 (1 chance in 2300).

Find the expected value of the amount won for one entry if the cost to enter is 70 cents.

A) $0.40  B) $299.30  C) $0.53  D) $1.23

Assume that a researcher randomly selects 14 newborn babies and counts the number of girls selected, x. The probabilities corresponding to the 14 possible values of x are summarized in the given table. Answer the question using the table.

<table>
<thead>
<tr>
<th>Probabilities of Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>x(girls)</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

10) Find the probability of selecting 12 or more girls.

A) 0.001  B) 0.022  C) 0.007  D) 0.006
CHAPTER 5 FORM C

Answer the question.

11) Suppose that weight of adolescents is being studied by a health organization and that the accompanying tables describes the probability distribution for three randomly selected adolescents, where x is the number who are considered morbidly obese. Is it unusual to have no obese subjects among three randomly selected adolescents?

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.111</td>
</tr>
<tr>
<td>1</td>
<td>0.215</td>
</tr>
<tr>
<td>2</td>
<td>0.450</td>
</tr>
<tr>
<td>3</td>
<td>0.224</td>
</tr>
</tbody>
</table>

A) No  B) Yes

Determine whether the given procedure results in a binomial distribution. If not, state the reason why.

12) Rolling a single "loaded" die 39 times, keeping track of the "fives" rolled.

A) Not binomial: the trials are not independent.
B) Not binomial: there are more than two outcomes for each trial.
C) Not binomial: there are too many trials.
D) Procedure results in a binomial distribution.

Assume that a procedure yields a binomial distribution with a trial repeated n times. Use the binomial probability formula to find the probability of x successes given the probability p of success on a single trial.

13) n = 12, x = 5, p = 0.25

A) 0.027  B) 0.103  C) 0.082  D) 0.091

Find the indicated probability.

14) An airline estimates that 90% of people booked on their flights actually show up. If the airline books 71 people on a flight for which the maximum number is 69, what is the probability that the number of people who show up will exceed the capacity of the plane?

A) 0.0050  B) 0.0044  C) 0.0223  D) 0.0006

Find the standard deviation, σ, for the binomial distribution which has the stated values of n and p. Round your answer to the nearest hundredth.

15) n = 721; p = .7

A) σ = 16.42  B) σ = 12.30  C) σ = 15.57  D) σ = 9.89
CHAPTER 5 FORM C
Use the given values of n and p to find the minimum usual value \( \mu - 2\sigma \) and the maximum usual value \( \mu + 2\sigma \).

16) \( n = 2220, p = 0.595 \)
   A) Minimum: 1274.641; maximum: 1367.159
   B) Minimum: 1297.771; maximum: 1344.029
   C) Minimum: 1367.159; maximum: 1274.641
   D) Minimum: 1288.19; maximum: 1353.61

Solve the problem.

17) The probability is 0.4 that a person shopping at a certain store will spend less than $20. For groups of size 18, find the mean number who spend less than $20.
   A) 10.8          B) 8.0          C) 12.0          D) 7.2

18) According to a college survey, 22% of all students work full time. Find the standard deviation for the number of students who work full time in samples of size 16.
   A) 1.88          B) 1.66         C) 3.52          D) 2.75

Determine if the outcome is unusual. Consider as unusual any result that differs from the mean by more than 2 standard deviations. That is, unusual values are either less than \( \mu - 2\sigma \) or greater than \( \mu + 2\sigma \).

19) A survey for brand recognition is done and it is determined that 68% of consumers have heard of Dull Computer Company. A survey of 800 randomly selected consumers is to be conducted. For such groups of 800, would it be unusual get 554 consumers who recognize the Dull Computer Company name?
   A) Yes          B) No

Use the Poisson Distribution to find the indicated probability.

20) For a certain type of fabric, the average number of defects in each square foot of fabric is 0.4. Find the probability that a randomly selected square foot of the fabric will contain more than one defect.
   A) 0.0536          B) 0.0616         C) 0.9384         D) 0.7319
Answer Key  
Testname: CHAPTER 5 FORM C

1) Based on the expected and actual results, at least one of the two dice must not be fair.
2) The Poisson computes probabilities for occurrences of events over some interval.
   The Poisson distribution is affected only by the mean $\mu$, whereas the binomial is affected by
   sample size $n$ and probability $p$.
   The Poisson distribution has discrete values from 1, 2, 3, . . . with no upper limit. A binomial
   distribution has discrete values from 1, 2, 3, to $n$; that is, the upper limit of values is $n$.
3) A
4) B
5) Not a probability distribution. The sum of the $P(x)$'s is not 1, since 1.088 $\neq$ 1.000.
6) Probability distribution
7) C
8) A
9) C
10) C
11) A
12) D
13) B
14) A
15) B
16) A
17) D
18) B
19) B
20) B
CHAPTER 6 FORM A

Name:_____________________________ Course Number: ________ Section Number: ________

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) The typical computer random-number generator yields numbers in a uniform distribution between 0 and 1 with a mean of 0.500 and a standard deviation of 0.289. (a) Suppose a sample of size 50 is randomly generated. Find the probability that the mean is below 0.300. (b) Suppose a sample size of 15 is randomly generated. Find the probability that the mean is below 0.300. These two problems appear to be very similar. Only one can be solved by the Central Limit theorem. Which one and why?

Using the following uniform density curve, answer the question.

\[ P(x) \]

\[ 0.125 \]

\[ 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad x \]

2) What is the probability that the random variable has a value between 0.4 and 0.7?
   A) 0.288  B) 0.038  C) 0.163  D) 0.088

Assume that the weight loss for the first month of a diet program varies between 6 pounds and 12 pounds, and is spread evenly over the range of possibilities, so that there is a uniform distribution. Find the probability of the given range of pounds lost.

3) More than 10 pounds
   A) \( \frac{2}{3} \)  B) \( \frac{1}{7} \)  C) \( \frac{1}{3} \)  D) \( \frac{5}{6} \)

If Z is a standard normal variable, find the probability.

4) The probability that Z is greater than \(-1.82\)
   A) \(-0.0344\)  B) 0.4656  C) 0.0344  D) 0.9656

5) \( P(-0.73 < Z < 2.27) \)
   A) 1.54  B) 0.2211  C) 0.7557  D) 0.4884
CHAPTER 6 FORM A

The Precision Scientific Instrument Company manufactures thermometers that are supposed to give readings of 0°C at the freezing point of water. Tests on a large sample of these thermometers reveal that at the freezing point of water, some give readings below 0°C (denoted by negative numbers) and some give readings above 0°C (denoted by positive numbers). Assume that the mean reading is 0°C and the standard deviation of the readings is 1.00°C. Also assume that the frequency distribution of errors closely resembles the normal distribution. A thermometer is randomly selected and tested. Find the temperature reading corresponding to the given information.

6) If 7% of the thermometers are rejected because they have readings that are too low, but all other thermometers are acceptable, find the temperature that separates the rejected thermometers from the others.

   A) −1.26°  B) −1.48°  C) −1.53°  D) −1.39°

7) A quality control analyst wants to examine thermometers that give readings in the bottom 4%. Find the reading that separates the bottom 4% from the others.

   A) −1.75°  B) −1.48°  C) −1.63°  D) −1.89°

Assume that X has a normal distribution, and find the indicated probability.

8) The mean is \( \mu = 60.0 \) and the standard deviation is \( \sigma = 4.0 \). Find the probability that \( X \) is less than 53.0.

   A) 0.0401  B) 0.0802  C) 0.9599  D) 0.5589

9) The mean is \( \mu = 137.0 \) and the standard deviation is \( \sigma = 5.3 \). Find the probability that \( X \) is between 134.4 and 140.1.

   A) 0.4069  B) 0.8138  C) 1.0311  D) 0.6242

Solve the problem.

10) Assume that women have heights that are normally distributed with a mean of 63.6 inches and a standard deviation of 2.5 inches. Find the value of the quartile \( Q_3 \).

   A) 64.3 inches  B) 66.1 inches  C) 65.3 inches  D) 67.8 inches

Find the indicated probability.

11) The diameters of pencils produced by a certain machine are normally distributed with a mean of 0.30 inches and a standard deviation of 0.01 inches. What is the probability that the diameter of a randomly selected pencil will be less than 0.285 inches?

   A) 0.0596  B) 0.0668  C) 0.9332  D) 0.4332
CHAPTER 6 FORM A

Solve the problem.

12) 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 40 different applicants are randomly selected, find the probability that their mean is above 215.

A) 0.1179  
B) 0.0287  
C) 0.3821  
D) 0.4713

Provide an appropriate response.

13) A poll of 1100 randomly selected students in grades 6 through 8 was conducted and found that 47% enjoy playing sports. Would confidence in the results increase if the sample size were 3600 instead of 1100? Why or why not?

List the different possible samples, and find the mean of each of them.

14) Personal phone calls received in the last three days by a new employee were 3, 5, and 6. Assume that samples of size 2 are randomly selected with replacement from this population of three values.

Solve the problem.

15) A study of the amount of time it takes a mechanic to rebuild the transmission for a 1992 Chevrolet Cavalier shows that the mean is 8.4 hours and the standard deviation is 1.8 hours. If 40 mechanics are randomly selected, find the probability that their mean rebuild time exceeds 9.1 hours.

A) 0.1046  
B) 0.0069  
C) 0.0046  
D) 0.1285

Use the continuity correction and describe the region of the normal curve that corresponds to the indicated binomial probability.

16) The probability of no more than 51 defective CD’s

A) The area to the left of 51  
B) The area to the left of 50.5 
C) The area to the right of 51.5  
D) The area to the left of 51.5
CHAPTER 6 FORM A

For the binomial distribution with the given values for n and p, state whether or not it is suitable
to use the normal distribution as an approximation.

17) n = 47 and p = .9
   A) Normal approximation is suitable.
   B) Normal approximation is not suitable.

Estimate the indicated probability by using the normal distribution as an approximation to the
binomial distribution.

18) Two percent of hair dryers produced in a certain plant are defective. Estimate the
probability that of 10,000 randomly selected hair dryers, at least 219 are defective.
   A) 0.0823    B) 0.0869    C) 0.9066    D) 0.0934

Solve the problem.

19) A normal probability plot is given below for a sample of scores on an aptitude test. Use
the plot to assess the normality of scores on this test. Explain your reasoning.

![Normal Score vs. Test Score plot](image-url)
CHAPTER 6 FORM A

Examine the given data set and determine whether the requirement of a normal distribution is satisfied. Assume that the requirement for a normal distribution is loose in the sense that the population distribution need not be exactly normal, but it must have a distribution which is basically symmetric with only one mode. Explain why you do or do not think that the requirement is satisfied.

20) The data below represents the amount of television watched per week (in hours) for 40 randomly selected teenagers.

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>4</td>
<td>17</td>
<td>14</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>16</td>
<td>0</td>
<td>15</td>
<td>10</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>9</td>
<td>13</td>
<td>3</td>
<td>14</td>
<td>24</td>
<td>15</td>
<td>17</td>
</tr>
</tbody>
</table>
Answer Key
Testname: CHAPTER 6 FORM A

1) The first (a) may be solved by the Central Limit theorem, because the sample size is large. The probability is 0.0001. The second (b) has a sample size that is too small, and the Central Limit theorem does not apply.

2) B
3) C
4) D
5) C
6) B
7) A
8) A
9) A
10) C
11) B
12) B
13) Yes. As the sample size increases, the sample statistics tend to vary less and they tend to be closer to the population parameter.
14) Possible samples: 3–3; 3–5; 3–6; 5–3; 5–5; 5–6; 6–3; 6–5; 6–6
   Means: 3, 4, 4.5, 4, 5, 5.5, 4.5, 5.5, 6

15) B
16) D
17) B
18) D
19) Since the normal probability plot is roughly linear, it appears that scores on this test are approximately normally distributed.

20) The requirement for normality is not satisfied since a histogram of the data is not bell shaped; it is essentially bimodal.
CHAPTER 6 FORM B

Name:______________________________ Course Number:_______ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Complete the following table for a distribution in which \( \mu = 16 \). It might be helpful to make a diagram to help you determine the continuity factor for each entry.

<table>
<thead>
<tr>
<th>Event</th>
<th>Probability</th>
<th>Continuity Correction Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>X is at least 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X is at most 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X is more than 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X is less than 12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using the following uniform density curve, answer the question.

Using the following uniform density curve, answer the question.

2) What is the probability that the random variable has a value between 2.6 and 5.4?

A) 0.4750  B) 0.2250  C) 0.3500  D) 0.6000

Assume that the weight loss for the first month of a diet program varies between 6 pounds and 12 pounds, and is spread evenly over the range of possibilities, so that there is a uniform distribution. Find the probability of the given range of pounds lost.

3) Less than 10 pounds

A) \( \frac{1}{3} \)  B) \( \frac{2}{3} \)  C) \( \frac{1}{6} \)  D) \( \frac{5}{7} \)

If \( Z \) is a standard normal variable, find the probability.

4) The probability that \( Z \) lies between –0.55 and 0.55

A) 0.9000  B) –0.9000  C) 0.4176  D) –0.4176

5) \( P(Z > 0.59) \)

A) 0.7224  B) 0.2224  C) 0.2190  D) 0.2776
CHAPTER 6 FORM B

The Precision Scientific Instrument Company manufactures thermometers that are supposed to give readings of 0°C at the freezing point of water. Tests on a large sample of these thermometers reveal that at the freezing point of water, some give readings below 0°C (denoted by negative numbers) and some give readings above 0°C (denoted by positive numbers). Assume that the mean reading is 0°C and the standard deviation of the readings is 1.00°C. Also assume that the frequency distribution of errors closely resembles the normal distribution. A thermometer is randomly selected and tested. Find the temperature reading corresponding to the given information.

6) If 9% of the thermometers are rejected because they have readings that are too high, but all other thermometers are acceptable, find the temperature that separates the rejected thermometers from the others.
   A) 1.45°   B) 1.26°   C) 1.39°   D) 1.34°

7) If 6.3% of the thermometers are rejected because they have readings that are too high and another 6.3% are rejected because they have readings that are too low, find the two readings that are cutoff values separating the rejected thermometers from the others.
   A) −1.53°, 1.53°   B) −1.46°, 1.46°   C) −1.39°, 1.39°   D) −1.45°, 1.45°

Assume that X has a normal distribution, and find the indicated probability.

8) The mean is μ = 15.2 and the standard deviation is σ = 0.9.
Find the probability that X is greater than 15.2.
   A) 0.0003   B) 0.9998   C) 0.5000   D) 1.0000

9) The mean is μ = 22.0 and the standard deviation is σ = 2.4.
Find the probability that X is between 19.7 and 25.3.
   A) 1.0847   B) 0.7477   C) 0.4107   D) 0.3370

Solve the problem.

10) Suppose that replacement times for washing machines are normally distributed with a mean of 8.7 years and a standard deviation of 1.6 years. Find the replacement time that separates the top 18% from the bottom 82%.
   A) 9.6 years   B) 10.2 years   C) 9.0 years   D) 7.2 years

Find the indicated probability.

11) The volumes of soda in quart soda bottles are normally distributed with a mean of 32.3 oz and a standard deviation of 1.2 oz. What is the probability that the volume of soda in a randomly selected bottle will be less than 32 oz?
   A) 0.4013   B) 0.5987   C) 0.0987   D) 0.3821
CHAPTER 6 FORM B

Provide an appropriate response.

12) A recent survey based on a random sample of \( n = 490 \) voters, predicted that the Independent candidate for the mayoral election will get 23\% of the vote, but he actually gets 28\%. Can it be concluded that the survey was done incorrectly?

List the different possible samples, and find the mean of each of them.

13) The number of books sold over the course of the four-day book fair were 156, 200, 293, and 55. Assume that samples of size 2 are randomly selected with replacement from this population of four values.

Solve the problem.

14) The weights of the fish in a certain lake are normally distributed with a mean of 15 lb and a standard deviation of 6. If 4 fish are randomly selected, what is the probability that the mean weight will be between 12.6 and 18.6 lb?
   A) 0.0968        B) 0.3270        C) 0.6730        D) 0.4032

15) A study of the amount of time it takes a mechanic to rebuild the transmission for a 1992 Chevrolet Cavalier shows that the mean is 8.4 hours and the standard deviation is 1.8 hours. If 40 mechanics are randomly selected, find the probability that their mean rebuild time exceeds 8.7 hours.
   A) 0.1285        B) 0.1346        C) 0.1469        D) 0.1946

Use the continuity correction and describe the region of the normal curve that corresponds to the indicated binomial probability.

16) The probability of fewer than 56 democrats
   A) The area to the left of 56        B) The area to the left of 55.5
   C) The area to the right of 56.5     D) The area to the left of 56.5
CHAPTER 6 FORM B

For the binomial distribution with the given values for n and p, state whether or not it is suitable to use the normal distribution as an approximation.

17) n = 15 and p = .5
   A) Normal approximation is not suitable.
   B) Normal approximation is suitable.

Estimate the indicated probability by using the normal distribution as an approximation to the binomial distribution.

18) A multiple choice test consists of 60 questions. Each question has 4 possible answers of which one is correct. If all answers are random guesses, estimate the probability of getting at least 20% correct.
   A) 0.0901   B) 0.3508   C) 0.8508   D) 0.1492

Solve the problem.

19) A normal probability plot is given below for the lifetimes (in hours) of a sample of batteries of a particular brand. Use the plot to assess the normality of the lifetimes of these batteries. Explain your reasoning.
Examine the given data set and determine whether the requirement of a normal distribution is satisfied. Assume that the requirement for a normal distribution is loose in the sense that the population distribution need not be exactly normal, but it must have a distribution which is basically symmetric with only one mode. Explain why you do or do not think that the requirement is satisfied.

20) The heart rates (in beats per minute) of 30 randomly selected students are given below.

| 78 | 64 | 69 | 75 | 80 |
| 63 | 70 | 72 | 72 | 68 |
| 77 | 71 | 74 | 84 | 70 |
| 62 | 67 | 71 | 69 | 58 |
| 74 | 70 | 80 | 63 | 88 |
| 60 | 68 | 69 | 70 | 71 |
1) Find the probability that \( x \) is at least 12

<table>
<thead>
<tr>
<th>The continuity correction factor is:</th>
<th>11.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x ) is at least 12</td>
<td>11.5</td>
</tr>
<tr>
<td>( x ) is at most 12</td>
<td>12.5</td>
</tr>
<tr>
<td>( x ) is more than 12</td>
<td>12.5</td>
</tr>
<tr>
<td>( x ) is less than 12</td>
<td>11.5</td>
</tr>
</tbody>
</table>

2) C  
3) B  
4) C  
5) D  
6) D  
7) A  
8) C  
9) B  
10) B  
11) A

12) No, because of sampling variability, sample proportions will naturally vary from the true population proportion, even if sampling is done with a perfectly valid procedure.


14) C
15) C
16) B
17) B
18) C

19) Since the normal probability plot displays curvature, it appears that lifetimes of these batteries are probably not normally distributed.

20) The requirement for normality is satisfied since a histogram of the data is roughly bell shaped; it is roughly symmetric with a single mode.
CHAPTER 6 FORM C

Name: __________________________ Course Number: ________ Section Number: ________

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) According to data from the American Medical Association, 10% of us are left-handed. Suppose groups of 500 people are randomly selected. Find the probability that at least 80 are left-handed. Describe the characteristics of this problem which help you to recognize that the problem is about a binomial distribution which you are to solve by estimating with the normal distribution. (Assume that you would not use a computer, a table, or the binomial probability formula.)

Using the following uniform density curve, answer the question.

![Uniform Density Curve]

2) What is the probability that the random variable has a value greater than 1.6?
   A) 0.9250   B) 0.7500   C) 0.6750   D) 0.8000

Assume that the weight loss for the first month of a diet program varies between 6 pounds and 12 pounds, and is spread evenly over the range of possibilities, so that there is a uniform distribution. Find the probability of the given range of pounds lost.

3) Between 9.5 pounds and 11 pounds
   A) \( \frac{1}{2} \)   B) \( \frac{1}{3} \)   C) \( \frac{1}{4} \)   D) \( \frac{3}{4} \)

If \( Z \) is a standard normal variable, find the probability.

4) The probability that \( Z \) lies between 0.7 and 1.98
   A) -0.2181   B) 0.2181   C) 1.7341   D) 0.2175

5) \( P(Z < 0.97) \)
   A) 0.8078   B) 0.8340   C) 0.8315   D) 0.1660
CHAPTER 6 FORM C

The Precision Scientific Instrument Company manufactures thermometers that are supposed to give readings of 0°C at the freezing point of water. Tests on a large sample of these thermometers reveal that at the freezing point of water, some give readings below 0°C (denoted by negative numbers) and some give readings above 0°C (denoted by positive numbers). Assume that the mean reading is 0°C and the standard deviation of the readings is 1.00°C. Also assume that the frequency distribution of errors closely resembles the normal distribution. A thermometer is randomly selected and tested. Find the temperature reading corresponding to the given information.

6) If 9% of the thermometers are rejected because they have readings that are too low, but all other thermometers are acceptable, find the temperature that separates the rejected thermometers from the others.
   A) −1.39°  B) −1.45°  C) −1.26°  D) −1.34°

7) A quality control analyst wants to examine thermometers that give readings in the bottom 7%. Find the reading that separates the bottom 7% from the others.
   A) −1.63°  B) −1.75°  C) −1.48°  D) −1.89°

Assume that X has a normal distribution, and find the indicated probability.

8) The mean is μ = 15.2 and the standard deviation is σ = 0.9. Find the probability that X is greater than 16.1.
   A) 0.1550  B) 0.1587  C) 0.8413  D) 0.1357

9) The mean is μ = 15.2 and the standard deviation is σ = 0.9. Find the probability that X is between 14.3 and 16.1.
   A) 0.1587  B) 0.3413  C) 0.8413  D) 0.6826

Solve the problem.

10) The serum cholesterol levels for men in one age group are normally distributed with a mean of 178.1 and a standard deviation of 40.8. All units are in mg/100 mL. Find the two levels that separate the top 9% and the bottom 9%.
    A) 107.1 mg/100mL and 249.1 mg/100mL
    B) 165.0 mg/100mL and 191.16 mg/100mL
    C) 161.4 mg/100mL and 194.8 mg/100mL
    D) 123.4 mg/100mL and 232.8 mg/100mL

Find the indicated probability.

11) The incomes of trainees at a local mill are normally distributed with a mean of $1100 and a standard deviation $150. What percentage of trainees earn less than $900 a month?
    A) 9.18%  B) 40.82%  C) 35.31%  D) 90.82%
CHAPTER 6 FORM C

Provide an appropriate response.

12) A poll of 1400 randomly selected students in grades 6 through 8 was conducted and found that 30% enjoy playing sports. Is the 30% result a statistic or a parameter? Explain.

List the different possible samples, and find the mean of each of them.

13) Flood insurance policies sold in the last three days by a new agent were 2, 6, and 7. Assume that samples of size 2 are randomly selected with replacement from this population of three values.

Solve the problem.

14) For women aged 18–24, systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1. If 23 women aged 18–24 are randomly selected, find the probability that their mean systolic blood pressure is between 119 and 122.

A) 0.0577  B) 0.9341  C) 0.3343  D) 0.0833

15) A study of the amount of time it takes a mechanic to rebuild the transmission for a 1992 Chevrolet Cavalier shows that the mean is 8.4 hours and the standard deviation is 1.8 hours. If 40 mechanics are randomly selected, find the probability that their mean rebuild time exceeds 7.7 hours.

A) 0.9931  B) 0.9712  C) 0.8531  D) 0.9634

Use the continuity correction and describe the region of the normal curve that corresponds to the indicated binomial probability.

16) The probability of at least 53 boys

A) The area to the left of 52.5  B) The area to the right of 53
C) The area to the right of 52.5  D) The area to the right of 53.5
CHAPTER 6 FORM C

For the binomial distribution with the given values for n and p, state whether or not it is suitable to use the normal distribution as an approximation.

17) n = 18 and p = .2
   A) Normal approximation is suitable.
   B) Normal approximation is not suitable.

Estimate the indicated probability by using the normal distribution as an approximation to the binomial distribution.

18) The probability that a radish seed will germinate is 0.7. Estimate the probability that of 140 randomly selected seeds, exactly 100 will germinate.
   A) 0.0669        B) 0.0769        C) 0.9331        D) 0.0679

Solve the problem.

19) A normal probability plot is given below for a sample of scores on an aptitude test. Use the plot to assess the normality of scores on this test. Explain your reasoning.
CHAPTER 6 FORM C

Examine the given data set and determine whether the requirement of a normal distribution is satisfied. Assume that the requirement for a normal distribution is loose in the sense that the population distribution need not be exactly normal, but it must have a distribution which is basically symmetric with only one mode. Explain why you do or do not think that the requirement is satisfied.

20) The numbers obtained on 50 rolls of a die.

1 5 5 3 6 4 5 6 3 4
2 5 3 5 4 2 1 4 3 1
6 1 2 6 1 2 5 3 3 4
4 1 3 1 6 2 2 5 5 3
3 5 1 6 2 1 1 4 6 5
Answer Key
Testname: CHAPTER 6 FORM C

1) The problem statement has a stated percent, 10%, which corresponds to the p value, and a fixed number of trials, 500, which corresponds to n. No mean or standard deviation is given, and they are to be computed using the formulas $\mu = n \cdot p$ and $\sigma = \sqrt{n \cdot p \cdot q}$. Also, the normal approximation to the binomial distribution is appropriate, since both np and nq are greater than 5. $P = 0.0001$.

2) D
3) C
4) B
5) B
6) D
7) C
8) B
9) D
10) D
11) A
12) Statistic, because it is calculated from a sample, not a population.
13) Possible samples: 2–2; 2–6; 2–7; 6–2; 6–6; 6–7; 7–2; 7–6; 7–7.
   Means: 2, 4, 4.5, 4, 6, 6.5, 4.5, 6.5, 7
14) A
15) A
16) C
17) B
18) A
19) Since the normal probability plot is roughly linear, it appears that scores on this test are approximately normally distributed.
20) The requirement for normality is not satisfied since a histogram of the data is not roughly bell shaped with a single mode. The data are roughly uniformly distributed.
CHAPTER 7 FORM A

Name:_____________________________ Course Number:__________ Section Number:_____

Directions: Circle the correct choice for multiple-choice items.

Solve the problem.

1) Find the value of $-z_{\alpha/2}$ that corresponds to a level of confidence of 94.76 percent.
   A) 0.0262   B) -1.62   C) -1.94   D) 1.94

2) The following confidence interval is obtained for a population proportion, $p$:
   $0.620 < p < 0.658$
   Use these confidence interval limits to find the margin of error, $E$.
   A) 0.020   B) 0.019   C) 0.639   D) 0.038

Find the margin of error for the 95% confidence interval used to estimate the population proportion.

3) In a survey of 7100 T.V. viewers, 40% said they watch network news programs.
   A) 0.0150   B) 0.0131   C) 0.00855   D) 0.0114

Use the given degree of confidence and sample data to construct a confidence interval for the population proportion $p$.

4) $n = 87$, $x = 48$; 98 percent
   A) $0.448 < p < 0.656$   B) $0.428 < p < 0.676$
   C) $0.427 < p < 0.677$   D) $0.447 < p < 0.657$

Find the minimum sample size you should use to assure that your estimate of $\hat{p}$ will be within the required margin of error around the population $p$.

5) Margin of error: 0.006; confidence level: 95%; $\hat{p}$ and $\hat{q}$ are unknown
   A) 38,416   B) 16,112   C) 26,678   D) 38,415

6) Margin of error: 0.04; confidence level: 90%; from a prior study, $\hat{p}$ is estimated by 0.28.
   A) 14   B) 341   C) 1023   D) 303

Solve the problem.

7) 50 people are selected randomly from a certain population and it is found that 14 people in the sample are over 6 feet tall. What is the point estimate of the true proportion of people in the population who are over 6 feet tall?
   A) 0.50   B) 0.28   C) 0.72   D) 0.20
CHAPTER 7 FORM A

Use the given degree of confidence and sample data to construct a confidence interval for the population proportion $p$.

8) A survey of 865 voters in one state reveals that 408 favor approval of an issue before the legislature. Construct the 95% confidence interval for the true proportion of all voters in the state who favor approval.

A) $0.444 < p < 0.500$
B) $0.435 < p < 0.508$
C) $0.438 < p < 0.505$
D) $0.471 < p < 0.472$

Solve the problem.

9) Suppose that $n$ trials of a binomial experiment result in no successes. According to the "Rule of Three", we have 95% confidence that the true population proportion has an upper bound of $3/n$. If a manufacturer randomly selects 25 computers for quality control and finds no defective computers, what statement can you make by using the rule of three, about the proportion $p$, of all its computers which are defective?

A) We are 95% confident that $p$ does not exceed $\frac{3}{25}$.
B) We are 95% confident that $p$ lies between $\frac{1}{25}$ and $\frac{3}{25}$.
C) We are 95% confident that $p$ is greater than $\frac{3}{25}$.
D) The value of $p$ cannot be greater than $\frac{3}{25}$.

Determine whether the given conditions justify using the margin of error $E = z_{\alpha/2} \sigma/\sqrt{n}$ when finding a confidence interval estimate of the population mean $\mu$.

10) The sample size is $n = 7$, $\sigma = 12.2$, and the original population is normally distributed.

A) Yes  
B) No

Use the confidence level and sample data to find the margin of error $E$.

11) Weights of eggs: 95% confidence; $n = 46$, $\bar{x} = 1.68$ oz, $\sigma = 0.46$ oz
A) 0.11 oz  
B) 0.02 oz  
C) 0.13 oz  
D) 6.78 oz

Use the confidence level and sample data to find a confidence interval for estimating the population $\mu$.

12) A random sample of 94 light bulbs had a mean life of $\bar{x} = 587$ hours with a standard deviation of $\sigma = 36$ hours. Construct a 90 percent confidence interval for the mean life, $\mu$, of all light bulbs of this type.

A) $578 < \mu < 596$  
B) $581 < \mu < 593$  
C) $577 < \mu < 597$  
D) $580 < \mu < 594$
CHAPTER 7 FORM A

Use the margin of error, confidence level, and standard deviation $\sigma$ to find the minimum sample size required to estimate an unknown population mean $\mu$.

13) Margin of error: $133, confidence level: 95\%, \sigma = 530$
   A) 20 B) 62 C) 54 D) 50

Do one of the following, as appropriate: (a) Find the critical value $z_{\alpha/2}$, (b) find the critical value $t_{\alpha/2}$, (c) state that neither the normal nor the $t$ distribution applies.

14) 95%; $n = 11; \sigma$ is known; population appears to be very skewed.
   A) Neither the normal nor the $t$ distribution applies.
   B) $z_{\alpha/2} = 1.96$
   C) $z_{\alpha/2} = 1.812$
   D) $t_{\alpha/2} = 2.228$

Find the margin of error.

15) 95% confidence interval; $n = 91; \bar{x} = 28, s = 14.1$
   A) 1.85 B) 4.25 C) 2.05 D) 2.93

Use the given degree of confidence and sample data to construct a confidence interval for the population mean $\mu$. Assume that the population has a normal distribution.

16) The principal randomly selected six students to take an aptitude test. Their scores were: 83.0 84.1 83.5 84.7 84.1 73.5
   Determine a 90 percent confidence interval for the mean score for all students.
   A) 85.52 < $\mu$ < 78.45 B) 85.42 < $\mu$ < 78.55
   C) 78.55 < $\mu$ < 85.42 D) 78.45 < $\mu$ < 85.52

Solve the problem.

17) Find the chi-square value $\chi^2$ corresponding to a sample size of 20 and a confidence level of 99 percent.
   A) 6.844 B) 7.633 C) 36.191 D) 38.582

Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation $\sigma$. Assume that the population has a normal distribution.

18) A sociologist develops a test to measure attitudes about public transportation, and 27 randomly selected subjects are given the test. Their mean score is 76.2 and their standard deviation is 21.4. Construct the 95% confidence interval for the standard deviation, $\sigma$, of the scores of all subjects.
   A) 16.6 < $\sigma$ < 28.6 B) 17.2 < $\sigma$ < 27.2
   C) 16.9 < $\sigma$ < 29.3 D) 17.5 < $\sigma$ < 27.8
Find the appropriate minimum sample size.

19) You want to be 99% confident that the sample standard deviation $s$ is within 5% of the population standard deviation.

A) 1,335  
B) 923  
C) 2,638  
D) 2,434

Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation $\sigma$. Assume that the population has a normal distribution.

20) The daily intakes of milk (in ounces) for ten randomly selected people were:

\[22.3 \ 12.9 \ 23.1 \ 26.3 \ 10.8 \]
\[26.8 \ 27.4 \ 29.7 \ 19.0 \ 19.9\]

Find a 99 percent confidence interval for the population standard deviation $\sigma$.

A) (3.74, 12.81)  
B) (0.89, 3.40)  
C) (3.87, 12.81)  
D) (3.87, 14.26)
Answer Key
Testname: CHAPTER 7 FORM A

1) C
2) B
3) D
4) B
5) C
6) B
7) B
8) C
9) A
10) A
11) C
12) B
13) B
14) A
15) D
16) C
17) A
18) C
19) A
20) D
CHAPTER 7 FORM B

Name:___________________________ Course Number:_______ Section Number:_____

Directions: Circle the correct choice for multiple-choice items.

Solve the problem.

1) Find the critical value \( z_{\alpha/2} \) that corresponds to a degree of confidence of 98%.
   A) 2.575      B) 2.05      C) 2.33      D) 1.75

2) The following confidence interval is obtained for a population proportion, \( p \): 
   (0.870, 0.894)
   Use these confidence interval limits to find the point estimate, \( \hat{p} \).
   A) 0.870      B) 0.882      C) 0.885      D) 0.894

Find the margin of error for the 95% confidence interval used to estimate the population proportion.

3) In a clinical test with 2440 subjects, 70% showed improvement from the treatment.
   A) 0.0236      B) 0.0196      C) 0.0175      D) 0.0182

Use the given degree of confidence and sample data to construct a confidence interval for the population proportion \( p \).

4) \( n = 107, x = 66; 88 \% \) percent
   A) 0.540 < \( p < 0.694 \)  
   B) 0.544 < \( p < 0.690 \)  
   C) 0.543 < \( p < 0.691 \)  
   D) 0.539 < \( p < 0.695 \)

Find the minimum sample size you should use to assure that your estimate of \( \hat{p} \) will be within the required margin of error around the population \( p \).

5) Margin of error: 0.07; confidence level: 97%; \( \hat{p} \) and \( \hat{q} \) unknown
   A) 241      B) 112      C) 240      D) 111

6) Margin of error: 0.008; confidence level: 99%; from a prior study, \( \hat{p} \) is estimated by 0.164
   A) 1145      B) 12,785      C) 14,205      D) 8230

Solve the problem.

7) 38 randomly picked people were asked if they rented or owned their own home, 11 said they rented. Obtain a point estimate of the true proportion of home owners.
   A) 0.289      B) 0.224      C) 0.737      D) 0.711
CHAPTER 7 FORM B

Use the given degree of confidence and sample data to construct a confidence interval for the population proportion p.

8) A survey of 300 union members in New York State reveals that 112 favor the Republican candidate for governor. Construct the 98% confidence interval for the true population proportion of all New York State union members who favor the Republican candidate.

\[ A) \ 0.316 < p < 0.430 \quad B) \ 0.304 < p < 0.442 \]
\[ C) \ 0.308 < p < 0.438 \quad D) \ 0.301 < p < 0.445 \]

Solve the problem.

9) A one-sided confidence interval for p can be written as \( p < \hat{p} + E \) or \( p > \hat{p} - E \) where the margin of error E is modified by replacing \( z_{\alpha/2} \) with \( z_{\alpha} \). If a teacher wants to report that the fail rate on a test is at most \( x \) with 90% confidence, construct the appropriate one-sided confidence interval. Assume that a simple random sample of 58 students results in 6 who fail the test.

\[ A) \ p < 0.052 \quad B) \ p > 0.155 \quad C) \ p < 0.155 \quad D) \ p > 0.052 \]

Determine whether the given conditions justify using the margin of error \( E = z_{\alpha/2} \sigma/\sqrt{n} \) when finding a confidence interval estimate of the population mean \( \mu \).

10) The sample size is \( n = 8 \) and \( \sigma \) is not known.

\[ A) \ \text{No} \quad \quad B) \ \text{Yes} \]

Use the confidence level and sample data to find the margin of error E.

11) College students' annual earnings: 99% confidence; \( n = 66, \bar{x} = $3903, \sigma = $801 \)

\[ A) \ $230 \quad B) \ $8 \quad C) \ $254 \quad D) \ $1237 \]

Use the confidence level and sample data to find a confidence interval for estimating the population \( \mu \).

12) A random sample of 100 full-grown lobsters had a mean weight of 22 ounces and a standard deviation of 3.7 ounces. Construct a 98 percent confidence interval for the population mean \( \mu \).

\[ A) \ 21 < \mu < 24 \quad B) \ 22 < \mu < 24 \quad C) \ 20 < \mu < 22 \quad D) \ 21 < \mu < 23 \]

Use the margin of error, confidence level, and standard deviation \( \sigma \) to find the minimum sample size required to estimate an unknown population mean \( \mu \).

13) Margin of error: $136, confidence level: 99%, \( \sigma = $503 \)

\[ A) \ 10 \quad B) \ 91 \quad C) \ 53 \quad D) \ 46 \]
CHAPTER 7 FORM B

Do one of the following, as appropriate: (a) Find the critical value $z_{\alpha/2}$; (b) find the critical value $t_{\alpha/2}$; (c) state that neither the normal nor the t distribution applies.

14) 90%; $n = 9$; $\sigma = 4.2$; population appears to be very skewed.
   A) Neither the normal nor the t distribution applies.
   B) $z_{\alpha/2} = 2.365$
   C) $z_{\alpha/2} = 2.306$
   D) $z_{\alpha/2} = 2.896$

Find the margin of error.

15) 95% confidence interval; $n = 12$; $\bar{x} = 35.6$; $s = 6.4$
   A) $3.659$  B) $3.050$  C) $4.879$  D) $4.066$

Use the given degree of confidence and sample data to construct a confidence interval for the population mean $\mu$. Assume that the population has a normal distribution.

16) The amounts (in ounces) of juice in eight randomly selected juice bottles are:
   15.1 15.6 15.6 15.9
   15.9 15.2 15.1 15.2
   Construct a 98 percent confidence interval for the mean amount of juice in all such bottles.
   A) $15.87 < \mu < 15.03$  B) $15.09 < \mu < 15.81$
   C) $15.77 < \mu < 15.13$  D) $15.03 < \mu < 15.87$

Solve the problem.

17) Find the critical value $\chi^2_L$ corresponding to a sample size of 3 and a confidence level of 90 percent.
   A) $0.103$  B) $0.0201$  C) $9.21$  D) $5.991$

Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation $\sigma$. Assume that the population has a normal distribution.

18) Weights of men: 90% confidence; $n = 14$, $\bar{x} = 156.9$ lb, $s = 11.1$ lb
   A) $8.5$ lb $< \sigma < 16.5$ lb  B) $9.0$ lb $< \sigma < 2.7$ lb
   C) $8.2$ lb $< \sigma < 15.6$ lb  D) $8.7$ lb $< \sigma < 14.3$ lb

Find the appropriate minimum sample size.

19) To be able to say with 95% confidence level that the standard deviation of a data set is within 10% of the population’s standard deviation, the number of observations within the data set must be greater than or equal to what quantity?
   A) $805$  B) $250$  C) $335$  D) $191$
CHAPTER 7 FORM B

Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation $\sigma$. Assume that the population has a normal distribution.

20) The amounts (in ounces) of juice in eight randomly selected juice bottles are:
   15.8  15.7  15.0  15.7  
   15.2  15.2  15.5  15.0

Find a 98 percent confidence interval for the population standard deviation $\sigma$.

A) (0.22, 0.84)  B) (0.20, 0.67)  C) (0.20, 0.78)  D) (0.19, 0.67)
Answer Key
Testname: CHAPTER 7 FORM B

1) C
2) B
3) D
4) B
5) A
6) C
7) D
8) C
9) C
10) A
11) C
12) D
13) B
14) A
15) D
16) B
17) A
18) A
19) D
20) C
CHAPTER 7 FORM C

Name:____________________________ Course Number:__________ Section Number:_____

Directions: Circle the correct choice for multiple-choice items.

Solve the problem.

1) Find the critical value $z_{\alpha/2}$ that corresponds to a degree of confidence of 91%.
   A) 1.70  B) 1.34  C) 1.645  D) 1.75

2) The following confidence interval is obtained for a population proportion, $p$: $0.817 < p < 0.855$
   Use these confidence interval limits to find the point estimate, $\hat{p}$.
   A) 0.839  B) 0.836  C) 0.817  D) 0.833

Find the margin of error for the 95% confidence interval used to estimate the population proportion.

3) $n = 186, x = 103$
   A) 0.0643  B) 0.125  C) 0.00260  D) 0.0714

Use the given degree of confidence and sample data to construct a confidence interval for the population proportion $p$.

4) $n = 158, x = 108; 95$ percent
   A) $0.625 < p < 0.743$  B) $0.626 < p < 0.742$
   C) $0.611 < p < 0.756$  D) $0.610 < p < 0.758$

Find the minimum sample size you should use to assure that your estimate of $\hat{p}$ will be within the required margin of error around the population $p$.

5) Margin of error: 0.002; confidence level: 93%; $\hat{p}$ and $\hat{q}$ unknown
   A) 204,757  B) 410  C) 204,750  D) 405

6) Margin of error: 0.07; confidence level: 95%; from a prior study, $\hat{p}$ is estimated by the decimal equivalent of 92%.
   A) 58  B) 174  C) 51  D) 4

Solve the problem.

7) 61 randomly selected light bulbs were tested in a laboratory, 50 lasted more than 500 hours. Find a point estimate of the true proportion of all light bulbs that last more than 500 hours.
   A) 0.803  B) 0.450  C) 0.820  D) 0.180
Use the given degree of confidence and sample data to construct a confidence interval for the population proportion \( p \).

8) When 343 college students are randomly selected and surveyed, it is found that 110 own a car. Find a 99% confidence interval for the true proportion of all college students who own a car.

A) \( 0.256 < p < 0.386 \)  
B) \( 0.279 < p < 0.362 \)  
C) \( 0.271 < p < 0.370 \)  
D) \( 0.262 < p < 0.379 \)

Solve the problem.

9) A researcher is interested in estimating the proportion of voters who favor a tax on e-commerce. Based on a sample of 250 people, she obtains the following 99% confidence interval for the population proportion \( p \):

\[ 0.113 < p < 0.171 \]

Which of the statements below is a valid interpretation of this confidence interval?

A: There is a 99% chance that the true value of \( p \) lies between 0.113 and 0.171.
B: If many different samples of size 250 were selected and, based on each sample, a confidence interval were constructed, 99% of the time the true value of \( p \) would lie between 0.113 and 0.171.
C: If many different samples of size 250 were selected and, based on each sample, a confidence interval were constructed, in the long run 99% of the confidence intervals would contain the true value of \( p \).
D: If 100 different samples of size 250 were selected and, based on each sample, a confidence interval were constructed, exactly 99 of these confidence intervals would contain the true value of \( p \).

A) C  
B) B  
C) D  
D) A

Determine whether the given conditions justify using the margin of error \( E = z_{\alpha/2} \sigma/\sqrt{n} \) when finding a confidence interval estimate of the population mean \( \mu \).

10) The sample size is \( n = 9 \), \( \sigma \) is not known, and the original population is normally distributed.

A) Yes  
B) No

Use the confidence level and sample data to find the margin of error \( E \).

11) Systolic blood pressures for women aged 18–24: 94% confidence; \( n = 92 \), \( \bar{x} = 114.9 \) mm Hg, \( \sigma = 13.2 \) mm Hg

A) 47.6 mm Hg  
B) 2.3 mm Hg  
C) 2.6 mm Hg  
D) 9.6 mm Hg
CHAPTER 7 FORM C

Use the confidence level and sample data to find a confidence interval for estimating the population \( \mu \).

12) A group of 52 randomly selected students have a mean score of 20.2 with a standard deviation of 4.6 on a placement test. What is the 90 percent confidence interval for the mean score, \( \mu \), of all students taking the test?
   - A) 19.1 < \( \mu \) < 21.3
   - B) 18.7 < \( \mu \) < 21.7
   - C) 19.0 < \( \mu \) < 21.5
   - D) 18.6 < \( \mu \) < 21.8

Use the margin of error, confidence level, and standard deviation \( \sigma \) to find the minimum sample size required to estimate an unknown population mean \( \mu \).

13) Margin of error: $100, confidence level: 95\%, \sigma = $403
   - A) 91
   - B) 63
   - C) 108
   - D) 44

Do one of the following, as appropriate: (a) Find the critical value \( z_{\alpha/2} \), (b) find the critical value \( t_{\alpha/2} \), (c) state that neither the normal nor the t distribution applies.

14) 99%; \( n = 17 \); \( \sigma \) is unknown; population appears to be normally distributed.
   - A) \( t_{\alpha/2} = 2.921 \)
   - B) \( z_{\alpha/2} = 2.583 \)
   - C) \( t_{\alpha/2} = 2.898 \)
   - D) \( z_{\alpha/2} = 2.567 \)

Find the margin of error.

15) 99% confidence interval; \( n = 201 \); \( \bar{x} = 175 \); \( s = 21 \)
   - A) 6.0
   - B) 4.6
   - C) 8.2
   - D) 3.9

Use the given degree of confidence and sample data to construct a confidence interval for the population mean \( \mu \). Assume that the population has a normal distribution.

16) The football coach randomly selected ten players and timed how long each player took to perform a certain drill. The times (in minutes) were:
    - 7.6 10.4 9.7 8.4 11.8
    - 7.0 6.5 11.1 10.4 12.4
    Determine a 95 percent confidence interval for the mean time for all players.
   - A) 8.03 < \( \mu \) < 11.03
   - B) 8.00 < \( \mu \) < 10.98
   - C) 8.06 < \( \mu \) < 11.00
   - D) 8.13 < \( \mu \) < 10.93

Solve the problem.

17) Find the critical value \( \chi^2_R \) corresponding to a sample size of 15 and a confidence level of 90 percent.
   - A) 31.319
   - B) 23.685
   - C) 29.141
   - D) 21.064
CHAPTER 7 FORM C

Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation $\sigma$. Assume that the population has a normal distribution.

18) College students' annual earnings: 98% confidence; $n = 9$, $\bar{x} = $3605, $s = $800
   A) $486 < \sigma < 1566$    B) $540 < \sigma < 1533$
   C) $629 < \sigma < 1044$    D) $505 < \sigma < 15764$

Find the appropriate minimum sample size.

19) You want to be 95% confident that the sample variance is within 40% of the population variance.
   A) 224    B) 14    C) 56    D) 11

Use the given degree of confidence and sample data to find a confidence interval for the population standard deviation $\sigma$. Assume that the population has a normal distribution.

20) The football coach randomly selected ten players and timed how long each player took to perform a certain drill. The times (in minutes) were:
   7   12   10   14   10
   14   9   11    7   14
   Find a 95 percent confidence interval for the population standard deviation $\sigma$.
   A) (1.9, 4.9)    B) (0.7, 2.2)    C) (1.8, 4.5)    D) (1.9, 4.5)
Answer Key
Testname: CHAPTER 7 FORM C

1) A
2) B
3) D
4) C
5) A
6) A
7) C
8) A
9) A
10) B
11) C
12) A
13) B
14) A
15) D
16) C
17) B
18) D
19) C
20) A
CHAPTER 8 FORM A

Name: ___________________________ Course Number: _______ Section Number: _______

**Directions:** Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

**Provide an appropriate response.**

1) Define Type I and Type II errors. Give an example of a Type I error which would have serious consequences. Give an example of a Type II error which would have serious consequences. What should be done to minimize the consequences of a serious Type I error?

**Solve the problem.**

2) Write the claim that is suggested by the given statement, then write a conclusion about the claim. Do not use symbolic expressions or formal procedures; use common sense.

A math teacher tries a new method for teaching her introductory statistics class. Last year the mean score on the final test was 73. This year the mean on the same final was 76.

**Express the null hypothesis H₀ and the alternative hypothesis H₁ in symbolic form. Use the correct symbol (μ, p, σ) for the indicated parameter.**

3) An entomologist writes an article in a scientific journal which claims that fewer than 12 in ten thousand male fireflies are unable to produce light due to a genetic mutation. Use the parameter p, the true proportion of fireflies unable to produce light.

A) H₀: p > 0.0012  
   H₁: p ≤ 0.0012 

B) H₀: p < 0.0012  
   H₁: p ≥ 0.0012 

C) H₀: p = 0.0012  
   H₁: p > 0.0012 

D) H₀: p = 0.0012  
   H₁: p < 0.0012
CHAPTER 8 FORM A

Assume that the data has a normal distribution and the number of observations is greater than fifty. Find the critical z value used to test a null hypothesis.

4) \( \alpha = 0.08; H_1 \) is \( \mu \neq 3.24 \)

A) 1.41  B) ±1.41  C) 1.75  D) ±1.75

Find the value of the test statistic \( z \) using \( z = \frac{\bar{X} - \mu}{\sqrt{\frac{pq}{n}}} \).

5) A claim is made that the proportion of children who play sports is less than 0.5, and the sample statistics include \( n = 1933 \) subjects with 30% saying that they play a sport.

A) 17.59  B) 35.90  C) −35.90  D) −17.59

Use the given information to find the P-value.

6) The test statistic in a left-tailed test is \( z = -2.05 \).

A) 0.0453  B) 0.0202  C) 0.5000  D) 0.4798

Formulate the indicated conclusion in nontechnical terms. Be sure to address the original claim.

7) Carter Motor Company claims that its new sedan, the Libra, will average better than 21 miles per gallon in the city. Assuming that a hypothesis test of the claim has been conducted and that the conclusion is to reject the null hypothesis, state the conclusion in nontechnical terms.

A) There is sufficient evidence to support the claim that the mean is less than 21 miles per gallon.
B) There is sufficient evidence to support the claim that the mean is greater than 21 miles per gallon.
C) There is not sufficient evidence to support the claim that the mean is greater than 21 miles per gallon.
D) There is not sufficient evidence to support the claim that the mean is less than 21 miles per gallon.

Assume that a hypothesis test of the given claim will be conducted. Identify the type I or type II error for the test.

8) The principal of a middle school claims that test scores of the seventh-graders at his school vary less than the test scores of seventh-graders at a neighboring school, which have variation described by \( \sigma = 14.7 \). Identify the type I error for the test.

A) The error of failing to reject the claim that the standard deviation is less than 14.7 when it actually is 14.7.
B) The error of rejecting the claim that the standard deviation is less than 14.7 when it actually is 14.7.
C) The error of rejecting the claim that the standard deviation is less than 14.7 when it actually is less than 14.7.
CHAPTER 8 FORM A

Solve the problem.

9) True or False: In a hypothesis test, an increase in $\alpha$ will cause a decrease in the power of the test provided the sample size is kept fixed.
   A) False  B) True

Identify the null hypothesis, alternative hypothesis, test statistic, $P$-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

10) A supplier of 3.5" disks claims that no more than 1% of the disks are defective. In a random sample of 600 disks, it is found that 3% are defective, but the supplier claims that this is only a sample fluctuation. At the 0.01 level of significance, test the supplier’s claim that no more than 1% are defective.

Find the $P$-value for the indicated hypothesis test.

11) A random sample of 139 forty-year-old men contains 26% smokers. Find the $P$-value for a test of the claim that the percentage of forty-year-old men that smoke is 22%.
   A) 0.2802  B) 0.1271  C) 0.2542  D) 0.1401

Determine whether the given conditions justify testing a claim about a population mean $\mu$.

12) The sample size is $n = 22$, $\sigma = 6.20$, and the original population is normally distributed.
   A) Yes  B) No

Identify the null hypothesis, alternative hypothesis, test statistic, $P$-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

13) A random sample of 100 pumpkins is obtained and the mean circumference is found to be 40.5 cm. Assuming that the population standard deviation is known to be 1.6 cm, use a 0.05 significance level to test the claim that the mean circumference of all pumpkins is equal to 39.9 cm.
CHAPTER 8 FORM A

Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, Student t distribution, or neither.

14) Claim: \( \mu = 105 \). Sample data: \( n = 18, \bar{x} = 101, s = 15.1 \). The sample data appear to come from a normally distributed population with unknown \( \mu \) and \( \sigma \).

   A) Student t  
   B) Neither  
   C) Normal

Assume that a simple random sample has been selected from a normally distributed population. Find the test statistic, P-value, critical value(s), and state the final conclusion.

15) Test the claim that the mean lifetime of car engines of a particular type is greater than 220,000 miles. Sample data are summarized as \( n = 23, \bar{x} = 226,450 \) miles, and \( s = 11,500 \) miles. Use a significance level of \( \alpha = 0.01 \).

Test the given claim using the traditional method of hypothesis testing. Assume that the sample has been randomly selected from a population with a normal distribution.

16) A large software company gives job applicants a test of programming ability and the mean for that test has been 160 in the past. Twenty-five job applicants are randomly selected from one large university and they produce a mean score and standard deviation of 183 and 12, respectively. Use a 0.05 level of significance to test the claim that this sample comes from a population with a mean score greater than 160.

17) In tests of a computer component, it is found that the mean time between failures is 520 hours. A modification is made which is supposed to increase the time between failures. Tests on a random sample of 10 modified components resulted in the following times (in hours) between failures.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>518</td>
<td>548</td>
<td>561</td>
<td>523</td>
<td>536</td>
</tr>
<tr>
<td>499</td>
<td>538</td>
<td>557</td>
<td>528</td>
<td>563</td>
</tr>
</tbody>
</table>

At the 0.05 significance level, test the claim that for the modified components, the mean time between failures is greater than 520 hours.
CHAPTER 8 FORM A

Find the critical value or values of $\chi^2$ based on the given information.

18) $H_1: \sigma \neq 9.3$
   
   $n = 28$
   
   $\alpha = 0.05$
   
   A) 14.573, 43.194  
   B) -14.573, 14.573  
   C) 16.151, 40.113  
   D) -40.113, 40.113

Use the traditional method to test the given hypothesis. Assume that the population is normally distributed and that the sample has been randomly selected.

19) The standard deviation of math test scores at one high school is 16.1. A teacher claims that the standard deviation of the girls' test scores is smaller than 16.1. A random sample of 22 girls results in scores with a standard deviation of 13.3. Use a significance level of 0.01 to test the teacher's claim.

20) Heights of men aged 25 to 34 have a standard deviation of 2.9. Use a 0.05 significance level to test the claim that the heights of women aged 25 to 34 have a different standard deviation. The heights (in inches) of 16 randomly selected women aged 25 to 34 are listed below.

<table>
<thead>
<tr>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>62.13</td>
<td>65.09</td>
<td>64.18</td>
<td>66.72</td>
<td>63.09</td>
<td>61.15</td>
<td>67.50</td>
<td>64.65</td>
</tr>
<tr>
<td>63.80</td>
<td>64.21</td>
<td>60.17</td>
<td>68.28</td>
<td>66.49</td>
<td>62.10</td>
<td>65.73</td>
<td>64.72</td>
</tr>
</tbody>
</table>
1) Type I: The mistake of rejecting the null hypothesis when it is true. Type II: The mistake of failing to reject the null hypothesis when it is false. Answers for examples will vary. To minimize a Type I error with serious consequences, make \( \alpha \) smaller. Also, make the sample size larger to minimize both \( \alpha \) and \( \beta \).

Possible example for Type I: Pharmaceutical Company A produces a new drug believed to be superior to the one it will replace. If a hypothesis test suggests this improved drug is better, the company will spend millions of dollars preparing and marketing its new drug, when in reality the former one was better. Possible example for Type II: Pharmaceutical Company A produces a new drug. This time a hypothesis test supports the null hypothesis with the result that Company A abandons production of this improved drug. In reality, the improved drug is better; and A loses millions of dollars, not only because of the cost of production of the improved drug but also because its existing drug is not competitive.

2) The claim is that the new teaching method is more effective than the old method and that on average students will score higher when she uses the new teaching method than when she uses the old teaching method. The small difference in the two means is not strong evidence that the new method is more effective. Even if both methods were equally effective, such a difference could easily occur by chance.

3) D
4) D
5) D
6) B
7) B
8) A
9) A

10) \( H_0: p = 0.01 \). \( H_1: p > 0.01 \). Test statistic: \( z = 4.92 \). P-value: \( p = 0.0001 \) (Table A-2), \( p = 0.0000 \) (STATDISK), \( p = 8.506511E-7 \) (TI-83/84 Plus).

Reject null hypothesis. There is sufficient evidence to warrant rejection of the claim that no more than 1% are defective.

Note: Since the term “no more than” is used, the translation is \( p \leq 0.01 \). Therefore, the competing hypothesis is \( p > 0.01 \).

11) C
12) A

13) \( H_0: \mu = 39.9 \). \( H_1: \mu \neq 39.9 \). Test statistic: \( z = 3.75 \). P-value: \( p = 0.0002 \) (Table A-2, STATDISK), \( p = 0.000177 \) (TI-83/84 Plus). Reject \( H_0 \). There is sufficient evidence to warrant rejection of the claim that the mean equals 39.9 cm.

14) A

15) \( \alpha = 0.01 \)

Test statistic: \( t = 2.6898 \)

P-value: 0.005 < P-value < 0.01 (Table A-3), \( p = 0.0067 \) (STATDISK), \( p = 0.00669 \) (TI-83/84 Plus).

Critical value: \( t = 2.508 \)

Because TS \( t > 2.508 \), we reject the null hypothesis. There is sufficient evidence to accept the

16) Test statistic: \( t = 9.583 \). Critical value: \( t = 1.711 \). Reject the null hypothesis. There is sufficient evidence to support the claim that the mean is greater than 160.

17) Test statistic: \( t = 2.612 \). Critical value: \( t = 1.833 \). Reject \( H_0 \). There is sufficient evidence to support the claim that the mean is greater than 520 hours.

18) A
19) Test statistic: $X^2 = 14.331$. Critical value: $X^2 = 8.897$. Fail to reject $H_0$. There is not sufficient evidence to support the claim that the standard deviation of the girls' test scores is smaller than 16.1.

20) Test statistic: $X^2 = 9.2597$. Critical values: $X^2 = 6.262, 27.488$. Fail to reject $H_0$. There is not sufficient evidence to support the claim that heights of women aged 25 to 34 have a standard deviation different from 2.9 in.
CHAPTER 8 FORM B

Name: __________________________ Course Number: ______ Section Number: ______

Directions: Write your answers to the short-answer items in the spaces provided.
Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Explain how to determine if a hypothesis test is one-tailed or two-tailed and explain how you know where to shade the critical region. Give an example for each which includes the claim, the hypotheses, and the diagram with the critical region shaded.

Solve the problem.

2) Write the claim that is suggested by the given statement, then write a conclusion about the claim. Do not use symbolic expressions or formal procedures; use common sense.

Last year an appliance manufacturer received many complaints about the high rate of defects among its washing machines. Approximately 9% of the machines were defective in some way. This year the company tightened up its quality control procedures. The latest shipment of 250 washing machines contained 2 defectives.

Express the null hypothesis $H_0$ and the alternative hypothesis $H_1$ in symbolic form. Use the correct symbol ($\mu$, $p$, $\sigma$) for the indicated parameter.

3) A researcher claims that 62% of voters favor gun control.

A) $H_0$: $p \neq 0.62$  B) $H_0$: $p = 0.62$  C) $H_0$: $p = 0.62$  D) $H_0$: $p < 0.62$

$H_1$: $p = 0.62$  $H_1$: $p < 0.62$  $H_1$: $p \neq 0.62$  $H_1$: $p \geq 0.62$

Assume that the data has a normal distribution and the number of observations is greater than fifty. Find the critical $z$ value used to test a null hypothesis.

4) $\alpha = 0.09$ for a right-tailed test.

A) $\pm 1.96$  B) $+1.34$  C) $+1.96$  D) $\pm 1.34$
CHAPTER 8 FORM B

Find the value of the test statistic $z$ using $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$.

5) The claim is that the proportion of drowning deaths of children attributable to beaches is more than 0.25, and the sample statistics include $n = 615$ drowning deaths of children with 30% of them attributable to beaches.
   A) 2.86  B) -2.71  C) -2.86  D) 2.71

Use the given information to find the P-value.

6) The test statistic in a right-tailed test is $z = 0.52$.
   A) 0.5530  B) 0.1915  C) 0.1950  D) 0.3015

Formulate the indicated conclusion in nontechnical terms. Be sure to address the original claim.

7) A psychologist claims that more than 56 percent of the population suffers from professional problems due to extreme shyness. Assuming that a hypothesis test of the claim has been conducted and that the conclusion is failure to reject the null hypothesis, state the conclusion in nontechnical terms.
   A) There is sufficient evidence to support the claim that the true proportion is greater than 56 percent.
   B) There is sufficient evidence to support the claim that the true proportion is less than 56 percent.
   C) There is not sufficient evidence to support the claim that the true proportion is less than 56 percent.
   D) There is not sufficient evidence to support the claim that the true proportion is greater than 56 percent.

Assume that a hypothesis test of the given claim will be conducted. Identify the type I or type II error for the test.

8) The manufacturer of a refrigerator system for beer kegs produces refrigerators that are supposed to maintain a true mean temperature, $\mu$, of 42°F, ideal for a certain type of German pilsner. The owner of the brewery does not agree with the refrigerator manufacturer, and claims he can prove that the true mean temperature is incorrect. Identify the type I error for the test.
   A) The error of rejecting the claim that the mean temperature equals 42°F when it really does equal 42°F.
   B) The error of rejecting the claim that the mean temperature equals 42°F when it is really different from 42°F.
   C) The error of failing to reject the claim that the mean temperature equals 42°F when it is really different from 42°F.
CHAPTER 8 FORM B

Solve the problem.

9) True or False: In a hypothesis test regarding a population mean, the probability of a type II error, $\beta$, depends on the true value of the population mean.
A) True
B) False

Identify the null hypothesis, alternative hypothesis, test statistic, $P$-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

10) A nationwide study of American homeowners revealed that 64% have one or more lawn mowers. A lawn equipment manufacturer, located in Omaha, feels the estimate is too low for households in Omaha. Can the value 0.64 be rejected if a survey of 490 homes in Omaha yields 331 with one or more lawn mowers? Use $\alpha = 0.05$.

Find the $P$-value for the indicated hypothesis test.

11) Find the $P$-value for a test of the claim that more than 50% of the people following a particular diet will experience increased energy. Of 100 randomly selected subjects who followed the diet, 47 noticed an increase in their energy level.
A) 0.2743
B) 0.5486
C) 0.7257
D) 0.2257

Determine whether the given conditions justify testing a claim about a population mean $\mu$.

12) The sample size is $n = 18$, $\sigma$ is not known, and the original population is normally distributed.
A) Yes
B) No

Identify the null hypothesis, alternative hypothesis, test statistic, $P$-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

13) Various temperature measurements are recorded at different times for a particular city. The mean of 20°C is obtained for 40 temperatures on 40 different days. Assuming that $\sigma = 1.5$°C, test the claim that the population mean is 22°C. Use a 0.05 significance level.
CHAPTER 8 FORM B

Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, Student t distribution, or neither.

14) Claim: \( \mu = 73 \). Sample data: \( n = 24, \bar{x} = 104, s = 15.1 \). The sample data appear to come from a population with a distribution that is very far from normal, and \( \sigma \) is unknown.
   A) Student t  
   B) Neither  
   C) Normal

Assume that a simple random sample has been selected from a normally distributed population. Find the test statistic, \( P\)-value, critical value(s), and state the final conclusion.

15) Test the claim that for the population of history exams, the mean score is 80. Sample data are summarized as \( n = 16, \bar{x} = 84.5, \) and \( s = 11.2 \). Use a significance level of \( \alpha = 0.01 \).

Test the given claim using the traditional method of hypothesis testing. Assume that the sample has been randomly selected from a population with a normal distribution.

16) A manufacturer makes ball bearings that are supposed to have a mean weight of 30 g. A retailer suspects that the mean weight is actually less than 30 g. The mean weight for a random sample of 16 ball bearings is 28.8 g with a standard deviation of 3.8 g. At the 0.05 significance level, test the claim that the mean is less than 30 g.

17) A cereal company claims that the mean weight of the cereal in its packets is 14 oz. The weights (in ounces) of the cereal in a random sample of 8 of its cereal packets are listed below.
   14.6 13.8 14.1 13.7 14.0 14.4 13.6 14.2
   Test the claim at the 0.01 significance level.
CHAPTER 8 FORM B

Find the critical value or values of $\chi^2$ based on the given information.

18) $H_1: \sigma > 26.1$
   
   $n = 9$
   
   $\alpha = 0.01$
   
   A) 1.646   B) 21.666   C) 2.088   D) 20.090

Use the traditional method to test the given hypothesis. Assume that the population is normally distributed and that the sample has been randomly selected.

19) In one town, monthly incomes for men with college degrees are found to have a standard deviation of $\$650$. Use a 0.01 significance level to test the claim that for men without college degrees in that town, incomes have a higher standard deviation. A random sample of 22 men without college degrees resulted in incomes with a standard deviation of $\$927$.

20) For randomly selected adults, IQ scores are normally distributed with a standard deviation of 15. The scores of 14 randomly selected college students are listed below. Use a 0.10 significance level to test the claim that the standard deviation of IQ scores of college students is less than 15.

   115  128  107  109  116  124  135
   127  115  104  118  126  129  133
1) Examples will vary. Relational operators in the alternative hypotheses indicate whether the test is one-tailed or two-tailed. Strict inequalities determine one-tailed tests, whereas "not equal to" and "different from" determine two-tailed tests. The critical region begins at the CV for one-tailed tests and at both CVs for two-tailed tests. Shading begins at the CV/CVs and extends to the tails of the distribution.

2) The claim is that the defect rate has decreased and is now less than 9%. If the overall defect rate were still 9%, it would be extremely unlikely that a shipment of 250 washing machines would contain as few as 2 defectives. Therefore, the claim that the defect rate has decreased is probably correct.

3) C
4) B
5) A
6) D
7) D
8) A
9) A
10) \( H_0: \mu = 22; H_1: \mu \neq 22 \). Test statistic: \( z = -8.43 \).
    P-value: \( p = 0.0002 \) (Table A-2), \( p = 3.42 \times 10^{-17} \) (TI-83/84 Plus). Because the P-value is less than the significance level of \( \alpha = 0.05 \), we reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the population mean temperature is 22 degrees C.

11) C
12) A
13) \( H_0: \mu = 22; H_1: \mu \neq 22 \). Test statistic: \( t = -1.263 \).
    P-value: \( p = 0.1369 \) (STATDISK, TI-83/84 Plus, 0.10 < P-value < 0.20 (Table A-3), Critical value: \( t = \pm 2.497 \).
    Because the TS \( t < 2.497 \), we do not reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the mean score is 80.
14) B
15) \( \alpha = 0.01 \)
    Test statistic: \( t = 1.607 \)
    P-value: \( p = 0.1289 \) (STATDISK, TI-83/84 Plus), 0.10 < P-value < 0.20 (Table A-3), Critical value: \( t = \pm 2.497 \).
    Because the TS \( t < 2.497 \), we do not reject the null hypothesis. There is not sufficient evidence to support the claim that the mean is less than 30 g.

16) Test statistic: \( t = -1.753 \). Critical value: \( t = \pm 3.499 \). Fail to reject \( H_0 \). There is not sufficient evidence to support the claim that the mean weight is 14 ounces.

17) Test statistic: \( t = 0.408 \). Critical value: \( t = \pm 3.499 \). Fail to reject \( H_0: \mu = 14 \) ounces. There is not sufficient evidence to warrant rejection of the claim that the mean weight is 14 ounces.

18) D

19) Test statistic: \( X^2 = 42.712 \). Critical value: \( X^2 = 38.932 \). Reject \( H_0 \). There is sufficient evidence to support the claim that incomes of men without college degrees have a standard deviation greater than $650.

20) Test statistic: \( X^2 = 5.571 \). Critical value: \( X^2 = 7.042 \) (Table A-4), Critical value: 7.041 (STATDISK). Reject \( H_0 \). There is sufficient evidence to support the claim that IQ scores of college students have a standard deviation smaller than 15.
CHAPTER 8 FORM C

Name:_____________________________ Course Number:_________ Section Number:____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Under what conditions do you reject $H_0$? Discuss both the traditional and the $P$-value approach.

Solve the problem.

2) Write the claim that is suggested by the given statement, then write a conclusion about the claim. Do not use symbolic expressions or formal procedures: use common sense.

A person claims to have extra sensory powers. A card is drawn at random from a deck of cards and without looking at the card, the person is asked to identify the suit of the card. He correctly identifies the suit 28 times out of 100.

Express the null hypothesis $H_0$ and the alternative hypothesis $H_1$ in symbolic form. Use the correct symbol ($\mu$, $p$, $\sigma$) for the indicated parameter.

3) The principal of a middle school claims that test scores of the seventh-graders at her school vary less than the test scores of seventh-graders at a neighboring school, which have variation described by $\sigma = 14.7$.

A) $H_0$: $\sigma < 14.7$  B) $H_0$: $\sigma < 14.7$  C) $H_0$: $\sigma = 14.7$  D) $H_0$: $\sigma > 14.7$

$H_1$: $\sigma > 14.7$  $H_1$: $\sigma > 14.7$  $H_1$: $\sigma < 14.7$  $H_1$: $\sigma \leq 14.7$

Assume that the data has a normal distribution and the number of observations is greater than fifty. Find the critical z value used to test a null hypothesis.

4) $\alpha = 0.1$ for a two-tailed test.

A) $\pm 2.33$  B) $\pm 1.645$  C) $\pm 1.4805$  D) $\pm 2.052$
CHAPTER 8 FORM C

Find the value of the test statistic \( z \) using \( z = \frac{\hat{p} - p}{\sqrt{pq/n}} \).

5) The claim is that the proportion of accidental deaths of the elderly attributable to residential falls is more than 0.10, and the sample statistics include \( n = 800 \) deaths of the elderly with 15\% of them attributable to residential falls.

A) 3.96  
B) 4.71  
C) -3.96  
D) -4.71

Use the given information to find the P-value.

6) The test statistic in a two-tailed test is \( z = 1.95 \).

A) 0.3415  
B) 0.4423  
C) 0.0244  
D) 0.0512

Formulate the indicated conclusion in nontechnical terms. Be sure to address the original claim.

7) A cereal company claims that the mean weight of the cereal in its packets is at least 14 oz. Assuming that a hypothesis test of the claim has been conducted and that the conclusion is to reject the null hypothesis, state the conclusion in nontechnical terms.

A) There is not sufficient evidence to warrant rejection of the claim that the mean weight is less than 14 oz.

B) There is not sufficient evidence to warrant rejection of the claim that the mean weight is at least 14 oz.

C) There is sufficient evidence to warrant rejection of the claim that the mean weight is less than 14 oz.

D) There is sufficient evidence to warrant rejection of the claim that the mean weight is at least 14 oz.

Assume that a hypothesis test of the given claim will be conducted. Identify the type I or type II error for the test.

8) A researcher claims that the amounts of acetaminophen in a certain brand of cold tablets have a standard deviation different from the \( \sigma = 3.3 \) mg claimed by the manufacturer. Identify the type II error for the test.

A) The error of failing to reject the claim that the standard deviation is 3.3 mg when it is actually different from 3.3 mg.

B) The error of rejecting the claim that the standard deviation is more than 3.3 mg when it really is more than 3.3 mg.

C) The error of rejecting the claim that the standard deviation is 3.3 mg when it really is 3.3 mg.
CHAPTER 8 FORM C

Solve the problem.

9) In a hypothesis test, which of the following will cause a decrease in $\beta$, the probability of making a type II error?

A: Increasing $\alpha$ while keeping the sample size $n$, fixed
B: Increasing the sample size $n$, while keeping $\alpha$ fixed
C: Decreasing $\alpha$ while keeping the sample size $n$, fixed
D: Decreasing the sample size $n$, while keeping $\alpha$ fixed

A) A and D B) C and D C) A and B D) B and C

Identify the null hypothesis, alternative hypothesis, test statistic, $P$-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

10) In a clinical study of an allergy drug, 108 of the 202 subjects reported experiencing significant relief from their symptoms. At the 0.01 significance level, test the claim that more than half of all those using the drug experience relief.

Find the $P$-value for the indicated hypothesis test.

11) An airline claims that the no-show rate for passengers booked on its flights is less than 6%. Of 380 randomly selected reservations, 18 were no-shows. Find the $P$-value for a test of the airline’s claim.

A) 0.1492 B) 0.1230 C) 0.3508 D) 0.0746

Determine whether the given conditions justify testing a claim about a population mean $\mu$.

12) The sample size is $n = 44$, $\sigma = 13.4$, and the original population is not normally distributed.

A) No B) Yes

Identify the null hypothesis, alternative hypothesis, test statistic, $P$-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

13) The health of employees is monitored by periodically weighing them in. A sample of 54 employees has a mean weight of 183.9 lb. Assuming that $\sigma$ is known to be 121.2 lb, use a 0.10 significance level to test the claim that the population mean of all such employees weights is less than 200 lb.
CHAPTER 8 FORM C

Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, Student t distribution, or neither.

14) Claim: \( \mu = 973 \). Sample data: \( n = 20, \bar{x} = 958, s = 29 \). The sample data appear to come from a normally distributed population with \( \sigma = 28 \).

A) Neither  B) Student t  C) Normal

Assume that a simple random sample has been selected from a normally distributed population. Find the test statistic, \( P \)-value, critical value(s), and state the final conclusion.

15) Test the claim that for the adult population of one town, the mean annual salary is given by \( \mu = $30,000 \). Sample data are summarized as \( n = 17, \bar{x} = $22,298 \), and \( s = $14,200 \). Use a significance level of \( \alpha = 0.05 \).

Test the given claim using the traditional method of hypothesis testing. Assume that the sample has been randomly selected from a population with a normal distribution.

16) A public bus company official claims that the mean waiting time for bus number 14 during peak hours is less than 10 minutes. Karen took bus number 14 during peak hours on 18 different occasions. Her mean waiting time was 7.2 minutes with a standard deviation of 1.6 minutes. At the 0.01 significance level, test the claim that the mean is less than 10 minutes.

17) A light-bulb manufacturer advertises that the average life for its light bulbs 900 hours. A random sample of 15 of its light bulbs resulted in the following lives in hours.

\[
\begin{array}{cccccccccc}
995 & 590 & 510 & 539 & 739 & 917 & 571 & 555 \\
916 & 728 & 664 & 693 & 708 & 887 & 849 \\
\end{array}
\]

At the 10% significance level, do the data provide evidence that the mean life for the company’s light bulbs differs from the advertised mean?
CHAPTER 8 FORM C

Find the critical value or values of $x^2$ based on the given information.

18) $H_1: \sigma < 0.14$
   \[ n = 23 \]
   \[ \alpha = 0.10 \]
   A) 30.813  B) -30.813  C) 14.848  D) 14.042

Use the traditional method to test the given hypothesis. Assume that the population is normally distributed and that the sample has been randomly selected.

19) A manufacturer uses a new production method to produce steel rods. A random sample of 17 steel rods resulted in lengths with a standard deviation of 2.1 cm. At the 0.10 significance level, test the claim that the new production method has lengths with a standard deviation different from 3.5 cm, which was the standard deviation for the old method.

20) With individual lines at the checkouts, a store manager finds that the standard deviation for the waiting times on Monday mornings is 5.7 minutes. After switching to a single waiting line, he finds that for a random sample of 29 customers, the waiting times have a standard deviation of 4.9 minutes. Use a 0.025 significance level to test the claim that with a single line, waiting times vary less than with individual lines.
Answer Key
Testname: CHAPTER 8 FORM C

1) For the traditional method, the test-statistic is the critical region. For the P-value method, the P-value is less than or equal to the significance level $\alpha$ and the test statistic is on the appropriate side in a one-tailed test.

2) The claim is that the person is using his extra sensory powers to determine the suit of the card, and that he correctly determines the suit more often than he would if he were guessing randomly. Even if he were just guessing randomly, he would have a reasonable chance of being correct 28 times out of a hundred; this is not improbable, since there are four suits. Therefore, identifying the suit correctly 28 times out of 100 does not constitute strong evidence in favor of his claim.

3) C
4) B
5) B
6) D
7) D
8) A
9) C
10) $H_0$: $p = 0.5$. $H_1$: $p > 0.5$. Test statistic: $z = 0.99$. P-value: $p = 0.1611$ (Table A-2), $p = 0.1623$ (STATDISK, TI-83/84 Plus).
    Critical value: $z = 2.33$. Fail to reject null hypothesis. There is not sufficient evidence to support the claim that more than half of all those using the drug experience relief.

11) A
12) B
13) $H_0$: $\mu = 200$. $H_1$: $\mu < 200$. Test statistic: $z = -0.98$. P-value: $p = 0.1635$ (Table A-2), $p = 0.1645$ (STATDISK, TI-83/84 Plus). Fail to reject $H_0$. There is not sufficient evidence to support the claim that the mean is less than 200 pounds.

14) C
15) $\alpha = 0.05$
    Test statistic: $t = -2.236$
    Critical values: $t = \pm 2.120$
    P-value: $p = 0.0399$ (STATDISK, TI-83/84 Plus), $0.02 < P-value < 0.05$ (Table A-3).
    Because $TS t < -2.120$, we reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that $\mu = 30,000$.

16) Test statistic: $t = -4.25$. Critical value: $t = -2.567$. Reject $H_0$. There is sufficient evidence to support the claim that the mean is less than 10 minutes.
17) Test statistic: $t = -4.342$. Critical values: $t = \pm 1.761$. Reject $H_0$: $\mu = 900$ hours. There is sufficient evidence to support the claim that the true mean life differs from the advertised mean.
18) D
19) Test statistic: $X^2 = 5.760$. Critical values: $X^2 = 7.962$, 26.296. Reject $H_0$. There is sufficient evidence to support the claim that the standard deviation is different from 3.5.
20) Test statistic: $X^2 = 20.692$. Critical value: $X^2 = 15.308$. Fail to reject $H_0$. There is not sufficient evidence to support the claim that with a single line waiting times have a smaller standard deviation.
CHAPTER 9 FORM A

Name:_____________________________ Course Number:_______ Section Number:_____ 

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) What is the effect on the P-value when a test is changed from a two-tailed hypothesis with = and ≠ to a one-tailed hypothesis such as ≥ and <?

Find the number of successes x suggested by the given statement.

2) Among 1410 randomly selected car drivers in one city, 7.52% said that they had been involved in an accident during the past year.

   A) 107   B) 104   C) 106   D) 105

From the sample statistics, find the value of \( \bar{p} \) used to test the hypothesis that the population proportions are equal.

3) \( n_1 = 100 \) \( n_2 = 100 \)
   \( x_1 = 34 \) \( x_2 = 30 \)

   A) 0.224   B) 0.288   C) 0.320   D) 0.352

Compute the test statistic used to test the null hypothesis that \( p_1 = p_2 \).

4) A report on the nightly news broadcast stated that 15 out of 140 households with pet dogs were burglarized and 20 out of 180 without pet dogs were burglarized.

   A) -0.045   B) -0.113   C) 0.000   D) -0.192

Find the appropriate p-value to test the null hypothesis, \( H_0: p_1 = p_2 \), using a significance level of 0.05.

5) \( n_1 = 50 \) \( n_2 = 75 \)
   \( x_1 = 20 \) \( x_2 = 15 \)

   A) .0001   B) .1201   C) .0146   D) .0032

152
CHAPTER 9 FORM A

Use the traditional method to test the given hypothesis. Assume that the samples are independent and that they have been randomly selected.

6) Use the given sample data to test the claim that $p_1 > p_2$. Use a significance level of 0.01.

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_1 = 85$</td>
<td>$n_2 = 90$</td>
</tr>
<tr>
<td>$x_1 = 38$</td>
<td>$x_2 = 23$</td>
</tr>
</tbody>
</table>

7) A researcher finds that of 1,000 people who said that they attend a religious service at least once a week, 31 stopped to help a person with car trouble. Of 1,200 people interviewed who had not attended a religious service at least once a month, 22 stopped to help a person with car trouble. At the 0.05 significance level, test the claim that the two proportions are equal.

Construct the indicated confidence interval for the difference between population proportions $p_1 - p_2$. Assume that the samples are independent and that they have been randomly selected.

8) $x_1 = 10$, $n_1 = 42$ and $x_2 = 27$, $n_2 = 47$; Construct a 90% confidence interval for the difference between population proportions $p_1 - p_2$.

A) $-0.497 < p_1 - p_2 < -0.176$  
B) $0.076 < p_1 - p_2 < 0.400$  
C) $0.047 < p_1 - p_2 < 0.429$  
D) $0.429 < p_1 - p_2 < 0.045$

Determine whether the samples are independent or consist of matched pairs.

9) The accuracy of verbal responses is tested in an experiment in which individuals report their heights and then are measured. The data consist of the reported height and measured height for each individual.

A) Matched pairs  
B) Independent samples
CHAPTER 9 FORM A

Test the indicated claim about the means of two populations. Assume that the two samples are independent and that they have been randomly selected.

10) Two types of flares are tested for their burning times (in minutes) and sample results are given below.

<table>
<thead>
<tr>
<th></th>
<th>Brand X</th>
<th>Brand Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>$\bar{x}$</td>
<td>19.4</td>
<td>15.1</td>
</tr>
<tr>
<td>s</td>
<td>1.4</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Refer to the sample data to test the claim that the two populations have equal means. Use a 0.05 significance level.

Construct the indicated confidence interval for the difference between the two population means. Assume that the two samples are independent and that they have been randomly selected.

11) A researcher wishes to determine whether people with high blood pressure can reduce their blood pressure by following a particular diet. Use the sample data below to construct a 99% confidence interval for $\mu_1 - \mu_2$ where $\mu_1$ and $\mu_2$ represent the mean for the treatment group and the control group respectively.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_1 = 85$</td>
<td>$n_2 = 75$</td>
</tr>
<tr>
<td>$\bar{x}_1 = 189.1$</td>
<td>$\bar{x}_2 = 203.7$</td>
</tr>
<tr>
<td>$s_1 = 38.7$</td>
<td>$s_2 = 39.2$</td>
</tr>
</tbody>
</table>

A) $-30.9 < \mu_1 - \mu_2 < 1.7$  
B) $-26.7 < \mu_1 - \mu_2 < -2.5$

C) $-1.3 < \mu_1 - \mu_2 < 30.5$  
D) $-29.0 < \mu_1 - \mu_2 < -0.2$

Find $s_d$.

12) Consider the set of differences between two dependent sets: 84, 85, 83, 63, 61, 100, 98. Round to the nearest tenth.

A) 15.3  
B) 16.2  
C) 13.1  
D) 15.7

The two data sets are dependent. Find $\bar{d}$ to the nearest tenth.

13) | A | 60 61 64 63 51  |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>25 23 29 25 22</td>
</tr>
</tbody>
</table>

A) 43.8  
B) 35.0  
C) 21.0  
D) 45.5
CHAPTER 9 FORM A

Use the computer display to solve the problem.

14) When testing for a difference between the means of a treatment group and a placebo group, the computer display below is obtained. Using a 0.04 significance level, is there sufficient evidence to support the claim that the treatment group (variable 1) comes from a population with a mean that is different from the mean for the placebo population? Explain.

<table>
<thead>
<tr>
<th>t-Test: Two Sample for Means</th>
<th>Variable 1</th>
<th>Variable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Mean</td>
<td>171.6392</td>
<td>168.7718</td>
</tr>
<tr>
<td>3 Known Variance</td>
<td>47.51672</td>
<td>41.08293</td>
</tr>
<tr>
<td>4 Observations</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>5 Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6 t</td>
<td>2.154057</td>
<td></td>
</tr>
<tr>
<td>7 P(T&gt;=t) one-tail</td>
<td>0.0158</td>
<td></td>
</tr>
<tr>
<td>8 T Critical one-tail</td>
<td>1.644853</td>
<td></td>
</tr>
<tr>
<td>9 P(T&gt;=t) two-tail</td>
<td>0.0316</td>
<td></td>
</tr>
<tr>
<td>10 t Critical two-tail</td>
<td>1.939961</td>
<td></td>
</tr>
</tbody>
</table>

Assume that you want to test the claim that the paired sample data come from a population for which the mean difference is $\mu_d = 0$. Compute the value of the t test statistic.

15) $\begin{array}{cccccccc}
 x & 28 & 35 & 25 & 25 & 32 & 30 & 28 & 34 \\
 y & 26 & 31 & 25 & 33 & 35 & 28 & 33 & 33 \\
\end{array}$

A) $t = 0.690$  B) $t = -1.480$  C) $t = -0.523$  D) $t = -0.185$
CHAPTER 9 FORM A

Determine the decision criterion for rejecting the null hypothesis in the given hypothesis test; i.e., describe the values of the test statistic that would result in rejection of the null hypothesis.

16) A farmer has decided to use a new additive to grow his crops. He divided his farm into 10 plots and kept records of the corn yield (in bushels) before and after using the additive. The results are shown below.

<table>
<thead>
<tr>
<th>Plot</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>After</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

You wish the test the following hypothesis at the 5 percent level of significance.

\[ H_0: \mu_D = 0 \text{ against } H_1: \mu_D > 0. \]

What decision rule would you use?

A) Reject \( H_0 \) if test statistic is less than 1.833.

B) Reject \( H_0 \) if test statistic is greater than 1.833 or less than 1.833.

C) Reject \( H_0 \) if test statistic is greater than 1.833.

D) Reject \( H_0 \) if test statistic is greater than -1.833.

Use the traditional method of hypothesis testing to test the given claim about the means of two populations. Assume that two dependent samples have been randomly selected from normally distributed populations.

17) A test of abstract reasoning is given to a random sample of students before and after they completed a formal logic course. The results are given below. At the 0.05 significance level, test the claim that the mean score is not affected by the course.

<table>
<thead>
<tr>
<th>Before</th>
<th>74</th>
<th>83</th>
<th>75</th>
<th>88</th>
<th>84</th>
<th>63</th>
<th>84</th>
<th>91</th>
<th>77</th>
</tr>
</thead>
<tbody>
<tr>
<td>After</td>
<td>73</td>
<td>77</td>
<td>70</td>
<td>77</td>
<td>74</td>
<td>67</td>
<td>95</td>
<td>83</td>
<td>84</td>
</tr>
</tbody>
</table>

Construct a confidence interval for \( \mu_d \), the mean of the differences \( d \) for the population of paired data. Assume that the population of paired differences is normally distributed.

18) If \( \bar{d} = 3.125, S_d = 2.911, \) and \( n = 8 \), determine a 99 percent confidence interval for \( \mu_d \).

A) \( 1.851 < \mu_d < 4.399 \)

B) \( -0.476 < \mu_d < 6.726 \)

C) \( 1.851 < \mu_d < 6.726 \)

D) \( -0.360 < \mu_d < 4.399 \)
CHAPTER 9 FORM A

Test the indicated claim about the variances or standard deviations of two populations. Assume that the populations are normally distributed. Assume that the two samples are independent and that they have been randomly selected.

19) Two types of flares are tested for their burning times (in minutes) and sample results are given below. Use a 0.05 significance level to test the claim that the two brands have equal variances.

<table>
<thead>
<tr>
<th>Brand X</th>
<th>Brand Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 35$</td>
<td>$n = 40$</td>
</tr>
<tr>
<td>$\bar{x} = 19.4$</td>
<td>$\bar{x} = 15.1$</td>
</tr>
<tr>
<td>$s = 1.4$</td>
<td>$s = 0.8$</td>
</tr>
</tbody>
</table>

Solve the problem.

20) A test for homogeneity of variance is conducted at the 5% level of significance. Sample sizes are $n_1 = 31$ and $n_2 = 31$. The test statistic is $F = 6.4071$. What do you know about the variance of the populations from which the samples were taken?
1) The P-value is cut in half.
2) C
3) C
4) B
5) C
6) $H_0: p_1 = p_2$. $H_1: p_1 > p_2$.
   Test statistic: $z = 2.66$. Critical value: $z = 2.33$.
   Reject the null hypothesis. There is sufficient evidence to support the claim that $p_1 > p_2$.
7) $H_0: p_1 = p_2$. $H_1: p_1 \neq p_2$.
   Test statistic: $z = 1.93$. Critical values: $z = \pm 1.96$.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the two proportions are equal.
8) A
9) A
10) $H_0: \mu_1 = \mu_2$. $H_1: \mu_1 \neq \mu_2$.
    Test statistic $t = 16.025$. Critical values: $t = \pm 2.032$ (Table A-3).
    [Optional: Critical values: $t = \pm 2.006217$ (STATDISK) with $df = 52.4700$ (STATDISK, TI-83/84 Plus). P-value: $p = 0.0000$ (STATDISK), $p \approx 7.350731E-22$ (TI-83/84 Plus).]
    Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the two populations have equal means.
11) A
12) A
13) B
14) Yes, the P-value for a two-tail test is 0.0316, which is smaller than the significance level of 0.04.
    There is sufficient evidence to support the claim that the two population means are different.
15) C
16) C
17) Test statistic $t = 2.366$. Critical values: $t = \pm 2.262$.
    Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the mean is not affected by the course.
18) B
19) $H_0: \sigma_1^2 = \sigma_2^2$. $H_1: \sigma_1^2 \neq \sigma_2^2$.
    Test statistic: $F = 3.0625$. Critical value: $1.8752 < F < 2.0739$ (Table A-5).
    [Optional: Critical value: $F = 1.921619$ (STATDISK).]
    Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the two brands have equal variances.
20) The critical F value for a two-tailed test of homogeneity of variance is $F = 2.0739$, $df = 30, 30$ and with 0.025 in the right tail. Since the TS F is in the critical region, the null hypothesis of equal variances is rejected. We conclude that the populations have significantly different variances.
CHAPTER 9 FORM B

Name: ___________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) The test statistic for testing hypothesis about two variances is \( F = \frac{s_1^2}{s_2^2} \) where \( s_1^2 > s_2^2 \).

Describe the numeric possibilities for this test statistic. Explain the circumstances under which the conclusion would be either that the variances are equal or that the variances are not equal.

Find the number of successes \( x \) suggested by the given statement.

2) Among 730 people selected randomly from among the eligible voters in one city, 64% were homeowners

A) 471  B) 472  C) 467  D) 463

From the sample statistics, find the value of \( \bar{p} \) used to test the hypothesis that the population proportions are equal.

3) \( n_1 = 216 \quad n_2 = 186 \)
\( x_1 = 76 \quad x_2 = 99 \)

A) 0.218  B) 0.305  C) 0.392  D) 0.435

Compute the test statistic used to test the null hypothesis that \( p_1 = p_2 \).

4) Information about movie ticket sales was printed in a movie magazine. Out of fifty PG-rated movies, 44% had ticket sales in excess of $3,000,000. Out of thirty-five R-rated movies, 22% grossed over $3,000,000.

A) 6.491  B) 2.089  C) 3.350  D) 4.188

159
CHAPTER 9 FORM B

Find the appropriate p-value to test the null hypothesis, \( H_0: p_1 = p_2 \), using a significance level of 0.05.

5) \( n_1 = 200 \quad n_2 = 100 \)
\( x_1 = 11 \quad x_2 = 8 \)

A) .1011 B) .0012 C) .0201 D) .4010

Use the traditional method to test the given hypothesis. Assume that the samples are independent and that they have been randomly selected

6) Use the given sample data to test the claim that \( p_1 < p_2 \). Use a significance level of 0.10.

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 = 462 )</td>
<td>( n_2 = 380 )</td>
</tr>
<tr>
<td>( x_1 = 84 )</td>
<td>( x_2 = 95 )</td>
</tr>
</tbody>
</table>

7) In a random sample of 360 women, 65% favored stricter gun control laws. In a random sample of 220 men, 60% favored stricter gun control laws. Test the claim that the proportion of women favoring stricter gun control is higher than the proportion of men favoring stricter gun control. Use a significance level of 0.05.

Construct the indicated confidence interval for the difference between population proportions \( p_1 - p_2 \). Assume that the samples are independent and that they have been randomly selected.

8) \( x_1 = 36, n_1 = 80 \) and \( x_2 = 44, n_2 = 85 \); Construct a 95% confidence interval for the difference between population proportions \( p_1 - p_2 \).

A) \( 0.269 < p_1 - p_2 < 0.631 \) B) \( -0.220 < p_1 - p_2 < 0.085 \)
C) \( 0.298 < p_1 - p_2 < 0.602 \) D) \( -0.249 < p_1 - p_2 < 0.631 \)
CHAPTER 9 FORM B

Determine whether the samples are independent or consist of matched pairs.

9) The effect of caffeine as an ingredient is tested with a sample of regular soda and another sample with decaffeinated soda.
   A) Independent samples  B) Matched pairs

Test the indicated claim about the means of two populations. Assume that the two samples are independent and that they have been randomly selected.

10) A researcher wishes to determine whether people with high blood pressure can reduce their blood pressure by following a particular diet. Use the sample data below to test the claim that the treatment population mean $\mu_1$ is smaller than the control population mean $\mu_2$. Test the claim using a significance level of 0.01.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_1 = 101$</td>
<td>$n_2 = 105$</td>
</tr>
<tr>
<td>$\bar{x}_1 = 120.5$</td>
<td>$\bar{x}_2 = 149.3$</td>
</tr>
<tr>
<td>$s_1 = 17.4$</td>
<td>$s_2 = 30.2$</td>
</tr>
</tbody>
</table>

Construct the indicated confidence interval for the difference between the two population means. Assume that the two samples are independent and that they have been randomly selected.

11) Independent samples from two different populations yield the following data. $\bar{x}_1 = 200$, $\bar{x}_2 = 963$, $s_1 = 19$, $s_2 = 86$. The sample size is 433 for both samples. Find the 90 percent confidence interval for $\mu_1 - \mu_2$.

   A) $-770 < \mu_1 - \mu_2 < -756$  B) $-763 < \mu_1 - \mu_2 < -763$
   C) $-768 < \mu_1 - \mu_2 < -758$  D) $-777 < \mu_1 - \mu_2 < -749$
CHAPTER 9 FORM B

Use the computer display to solve the problem.

12) When testing for a difference between the means of a treatment group and a placebo group, the computer display below is obtained. Using a 0.01 significance level, is there sufficient evidence to support the claim that the treatment group (variable 1) comes from a population with a mean that is greater than the mean for the placebo population? Explain.

<table>
<thead>
<tr>
<th>t-Test: Two Sample for Means</th>
<th>Variable 1</th>
<th>Variable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Mean</td>
<td>171.6392</td>
<td>168.7718</td>
</tr>
<tr>
<td>3 Known Variance</td>
<td>47.51672</td>
<td>41.08293</td>
</tr>
<tr>
<td>4 Observations</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>5 Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6 t</td>
<td>2.154057</td>
<td></td>
</tr>
<tr>
<td>7 P(T≥t) one-tail</td>
<td>0.0158</td>
<td></td>
</tr>
<tr>
<td>8 T Critical one-tail</td>
<td>1.644853</td>
<td></td>
</tr>
<tr>
<td>9 P(T≥t) two-tail</td>
<td>0.0316</td>
<td></td>
</tr>
<tr>
<td>10 t Critical two-tail</td>
<td>1.959961</td>
<td></td>
</tr>
</tbody>
</table>

The two data sets are dependent. Find \( \bar{d} \) to the nearest tenth.

13) \[
\begin{array}{c|cccccccc}
X & 223 & 196 & 220 & 182 & 278 & 298 & 302 \\
Y & 205 & 140 & 195 & 153 & 235 & 247 & 284 \\
\end{array}
\]
A) 205.8 \hspace{1cm} B) 20.6 \hspace{1cm} C) 34.3 \hspace{1cm} D) 44.6

Find \( s_d \).

14) The differences between two sets of dependent data are -8, -9, -6, -7. Round to the nearest tenth.
A) 2.6 \hspace{1cm} B) 1.7 \hspace{1cm} C) 1.3 \hspace{1cm} D) 1.0

Assume that you want to test the claim that the paired sample data come from a population for which the mean difference is \( \mu_d = 0 \). Compute the value of the t test statistic.

15) \[
\begin{array}{c|cccc}
x & 7.4 & 5.9 & 5.1 & 10.2 & 5.8 & 10.2 & 10.5 & 8.4 \\
y & 5.6 & 5.6 & 6.2 & 5.5 & 6.2 & 4.9 & 6.6 & 6.7 \\
\end{array}
\]
A) \( t = 2.391 \) \hspace{1cm} B) \( t = 0.998 \) \hspace{1cm} C) \( t = 6.792 \) \hspace{1cm} D) \( t = 0.845 \)
CHAPTER 9 FORM B

Determine the decision criterion for rejecting the null hypothesis in the given hypothesis test; i.e., describe the values of the test statistic that would result in rejection of the null hypothesis.

16) A farmer has decided to use a new additive to grow his crops. He divided his farm into 10 plots and kept records of the corn yield (in bushels) before and after using the additive. The results are shown below.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>11</td>
</tr>
</tbody>
</table>

You wish to test the following hypothesis at the 1 percent level of significance.

\[ H_0: \mu_D = 0 \text{ against } H_1: \mu_D \neq 0. \]

What decision rule would you use?

A) Reject \( H_0 \) if test statistic is greater than 3.250.

B) Reject \( H_0 \) if test statistic is greater than –3.250 or less than 3.250.

C) Reject \( H_0 \) if test statistic is less than –3.250.

D) Reject \( H_0 \) if test statistic is less than –3.250 or greater than 3.250.

Use the traditional method of hypothesis testing to test the given claim about the means of two populations. Assume that two dependent samples have been randomly selected from normally distributed populations.

17) A coach uses a new technique to train gymnasts. 7 gymnasts were randomly selected and their competition scores were recorded before and after the training. The results are shown below.

<table>
<thead>
<tr>
<th>Subject</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>9.7</td>
<td>9.6</td>
<td>9.6</td>
<td>9.6</td>
<td>9.7</td>
<td>9.4</td>
<td></td>
</tr>
<tr>
<td>After</td>
<td>9.8</td>
<td>9.8</td>
<td>9.6</td>
<td>9.5</td>
<td>9.7</td>
<td>10</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Using a 0.01 level of significance, test the claim that the training technique is effective in raising the gymnasts’ scores.

Construct a confidence interval for \( \mu_D \), the mean of the differences \( d \) for the population of paired data. Assume that the population of paired differences is normally distributed.

18) Using the sample paired data below, construct a 90% confidence interval for the population mean of all differences \( x - y \).

<table>
<thead>
<tr>
<th>x</th>
<th>5.1</th>
<th>6.5</th>
<th>6.2</th>
<th>4.8</th>
<th>6.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>4.8</td>
<td>5.3</td>
<td>5.8</td>
<td>5.5</td>
<td>3.9</td>
</tr>
</tbody>
</table>

A) \( 0.22 < \mu_D < 7.48 \)  
B) \( -0.31 < \mu_D < 1.71 \)  
C) \( -0.37 < \mu_D < 1.77 \)  
D) \( -0.07 < \mu_D < 1.47 \)
CHAPTER 9 FORM B

Test the indicated claim about the variances or standard deviations of two populations. Assume that the populations are normally distributed. Assume that the two samples are independent and that they have been randomly selected.

19) Test the claim that populations A and B have different variances. Use a significance level of 0.10.

<table>
<thead>
<tr>
<th>Sample A</th>
<th>Sample B</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 28</td>
<td>n = 41</td>
</tr>
<tr>
<td>( \bar{x}_1 = 19.2 )</td>
<td>( \bar{x}_2 = 23.7 )</td>
</tr>
<tr>
<td>s = 5.38</td>
<td>s = 5.89</td>
</tr>
</tbody>
</table>

Solve the problem.

20) A test of homogeneity of variance is conducted at the 5% level of significance. Sample sizes are \( n_1 = 250 \) and \( n_2 = 275 \). The test statistic is 5.8231. What do you know about the variance of the populations from which the samples were taken?
Answer Key
Testname: CHAPTER 9 FORM B

1) The value for the test statistic F will be 1 or greater. If the value is reasonably close to 1, the conclusion is that the two variances are equal. If the value is significantly greater than 1, the conclusion is that the two variances are not equal.

2) C
3) D
4) B
5) D

6) \( H_0: \sigma_1^2 = \sigma_2^2 \quad H_1: \sigma_1^2 \neq \sigma_2^2 \).
   Test statistic: \( F = 1.20 \). Critical value: \( F = 1.84 \).
   Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that \( \sigma_1^2 \neq \sigma_2^2 \).

7) \( H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 > \mu_2 \).
   Test statistic: \( t = -3.143 \). Critical value: \( t = -2.364 \) (Table A-3).
   [Optional: Critical value: \( t = -2.348828 \) (STATDISK) with df = 167.43 (STATDISK, TI-83/84 Plus). P-value: \( p = 0.0000 \) (STATDISK, TI-83/84 Plus).]
   Reject the null hypothesis. There is sufficient evidence to support the claim that the treatment population mean \( \mu_1 \) is smaller than the control population \( \mu_2 \). On average, the diet appears effective in lowering blood pressure.

8) B
9) A
10) \( H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 < \mu_2 \).
    Test statistic t = -0.880. Critical value: \( t = -2.364 \) (Table A-3).
    [Optional: Critical value: \( t = -2.348828 \) (STATDISK) with df = 167.43 (STATDISK, TI-83/84 Plus). P-value: \( p = 0.0000 \) (STATDISK, TI-83/84 Plus).]
    Reject the null hypothesis. There is sufficient evidence to support the claim that the treatment group mean is greater than the mean for the placebo group.

11) A
12) No, the p-value for a one-tail test is 0.0158, which is larger than the significance level of 0.01. There is not sufficient evidence to support the claim that the mean for the treatment group is greater than the mean for the placebo group.

13) C
14) C
15) A
16) D

17) Test statistic \( t = -0.880 \). Critical value: \( t = -3.143 \).
    Fail to reject \( H_0: \mu_d = 0 \). There is not sufficient evidence to support the claim that the technique is effective in raising the gymnasts’ scores.

18) C

19) \( H_0: \sigma^2_1 = \sigma^2_2 \quad H_1: \sigma^2_1 \neq \sigma^2_2 \).
    Test statistic: \( F = 1.20 \). Critical value: \( F = 1.84 \).
    Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that populations A and B have different variances.

20) The critical F value for a two-tailed test of homogeneity of variance is \( F = 1.4327 \), df = 120, 120 and with 0.025 in the right tail. Since the test F is in the critical region, the null hypothesis of equal variances is rejected. We conclude that the populations have significantly different variances.
CHAPTER 9 FORM C

Name:_____________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Compare the technique for decision making about populations using the hypothesis test method and the confidence interval method.

Find the number of successes x suggested by the given statement.

2) Among 690 adults selected randomly from among the residents of one town, 28.3% said that they favor stronger gun-control laws.

A) 195   B) 193   C) 194   D) 196

From the sample statistics, find the value of $\bar{p}$ used to test the hypothesis that the population proportions are equal.

3) $n_1 = 541$   $n_2 = 1927$
   $x_1 = 160$   $x_2 = 689$

A) 0.587   B) 0.344   C) 0.275   D) 0.688

Compute the test statistic used to test the null hypothesis that $p_1 = p_2$.

4) In a vote on the Clean Water bill, 41% of the 205 Democrats voted for the bill while 42% of the 230 Republicans voted for it.

A) -0.232   B) -0.127   C) -0.179   D) -0.253

Find the appropriate p-value to test the null hypothesis, $H_0: p_1 = p_2$, using a significance level of 0.05.

5) $n_1 = 100$   $n_2 = 100$
   $x_1 = 38$   $x_2 = 40$

A) .0412   B) .2130   C) .7718   D) .1610
Use the traditional method to test the given hypothesis. Assume that the samples are independent and that they have been randomly selected

6) A marketing survey involves product recognition in New York and California. Of 558 New Yorkers surveyed, 193 knew the product while 196 out of 614 Californians knew the product. At the 0.05 significance level, test the claim that the recognition rates are the same in both states.

7) 7 of 8,500 people vaccinated against a certain disease later developed the disease. 18 of 10,000 people vaccinated with a placebo later developed the disease. Test the claim that the vaccine is effective in lowering the incidence of the disease. Use a significance level of 0.02.

Construct the indicated confidence interval for the difference between population proportions \( p_1 - p_2 \). Assume that the samples are independent and that they have been randomly selected.

8) \( x_1 = 62, n_1 = 124 \) and \( x_2 = 65, n_2 = 125 \); Construct a 98% confidence interval for the difference between population proportions \( p_1 - p_2 \).

A) \( 0.352 < p_1 - p_2 < 0.648 \) \hspace{1cm} B) \( 0.376 < p_1 - p_2 < 0.624 \)

C) \( -0.168 < p_1 - p_2 < 0.128 \) \hspace{1cm} D) \( -0.144 < p_1 - p_2 < 0.624 \)

Determine whether the samples are independent or consist of matched pairs.

9) The effectiveness of a new headache medicine is tested by measuring the amount of time before the headache is cured for patients who use the medicine and another group of patients who use a placebo drug.

A) Independent samples \hspace{1cm} B) Matched pairs
CHAPTER 9 FORM C

Test the indicated claim about the means of two populations. Assume that the two samples are independent and that they have been randomly selected.

10) A researcher wishes to determine whether people with high blood pressure can reduce their blood pressure by following a particular diet. Use the sample data below to test the claim that the treatment population mean \( \mu_1 \) is smaller than the control population mean \( \mu_2 \). Test the claim using a significance level of 0.01.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 = 85 )</td>
<td>( n_2 = 75 )</td>
</tr>
<tr>
<td>( \bar{x}_1 = 189.1 )</td>
<td>( \bar{x}_2 = 203.7 )</td>
</tr>
<tr>
<td>( s_1 = 38.7 )</td>
<td>( s_2 = 39.2 )</td>
</tr>
</tbody>
</table>

Construct the indicated confidence interval for the difference between the two population means. Assume that the two samples are independent and that they have been randomly selected.

11) Two types of flares are tested for their burning times (in minutes) and sample results are given below.

<table>
<thead>
<tr>
<th>Brand X</th>
<th>Brand Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = 35 )</td>
<td>( n = 40 )</td>
</tr>
<tr>
<td>( \bar{x} = 19.4 )</td>
<td>( \bar{x} = 15.1 )</td>
</tr>
<tr>
<td>( s = 1.4 )</td>
<td>( s = 0.8 )</td>
</tr>
</tbody>
</table>

Construct a 95% confidence interval for the differences \( \mu_X - \mu_Y \) based on the sample data.

A) \( 3.2 < \mu_X - \mu_Y < 5.4 \)  
B) \( 3.8 < \mu_X - \mu_Y < 4.8 \)

C) \( 3.5 < \mu_X - \mu_Y < 5.1 \)  
D) \( 3.6 < \mu_X - \mu_Y < 5.0 \)
CHAPTER 9 FORM C

Use the computer display to solve the problem.

12) When testing for a difference between the means of a treatment group and a placebo group, the computer display below is obtained. Using a 0.05 significance level, is there sufficient evidence to support the claim that the treatment group (variable 1) comes from a population with a mean that is different from the mean for the placebo population? Explain.

<table>
<thead>
<tr>
<th></th>
<th>t-Test: Two Sample for Means</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Variable 1</td>
<td>Variable 2</td>
</tr>
<tr>
<td>2</td>
<td>Mean</td>
<td>65.10738</td>
</tr>
<tr>
<td>3</td>
<td>Known Variance</td>
<td>8.102938</td>
</tr>
<tr>
<td>4</td>
<td>Observations</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>Hypothesized Mean Difference</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>t</td>
<td>-1.773417</td>
</tr>
<tr>
<td>7</td>
<td>P(T&lt;=t) one-tail</td>
<td>0.0384</td>
</tr>
<tr>
<td>8</td>
<td>T Critical one-tail</td>
<td>1.644853</td>
</tr>
<tr>
<td>9</td>
<td>P(T&lt;=t) two-tail</td>
<td>0.0768</td>
</tr>
<tr>
<td>10</td>
<td>t Critical two-tail</td>
<td>1.959961</td>
</tr>
</tbody>
</table>

The two data sets are dependent. Find \( \overline{d} \) to the nearest tenth.

13) \begin{tabular}{c|cccccc}
X & 8.0 & 7.3 & 7.1 & 5.7 & 6.8 & 5.7 \\
Y & 9.5 & 7.7 & 8.9 & 9.6 & 8.1 & 9.3 \\
\end{tabular}

A) -2.7 B) -12.6 C) -2.1 D) -1.3

Find \( s_d \).

14) The differences between two sets of dependent data are -9, 3, -9, 6. Round to the nearest tenth.

A) 7.9 B) 4.0 C) 6.3 D) 181.7
CHAPTER 9 FORM C

Assume that you want to test the claim that the paired sample data come from a population for which the mean difference is \( \mu_d = 0 \). Compute the value of the t test statistic.

\[
\begin{array}{c|cc}
\text{x} & 7 & 6 & 7 & 3 & 9 \\
\text{y} & 4 & 8 & 3 & 4 & 4 \\
\end{array}
\]

A) \( t = 1.292 \)  
B) \( t = 0.578 \)  
C) \( t = 2.890 \)  
D) \( t = 0.415 \)

Determine the decision criterion for rejecting the null hypothesis in the given hypothesis test; i.e., describe the values of the test statistic that would result in rejection of the null hypothesis.

16) Suppose you wish to test the claim that \( \mu_d \), the mean value of the differences \( d \) for a population of paired data, is greater than \( 0 \). Given a sample of \( n = 15 \) and a significance level of \( \alpha = 0.01 \), what criterion would be used for rejecting the null hypothesis?

A) Reject null hypothesis if test statistic > 2.977 or < -2.977.
B) Reject null hypothesis if test statistic > 2.602.
C) Reject null hypothesis if test statistic < 2.624.
D) Reject null hypothesis if test statistic > 2.624.

Use the traditional method of hypothesis testing to test the given claim about the means of two populations. Assume that two dependent samples have been randomly selected from normally distributed populations.

17) Ten different families are tested for the number of gallons of water a day they use before and after viewing a conservation video. At the 0.05 significance level, test the claim that the mean is the same before and after the viewing.

<table>
<thead>
<tr>
<th>Before</th>
<th>33</th>
<th>33</th>
<th>38</th>
<th>33</th>
<th>35</th>
<th>40</th>
<th>40</th>
<th>40</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>After</td>
<td>34</td>
<td>28</td>
<td>25</td>
<td>28</td>
<td>35</td>
<td>33</td>
<td>31</td>
<td>28</td>
<td>35</td>
</tr>
</tbody>
</table>

Construct a confidence interval for \( \mu_d \), the mean of the differences \( d \) for the population of paired data. Assume that the population of paired differences is normally distributed.

18) A test of writing ability is given to a random sample of students before and after they completed a formal writing course. The results are given below. Construct a 99% confidence interval for the mean difference between the before and after scores.

<table>
<thead>
<tr>
<th>Before</th>
<th>70</th>
<th>80</th>
<th>92</th>
<th>99</th>
<th>93</th>
<th>97</th>
<th>76</th>
<th>63</th>
<th>68</th>
<th>71</th>
<th>74</th>
</tr>
</thead>
<tbody>
<tr>
<td>After</td>
<td>69</td>
<td>79</td>
<td>90</td>
<td>96</td>
<td>91</td>
<td>95</td>
<td>75</td>
<td>64</td>
<td>62</td>
<td>64</td>
<td>76</td>
</tr>
</tbody>
</table>

A) \( 1.2 < \mu_d < 2.8 \)  
B) \( -0.2 < \mu_d < 4.2 \)
C) \( -0.1 < \mu_d < 4.1 \)  
D) \( -0.5 < \mu_d < 4.5 \)
CHAPTER 9 FORM C

Test the indicated claim about the variances or standard deviations of two populations. Assume that the populations are normally distributed. Assume that the two samples are independent and that they have been randomly selected.

19) A random sample of 16 women resulted in blood pressure levels with a standard deviation of 22.7 mm Hg. A random sample of 17 men resulted in blood pressure levels with a standard deviation of 20.6 mm Hg. Use a 0.025 significance level to test the claim that blood pressure levels for women have a larger variance than those for men.

20) A test for homogeneity of variance is conducted at the 5% level of significance. Samples sizes are \( n_1 = 25 \) and \( n_2 = 25 \). The test statistic is \( F = 1.0052 \). What do you know about the variance of the populations from which the samples were taken?
Answer Key
Testname: CHAPTER 9 FORM C

1) In the hypothesis test method, the comparison is made between a test statistic from the sample data and the critical value from the table. The conclusion is either to reject or fail to reject the null hypothesis based on whether or not the test statistic is in the reject region. With the confidence interval method, the confidence interval is constructed and the population mean is compared against the interval. The conclusion is made based on whether or not the population mean is within the confidence interval.

2) A
3) B
4) D
5) C
6) \( H_0: p_1 = p_2. \quad H_1: p_1 \neq p_2. \)
   Test statistic \( z = 0.97. \) Critical values: \( z = \pm 1.96. \)
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the recognition rates are the same in both states.

7) \( H_0: p_1 = p_2. \quad H_1: p_1 < p_2. \)
   Test statistic \( z = -1.80. \) Critical value: \( z = -2.05. \)
   Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the vaccine is effective in lowering the incidence of the disease.

8) C
9) A
10) \( H_0: \mu_1 = \mu_2. \quad H_1: \mu_1 < \mu_2. \)
    Test statistic \( t = -2.365. \) Critical value: \(-2.381 < t < -2.377 \) (Table A-3).
    [Optional: Critical value: \(-2.351 \) (STATDISK) with \( df = 155.0116712 \) (STATDISK, TI-83/84 Plus).
    P-value: \( p = 0.0096 \) (STATDISK, TI-83/84 Plus).]
    Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the treatment population mean \( \mu_1 \) is smaller than the control population \( \mu_2. \)
    [Optional: The technology procedure favors the diet, though barely.]

11) B
12) No, the P-value for a two-tail test is 0.0768, which is greater than the significance level of 0.05.
    There is not sufficient evidence to support the claim that the two population means are different.

13) C
14) A
15) A
16) D
17) Test statistic \( t = 2.894. \) Critical values: \( t = \pm 2.262. \)
    Reject \( H_0: \mu_d = 0. \) There is sufficient evidence to warrant rejection of the claim that the mean is the same before and after viewing.

18) D
19) \( H_0: \sigma_1^2 = \sigma_2^2. \quad H_1: \sigma_1^2 > \sigma_2^2. \)
    Test statistic \( F = 1.2143. \) Critical value: \( F = 2.7875 \)
    Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that blood pressure levels for women have a larger variance than those for men.

20) The critical F value for a two–tailed test of homogeneity of variance is \( F = 2.2693, df = 24, 24 \) and with 0.025 in the right tail. Since the TS F is in the critical region, we conclude that the populations have the same variation.

172
CHAPTER 10 FORM A

Name:__________________________________ Course Number:_______ Section Number:_____ 

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Describe the standard error of estimate, $s_e$. How do smaller values of $s_e$ relate to the dispersion of data points about the line determined by the linear regression equation? What does it mean when $s_e$ is 0?

Given the linear correlation coefficient $r$ and the sample size $n$, determine the critical values of $r$ and use your finding to state whether or not the given $r$ represents a significant linear correlation. Use a significance level of 0.05.

2) $r = 0.812$, $n = 9$
   
   A) Critical values: $r = \pm 0.666$, no significant linear correlation

   B) Critical value: $r = 0.666$, no significant linear correlation

   C) Critical values: $r = \pm 0.666$, significant linear correlation

   D) Critical value: $r = -0.666$, no significant linear correlation

Describe the error in the stated conclusion.

3) Given: Each school in a state reports the average SAT score of its students. There is a significant linear correlation between the average SAT score of a school and the average annual income in the district in which the school is located.

   Conclusion: There is a significant linear correlation between individual SAT scores and family income.
CHAPTER 10 FORM A

Determine which plot shows the strongest linear correlation.

4)  

A) 

B) 

C)
CHAPTER 10 FORM A

Find the value of the linear correlation coefficient \( r \).

5) A study was conducted to compare the average time spent in the lab each week versus course grade for computer students. The results are recorded in the table below.

<table>
<thead>
<tr>
<th>Number of hours spent in lab</th>
<th>Grade (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>96</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
</tr>
<tr>
<td>16</td>
<td>62</td>
</tr>
<tr>
<td>9</td>
<td>58</td>
</tr>
<tr>
<td>7</td>
<td>89</td>
</tr>
<tr>
<td>15</td>
<td>81</td>
</tr>
<tr>
<td>16</td>
<td>46</td>
</tr>
<tr>
<td>10</td>
<td>51</td>
</tr>
</tbody>
</table>

A) 0.017  B) 0.462  C) -0.284  D) -0.335

Find the best predicted value of \( y \) corresponding to the given value of \( x \).

6) Based on the data from six students, the regression equation relating number of hours of preparation (\( x \)) and test score (\( y \)) is \( y = 67.3 + 1.07x \). The same data yield \( r = 0.224 \) and \( \bar{y} = 75.2 \). What is the best predicted test score for a student who spent 3 hours preparing for the test?

A) 78.1  B) 75.2  C) 59.7  D) 70.5

Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.

7) \[
\begin{array}{c|cccc}
\hline
x & 6 & 8 & 20 & 28 & 36 \\
\hline
y & 2 & 4 & 13 & 20 & 30 \\
\hline
\end{array}
\]

A) \( \hat{y} = -3.79 + 0.801x \)  B) \( \hat{y} = -2.79 + 0.950x \)

C) \( \hat{y} = -3.79 + 0.897x \)  D) \( \hat{y} = -2.79 + 0.897x \)

Is the data point, \( P \), an outlier, an influential point, both, or neither?

8)

A) Both  B) Outlier

C) Influential point  D) Neither
CHAPTER 10 FORM A

Use the given information to find the coefficient of determination.

9) The test scores of 6 randomly picked students and the numbers of hours they prepared are as follows:

<table>
<thead>
<tr>
<th>Hours</th>
<th>5</th>
<th>10</th>
<th>4</th>
<th>6</th>
<th>10</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>64</td>
<td>86</td>
<td>69</td>
<td>86</td>
<td>59</td>
<td>87</td>
</tr>
</tbody>
</table>

The equation of the regression line is $\hat{y} = 1.06604x + 67.3491$. Find the coefficient of determination.

A) 0.6781  B) 0.2242  C) 0.0503  D) -0.2242

Use the computer display to answer the question.

10) A collection of paired data consists of the number of years that students have studied Spanish and their scores on a Spanish language proficiency test. A computer program was used to obtain the least squares linear regression line and the computer output is shown below. Along with the paired sample data, the program was also given an $x$ value of 2 (years of study) to be used for predicting test score.

The regression equation is

$\text{Score} = 31.55 + 10.90 \text{Years.}$

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>StDev</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>31.55</td>
<td>6.360</td>
<td>4.96</td>
<td>0.000</td>
</tr>
<tr>
<td>Years</td>
<td>10.90</td>
<td>1.744</td>
<td>6.25</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$S = 5.651 \quad \text{R-Sq} = 83.0\% \quad \text{R-Sq (Adj)} = 82.7\%$

Predicted values

<table>
<thead>
<tr>
<th>Fit</th>
<th>StDev Fit</th>
<th>95.0% CI</th>
<th>95.0% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.35</td>
<td>3.168</td>
<td>(42.72, 63.98)</td>
<td>(31.61, 75.09)</td>
</tr>
</tbody>
</table>

Use the information in the display to find the value of the linear correlation coefficient $r$. Determine whether there is significant linear correlation between years of study and test scores. Use a significance level of 0.05. There are 10 pairs of data.

A) $r = 0.83$; There is significant linear correlation.

B) $r = 0.91$; There is significant linear correlation.

C) $r = 0.91$; There is no significant linear correlation.

D) $r = 0.83$; There is no significant linear correlation.
CHAPTER 10 FORM A

Find the explained variation for the paired data.

11) The paired data below consists of test scores and hours of preparation for 5 randomly selected students. The equation of the regression line is \( \hat{y} = 44.8447 + 3.52427x \). Find the explained variation.

<table>
<thead>
<tr>
<th>x Hours of preparation</th>
<th>5</th>
<th>2</th>
<th>9</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>y Test of score</td>
<td>64</td>
<td>48</td>
<td>72</td>
<td>73</td>
<td>80</td>
</tr>
</tbody>
</table>

A) 599.2 B) 511.724 C) 87.4757 D) 498.103

Find the standard error of estimate for the given paired data.

12) The equation of the regression line for the paired data below is \( \hat{y} = 6.1826 + 4.33937x \). Find the standard error of estimate.

<table>
<thead>
<tr>
<th>x</th>
<th>9</th>
<th>7</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>22</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>43</td>
<td>35</td>
<td>16</td>
<td>21</td>
<td>23</td>
<td>102</td>
<td>81</td>
</tr>
</tbody>
</table>

A) 3.1270 B) 2.6959 C) 1.6419 D) 0.8275

Construct the indicated prediction interval for an individual y.

13) The equation of the regression line for the paired data below is \( \hat{y} = 6.1829 + 4.3394x \) and the standard error of estimate is \( s_e = 1.6419 \). Find the 99% prediction interval of y for \( x = 14 \).

<table>
<thead>
<tr>
<th>x</th>
<th>9</th>
<th>7</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>22</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>43</td>
<td>35</td>
<td>16</td>
<td>21</td>
<td>23</td>
<td>102</td>
<td>81</td>
</tr>
</tbody>
</table>

A) 59.6 < y < 74.2 B) 64.0 < y < 82.1
C) 66.9 < y < 79.2 D) 70.5 < y < 75.6

177
CHAPTER 10 FORM A

Use computer software to find the regression equation. Can the equation be used for prediction?

14) A wildlife analyst gathered the data in the table to develop an equation to predict the weights of bears. He used WEIGHT as the dependent variable and CHEST, LENGTH, and SEX as the independent variables. For SEX, he used male = 1 and female = 2.

<table>
<thead>
<tr>
<th>WEIGHT</th>
<th>CHEST</th>
<th>LENGTH</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>344</td>
<td>45.0</td>
<td>67.5</td>
<td>1</td>
</tr>
<tr>
<td>416</td>
<td>54.0</td>
<td>72.0</td>
<td>1</td>
</tr>
<tr>
<td>220</td>
<td>41.0</td>
<td>70.0</td>
<td>2</td>
</tr>
<tr>
<td>360</td>
<td>49.0</td>
<td>68.5</td>
<td>1</td>
</tr>
<tr>
<td>332</td>
<td>44.0</td>
<td>73.0</td>
<td>1</td>
</tr>
<tr>
<td>140</td>
<td>32.0</td>
<td>63.0</td>
<td>2</td>
</tr>
<tr>
<td>436</td>
<td>48.0</td>
<td>72.0</td>
<td>1</td>
</tr>
<tr>
<td>132</td>
<td>33.0</td>
<td>61.0</td>
<td>2</td>
</tr>
<tr>
<td>356</td>
<td>48.0</td>
<td>64.0</td>
<td>2</td>
</tr>
<tr>
<td>150</td>
<td>35.0</td>
<td>59.0</td>
<td>1</td>
</tr>
<tr>
<td>202</td>
<td>40.0</td>
<td>63.0</td>
<td>2</td>
</tr>
<tr>
<td>365</td>
<td>50.0</td>
<td>70.5</td>
<td>1</td>
</tr>
</tbody>
</table>

A) WEIGHT = 196 + 2.35CHEST + 3.40LENGTH + 25SEX;
Yes, because the $R^2$ is high

B) WEIGHT = -320 + 10.6CHEST + 7.3LENGTH - 10.7SEX;
Yes, because the P-value is high

C) WEIGHT = -442.6 + 12.1CHEST + 3.6LENGTH - 23.8SEX;
Yes, because the adjusted $R^2$ is high

D) WEIGHT = 442.6 + 12.1CHEST + 4.2LENGTH - 21SEX;
Yes, because the P-value is low

Construct a scatterplot and identify the mathematical model that best fits the data. Assume that the model is to be used only for the scope of the given data and consider only linear, quadratic, logarithmic, exponential, and power models. Use a calculator or computer to obtain the regression equation of the model that best fits the data. You may need to fit several models and compare the values of $R^2$.

15) The table below shows the population of a city (in millions) in each year during the period 1990 - 1995. Using the number of years since 1990 as the independent variable, find the regression equation of the best model.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>1.08</td>
<td>1.37</td>
<td>1.68</td>
<td>2.19</td>
<td>2.73</td>
<td>3.34</td>
</tr>
</tbody>
</table>

A) $y = 1.08 e^{0.228 x}$
B) $y = 0.930 + 0.454 x$
C) $y = 0.05 x^2 + 0.27 x + 1.06$
D) $y = 1.27 x^{0.550}$
Use computer software to obtain the regression and identify $R^2$, adjusted $R^2$, and the $P$-value.

16) An anti-smoking group used data in the table to relate the carbon monoxide of various brands of cigarettes to their tar and nicotine content.

<table>
<thead>
<tr>
<th>CO</th>
<th>TAR</th>
<th>NICOTINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1.2</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>1.2</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>1.0</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>0.8</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.8</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>0.8</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>1.0</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>1.2</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>0.7</td>
<td>9</td>
</tr>
<tr>
<td>18</td>
<td>1.4</td>
<td>18</td>
</tr>
<tr>
<td>16</td>
<td>1.0</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>0.8</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>1.1</td>
<td>16</td>
</tr>
</tbody>
</table>

A) .861, .900, .015  
B) .931, .902, .000  
C) .976, .921, .002  
D) .943, .934, .000

Use computer software to obtain the regression equation. Use the estimated equation to find the predicted value.

17) A health specialist gathered the data in the table to see if pulse rates can be explained by exercise and smoking. For exercise, he assigns 1 for yes, 2 for no. For smoking, he assigns 1 for yes, 2 for no. He then used his results to predict the pulse rate of a person whose exercise value was 1 and whose smoking value was 1.

<table>
<thead>
<tr>
<th>PULSE</th>
<th>EXERCISE</th>
<th>SMOKE</th>
</tr>
</thead>
<tbody>
<tr>
<td>97</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>88</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>69</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>67</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>83</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>77</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>66</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>78</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>73</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>67</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>82</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>70</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>76</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

A) 70 beats per minute  
B) 74 beats per minute  
C) 81 beats per minute  
D) 77 beats per minute
CHAPTER 10 FORM A

Find the indicated multiple regression equation.

18) Below are the productivity, dexterity, and job satisfaction ratings of ten randomly selected employees.

<table>
<thead>
<tr>
<th>Productivity</th>
<th>Dexterity</th>
<th>Job satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>49</td>
<td>56</td>
</tr>
<tr>
<td>25</td>
<td>53</td>
<td>58</td>
</tr>
<tr>
<td>28</td>
<td>59</td>
<td>60</td>
</tr>
<tr>
<td>21</td>
<td>42</td>
<td>50</td>
</tr>
<tr>
<td>21</td>
<td>47</td>
<td>54</td>
</tr>
<tr>
<td>25</td>
<td>53</td>
<td>61</td>
</tr>
<tr>
<td>30</td>
<td>63</td>
<td>59</td>
</tr>
<tr>
<td>34</td>
<td>67</td>
<td>63</td>
</tr>
<tr>
<td>36</td>
<td>75</td>
<td>67</td>
</tr>
</tbody>
</table>

Find the multiple regression equation that expresses the job satisfaction scores in terms of the productivity and dexterity scores.

A) \( \hat{S} = 28.28 + 0.517P + 0.086D \)  
B) \( \hat{S} = 28.28 + 0.086P + 0.517D \)  
C) \( \hat{S} = 28.28 + 0.011P + 0.437D \)  
D) \( \hat{S} = 28.28 + 0.237P + 0.728D \)

Use computer software to find the best regression equation to explain the variation in the dependent variable, \( Y \), in terms of the independent variables, \( X_1, X_2, X_3 \).

19) \[
\begin{array}{cccc}
Y & X_1 & X_2 & X_3 \\
344 & 45.0 & 67.5 & 1 \\
416 & 54.0 & 72.0 & 1 \\
220 & 41.0 & 70.0 & 2 \\
360 & 49.0 & 68.5 & 1 \\
332 & 44.0 & 73.0 & 1 \\
140 & 32.0 & 63.0 & 2 \\
436 & 48.0 & 72.0 & 1 \\
132 & 33.0 & 61.0 & 2 \\
356 & 48.0 & 64.0 & 2 \\
150 & 35.0 & 59.0 & 1 \\
202 & 40.0 & 63.0 & 2 \\
365 & 50.0 & 70.5 & 1 \\
\end{array}
\]

CORRELATION COEFFICIENTS

\( Y / X_1 = .951 \)  
\( Y / X_2 = .790 \)  
\( Y / X_3 = -.616 \)

COEFFICIENTS OF DETERMINATION

\( Y / X_1 = .905 \)  
\( Y / X_1, X_2 = .919 \)  
\( Y / X_1, X_2, X_3 = .927 \)

A) \( \hat{Y} = -543 + 12.8 X_1 + 4.15 X_2 \)  
B) \( \hat{Y} = -442 + 12.1 X_1 + 3.58 X_2 - 23.8 X_3 \)  
C) \( \hat{Y} = -412 + 13.6 X_1 + 3.15 X_2 \)  
D) \( \hat{Y} = -355 + 14.9 X_1 \)

180
CHAPTER 10 FORM A

Construct a scatter diagram for the given data.

\[
\begin{array}{cccccccc}
 x & 3 & -1 & -1 & -7 & 2 & 1 & 2 & -5 & -1 \\
 y & 3 & -3 & -1 & -7 & 3 & 2 & -10 & 2 & -8 & -2
\end{array}
\]

\[\begin{array}{c}
\text{A)}\\
\text{B)}\\
\text{C)}\\
\text{D)}
\end{array}\]
Answer Key
Testname: CHAPTER 10 FORM A

1) The standard error of estimate, \( s_\hat{e} \), is a measure of the distances between the observed sample \( y \) values and the predicted values \( \hat{y} \). Smaller values of \( s_\hat{e} \) indicate that the actual values of \( y \) will be closer to the regression line, whereas larger values of \( s_\hat{e} \) indicate a greater dispersion of the \( y \) values from the regression line. When the standard error of estimate is 0, the \( y \) values lie on the regression line.

2) C

3) Averages suppress individual variation and tend to inflate the correlation coefficient. The fact that there is significant linear correlation between average SAT scores and average incomes in the district does not necessarily imply that there is significant linear correlation between individual SAT scores and family incomes.

4) C
5) D
6) B
7) C
8) D
9) C
10) B
11) B
12) C
13) A
14) C
15) A
16) D
17) B
18) B
19) D
20) D
CHAPTER 10 FORM B

Name: ___________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided.
Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Suppose there is significant correlation between two variables. Describe two cases under which it might be inappropriate to use the linear regression equation for prediction. Give examples to support these cases.

Given the linear correlation coefficient $r$ and the sample size $n$, determine the critical values of $r$ and use your finding to state whether or not the given $r$ represents a significant linear correlation. Use a significance level of 0.05.

2) $r = -0.174$, $n = 15$
   A) Critical value: $r = 0.514$, no significant linear correlation
   B) Critical values: $r = \pm 0.532$, no significant linear correlation
   C) Critical values: $r = \pm 0.514$, significant linear correlation
   D) Critical values: $r = \pm 0.514$, no significant linear correlation

Describe the error in the stated conclusion.

3) Given: There is no significant linear correlation between scores on a math test and scores on a verbal test.

Conclusion: There is no relationship between scores on the math test and scores on the verbal test.
CHAPTER 10 FORM B

Determine which plot shows the strongest linear correlation.

4)

A)

B)

C)

Find the value of the linear correlation coefficient r.

5) Managers rate employees according to job performance and attitude. The results for several randomly selected employees are given below.

<table>
<thead>
<tr>
<th>Performance</th>
<th>59</th>
<th>63</th>
<th>65</th>
<th>69</th>
<th>58</th>
<th>77</th>
<th>69</th>
<th>70</th>
<th>64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude</td>
<td>72</td>
<td>67</td>
<td>78</td>
<td>82</td>
<td>75</td>
<td>87</td>
<td>92</td>
<td>83</td>
<td>78</td>
</tr>
</tbody>
</table>

A) 0.610         B) 0.729         C) 0.916         D) 0.863

Find the best predicted value of y corresponding to the given value of x.

6) The regression equation relating attitude rating (x) and job performance rating (y) for the employees of a company is \( \hat{y} = 11.7 + 1.02x \). Ten pairs of data were used to obtain the equation. The same data yield \( r = 0.863 \) and \( \bar{y} = 80.1 \). What is the best predicted job performance rating for a person whose attitude rating is 80? 

A) 91.9         B) 80.1         C) 12.6         D) 93.3

Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.

7) \[
\begin{array}{c|c|c|c|c|c}
\hline
x & 0 & 3 & 4 & 5 & 12 \\
\hline
y & 8 & 2 & 6 & 9 & 12 \\
\hline
\end{array}
\]

A) \( \hat{y} = 4.88 + 0.625x \)

B) \( \hat{y} = 4.88 + 0.525x \)

C) \( \hat{y} = 4.98 + 0.425x \)

D) \( \hat{y} = 4.98 + 0.725x \)
CHAPTER 10 FORM B

Is the data point, P, an outlier, an influential point, both, or neither?
8)

Use the given information to find the coefficient of determination.
9) The following are costs of advertising (in thousands of dollars) and the numbers of products sold (in thousands):

<table>
<thead>
<tr>
<th>Cost (in thousands)</th>
<th>9 2 3 4 2 5 9 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>85 52 55 68 67 83 73</td>
</tr>
</tbody>
</table>

The equation of the regression line is $\hat{y} = 2.7884x + 55.7885$. Find the coefficient of determination.

A) 0.5009  B) 0.2353  C) −0.0707  D) 0.7077

Construct the indicated prediction interval for an individual y.
10) The paired data below consists of heights and weights of 6 randomly selected adults.

The equation of the regression line is $\hat{y} = −181.342 + 144.46x$ and the standard error of estimate is $s_e = 5.0015$. Find the 95% prediction interval for the weight of a person whose height is 1.75 m.

<table>
<thead>
<tr>
<th>x Height (meters)</th>
<th>1.61 1.72 1.78 1.80 1.67 1.88</th>
</tr>
</thead>
<tbody>
<tr>
<td>y Weight (kg)</td>
<td>54 62 70 84 61 92</td>
</tr>
</tbody>
</table>

A) $56.5 < y < 86.5$  B) $52.1 < y < 90.9$
C) $58.5 < y < 84.5$  D) $65.4 < y < 77.6$
CHAPTER 10 FORM B

Use the computer display to answer the question.

11) A collection of paired data consists of the number of years that students have studied Spanish and their scores on a Spanish language proficiency test. A computer program was used to obtain the least squares linear regression line and the computer output is shown below. Along with the paired sample data, the program was also given an x value of 2 (years of study) to be used for predicting test score.

The regression equation is

\[ \text{Score} = 31.55 + 10.90 \text{Years} \]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>StDev</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>31.55</td>
<td>6.360</td>
<td>4.96</td>
<td>0.000</td>
</tr>
<tr>
<td>Years</td>
<td>10.90</td>
<td>1.744</td>
<td>6.25</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\[ S = 5.651 \quad \text{R-Sq} = 83.0\% \quad \text{R-Sq (Adj)} = 82.7\% \]

Predicted values

<table>
<thead>
<tr>
<th>Fit</th>
<th>StDev Fit</th>
<th>95.0% CI</th>
<th>95.0% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.3</td>
<td>3.168</td>
<td>(42.72, 63.98)</td>
<td>(31.61, 75.09)</td>
</tr>
</tbody>
</table>

For a person who studies for 2 years, obtain the 95\% prediction interval and write a statement interpreting the interval.

A) (31.61, 75.09); We can be 95\% confident that the mean test score of all individuals who study 2 years will lie in the interval (31.61, 75.09)

B) (31.61, 75.09); We can be 95\% confident that the test score of an individual who studies 2 years will lie in the interval (31.61, 75.09)

C) (42.72, 63.98); We can be 95\% confident that the test score of an individual who studies 2 years will lie in the interval (42.72, 63.98)

D) (42.72, 63.98); We can be 95\% confident that the mean test score of all individuals who study 2 years will lie in the interval (42.72, 63.98)

Find the unexplained variation for the paired data.

12) The paired data below consists of heights and weights of 6 randomly selected adults. The equation of the regression line is \( y = -181.342 + 144.46x \). Find the unexplained variation.

<table>
<thead>
<tr>
<th>x Height (meters)</th>
<th>y Weight (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.61</td>
<td>54</td>
</tr>
<tr>
<td>1.72</td>
<td>62</td>
</tr>
<tr>
<td>1.78</td>
<td>70</td>
</tr>
<tr>
<td>1.80</td>
<td>84</td>
</tr>
<tr>
<td>1.67</td>
<td>61</td>
</tr>
<tr>
<td>1.88</td>
<td>92</td>
</tr>
</tbody>
</table>

A) 1,079.5          B) 119.3          C) 100.06        D) 979.44
CHAPTER 10 FORM B

Use computer software to find the regression equation. Can the equation be used for prediction?

13) FPEA, the Farm Production Enhancement Agency, regressed corn output against acreage, rainfall, and a trend line. The trend line is proxy for technological advancement in farming from improved pest control, fertilization, land management, and farming implements.

<table>
<thead>
<tr>
<th>CORNPROD</th>
<th>ACRES</th>
<th>RAINFALL</th>
<th>TREND</th>
</tr>
</thead>
<tbody>
<tr>
<td>456</td>
<td>9896</td>
<td>29.1</td>
<td>1</td>
</tr>
<tr>
<td>421</td>
<td>9680</td>
<td>42.3</td>
<td>2</td>
</tr>
<tr>
<td>653</td>
<td>10449</td>
<td>29.8</td>
<td>3</td>
</tr>
<tr>
<td>573</td>
<td>10811</td>
<td>26.0</td>
<td>4</td>
</tr>
<tr>
<td>546</td>
<td>10014</td>
<td>34.3</td>
<td>5</td>
</tr>
<tr>
<td>499</td>
<td>10293</td>
<td>22.7</td>
<td>6</td>
</tr>
<tr>
<td>504</td>
<td>9413</td>
<td>24.2</td>
<td>7</td>
</tr>
<tr>
<td>611</td>
<td>9860</td>
<td>31.6</td>
<td>8</td>
</tr>
<tr>
<td>646</td>
<td>9782</td>
<td>25.6</td>
<td>9</td>
</tr>
<tr>
<td>789</td>
<td>12139</td>
<td>37.9</td>
<td>10</td>
</tr>
<tr>
<td>773</td>
<td>12166</td>
<td>33.9</td>
<td>11</td>
</tr>
<tr>
<td>753</td>
<td>9976</td>
<td>37.4</td>
<td>12</td>
</tr>
<tr>
<td>852</td>
<td>10645</td>
<td>27.0</td>
<td>13</td>
</tr>
<tr>
<td>755</td>
<td>9738</td>
<td>31.5</td>
<td>14</td>
</tr>
<tr>
<td>815</td>
<td>9933</td>
<td>39.9</td>
<td>15</td>
</tr>
<tr>
<td>902</td>
<td>10132</td>
<td>25.3</td>
<td>16</td>
</tr>
<tr>
<td>986</td>
<td>11145</td>
<td>30.4</td>
<td>17</td>
</tr>
<tr>
<td>909</td>
<td>9775</td>
<td>32.7</td>
<td>18</td>
</tr>
<tr>
<td>945</td>
<td>9549</td>
<td>35.0</td>
<td>19</td>
</tr>
<tr>
<td>866</td>
<td>10077</td>
<td>33.8</td>
<td>20</td>
</tr>
<tr>
<td>1178</td>
<td>11550</td>
<td>29.4</td>
<td>21</td>
</tr>
<tr>
<td>1230</td>
<td>10600</td>
<td>37.1</td>
<td>22</td>
</tr>
<tr>
<td>1207</td>
<td>11280</td>
<td>42.9</td>
<td>23</td>
</tr>
<tr>
<td>968</td>
<td>12100</td>
<td>32.2</td>
<td>24</td>
</tr>
<tr>
<td>1118</td>
<td>12420</td>
<td>30.5</td>
<td>25</td>
</tr>
</tbody>
</table>

A) CORNPROD = -16.3 + 2.6ACRES + 3.9RAINFALL + 21.3TREND;
   Yes, because the R^2 is high

B) CORNPROD = -21.1 + .036ACRES + 2.62RAINFALL + 27.6TREND;
   No, because the P-value is low

C) CORNPROD = -21.1 + .036ACRES + 2.62RAINFALL + 27.6TREND;
   Yes, because the R^2 is high

D) CORNPROD = -.9 + 1.68ACRES + .79RAINFALL + 10.2TREND;
   Yes, because the adjusted R^2 is high
Use computer software to obtain the regression and identify $R^2$, adjusted $R^2$, and the $P$-value.

14) A visitor to Yellowstone National Park sat down one day and observed Old Faithful, which faithfully spurs throughout the day, day in and day out. He surmised that the height of a given spurt was associated with the pressure build-up during the interval between spurs and by the momentum build-up during the duration of the spurt. He wrote down the data to test his hypothesis, and later ran a multiple regression program. What did he find?

<table>
<thead>
<tr>
<th>HEIGHT</th>
<th>INTERVAL</th>
<th>DURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>86</td>
<td>240</td>
</tr>
<tr>
<td>154</td>
<td>86</td>
<td>237</td>
</tr>
<tr>
<td>140</td>
<td>62</td>
<td>122</td>
</tr>
<tr>
<td>140</td>
<td>104</td>
<td>267</td>
</tr>
<tr>
<td>160</td>
<td>62</td>
<td>113</td>
</tr>
<tr>
<td>140</td>
<td>95</td>
<td>258</td>
</tr>
<tr>
<td>150</td>
<td>79</td>
<td>232</td>
</tr>
<tr>
<td>150</td>
<td>62</td>
<td>105</td>
</tr>
<tr>
<td>160</td>
<td>94</td>
<td>276</td>
</tr>
<tr>
<td>155</td>
<td>79</td>
<td>248</td>
</tr>
<tr>
<td>125</td>
<td>86</td>
<td>243</td>
</tr>
<tr>
<td>136</td>
<td>85</td>
<td>241</td>
</tr>
<tr>
<td>140</td>
<td>86</td>
<td>214</td>
</tr>
<tr>
<td>155</td>
<td>58</td>
<td>114</td>
</tr>
<tr>
<td>130</td>
<td>89</td>
<td>272</td>
</tr>
<tr>
<td>125</td>
<td>79</td>
<td>227</td>
</tr>
<tr>
<td>125</td>
<td>83</td>
<td>237</td>
</tr>
<tr>
<td>139</td>
<td>82</td>
<td>238</td>
</tr>
<tr>
<td>125</td>
<td>84</td>
<td>203</td>
</tr>
<tr>
<td>140</td>
<td>82</td>
<td>270</td>
</tr>
<tr>
<td>140</td>
<td>78</td>
<td>218</td>
</tr>
<tr>
<td>135</td>
<td>87</td>
<td>270</td>
</tr>
<tr>
<td>140</td>
<td>70</td>
<td>241</td>
</tr>
<tr>
<td>100</td>
<td>56</td>
<td>102</td>
</tr>
<tr>
<td>105</td>
<td>81</td>
<td>271</td>
</tr>
</tbody>
</table>

A) .213, .182, .213  B) .049, -.021, .123  
C) .089, .032, .634  D) .025, -.060, .750
CHAPTER 10 FORM B

Use computer software to obtain the regression equation. Use the estimated equation to find the predicted value.

15) A wildlife analyst gathered the data in the table to develop an equation to predict the weights of bears. He used WEIGHT in pounds as the dependent variable and CHEST, LENGTH, and SEX as the independent variables. For SEX, he used male = 1 and female = 2. He took his equation “to the forest” and found a male bear whose chest measured 40.3 inches and who was 64.0 inches long.

<table>
<thead>
<tr>
<th>WEIGHT</th>
<th>CHEST</th>
<th>LENGTH</th>
<th>SEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>344</td>
<td>45.0</td>
<td>67.5</td>
<td>1</td>
</tr>
<tr>
<td>416</td>
<td>54.0</td>
<td>72.0</td>
<td>1</td>
</tr>
<tr>
<td>220</td>
<td>41.0</td>
<td>70.0</td>
<td>2</td>
</tr>
<tr>
<td>360</td>
<td>49.0</td>
<td>68.5</td>
<td>1</td>
</tr>
<tr>
<td>332</td>
<td>44.0</td>
<td>73.0</td>
<td>1</td>
</tr>
<tr>
<td>140</td>
<td>32.0</td>
<td>63.0</td>
<td>2</td>
</tr>
<tr>
<td>436</td>
<td>48.0</td>
<td>72.0</td>
<td>1</td>
</tr>
<tr>
<td>132</td>
<td>33.0</td>
<td>61.0</td>
<td>2</td>
</tr>
<tr>
<td>356</td>
<td>48.0</td>
<td>64.0</td>
<td>2</td>
</tr>
<tr>
<td>150</td>
<td>35.0</td>
<td>59.0</td>
<td>1</td>
</tr>
<tr>
<td>202</td>
<td>40.0</td>
<td>63.0</td>
<td>2</td>
</tr>
<tr>
<td>365</td>
<td>50.0</td>
<td>70.5</td>
<td>1</td>
</tr>
</tbody>
</table>

A) 252 lb  B) 405 lb  C) 415 lb  D) 293 lb

Use computer software to find the best regression equation to explain the variation in the dependent variable, Y, in terms of the independent variables, X_1 and X_2.

16) Y  X_1  X_2
15  1.2  16
15  1.2  16
17  1.0  16  CORRELATION COEFFICIENT
  6  0.8  9
 1  0.1  1  Y/ X_1 = .886
 8  0.8  8  Y/ X_2 = .965
10  0.8  10
17  1.0  16  COEFFICIENTS OF DETERMINATION
15  1.2  15
11  0.7  9  Y/ X_2 = .932
18  1.4  18  Y/ X_2, X_1 = .943
16  1.0  15
10  0.8  9
 7  0.5  5
18  1.1  16

A) \( \hat{Y} = -0.49 + 14.07 \times X_1 \)
B) \( \hat{Y} = 1.38 - 5.53 \times X_1 + 1.33 \times X_2 \)
C) \( \hat{Y} = 0.42 + 0.99 \times X_2 \)
D) \( \hat{Y} = 1.25 - 1.55 \times X_1 + 5.79 \times X_2 \)

189
CHAPTER 10 FORM B

Construct a scatterplot and identify the mathematical model that best fits the data. Assume that the model is to be used only for the scope of the given data and consider only linear, quadratic, logarithmic, exponential, and power models. Use a calculator or computer to obtain the regression equation of the model that best fits the data. You may need to fit several models and compare the values of $R^2$.

17) A rock is dropped from a tall building and its distance (in feet) below the point of release is recorded as accurately as possible at various times after the moment of release. The results are shown in the table. Find the regression equation of the best model.

<table>
<thead>
<tr>
<th>x (seconds after release)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>y (distance in feet)</td>
<td>16</td>
<td>63</td>
<td>146</td>
<td>255</td>
<td>403</td>
<td>572</td>
</tr>
</tbody>
</table>

A) $y = 13.0 e^{0.686 x}$  
B) $y = -148.4 + 112 x$  
C) $y = 15.95 x^2$  
D) $y = -74.9 + 290 \ln x$

Find the indicated multiple regression equation.

18) Below are performance and attitude ratings of employees.

<table>
<thead>
<tr>
<th>Performance</th>
<th>59</th>
<th>63</th>
<th>65</th>
<th>69</th>
<th>58</th>
<th>77</th>
<th>76</th>
<th>69</th>
<th>70</th>
<th>64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude</td>
<td>72</td>
<td>67</td>
<td>78</td>
<td>82</td>
<td>75</td>
<td>87</td>
<td>92</td>
<td>83</td>
<td>87</td>
<td>78</td>
</tr>
</tbody>
</table>

Managers also rate the same employees according to adaptability, and below are the results that correspond to those given above.

| Adaptability | 50 | 52 | 54 | 60 | 66 | 59 | 62 | 55 |

Find the multiple regression equation that expresses performance in terms of attitude and adaptability.

A) $\hat{y} = 14.09 + 0.907(\text{Att.}) + 0.014(\text{Adapt.})$  
B) $\hat{y} = 14.09 + 0.213(\text{Att.}) + 0.895(\text{Adapt.})$  
C) $\hat{y} = 14.09 + 0.895(\text{Att.}) + 0.213(\text{Adapt.})$  
D) $\hat{y} = 14.09 + 0.014(\text{Att.}) + 0.907(\text{Adapt.})$

Find the standard error of estimate for the given paired data.

19) The paired data below consists of test scores and hours of preparation for 5 randomly selected students. The equation of the regression line is $y = 44.8447 + 3.52427x$. Find the standard error of estimate.

<table>
<thead>
<tr>
<th>x Hours of preparation</th>
<th>5</th>
<th>2</th>
<th>9</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>y Test score</td>
<td>64</td>
<td>48</td>
<td>72</td>
<td>73</td>
<td>80</td>
</tr>
</tbody>
</table>

A) 7.1720  
B) 13.060  
C) 4.1097  
D) 5.3999
Construct a scatter diagram for the given data.

<table>
<thead>
<tr>
<th>x</th>
<th>-5</th>
<th>5</th>
<th>10</th>
<th>8</th>
<th>9</th>
<th>7</th>
<th>-2</th>
<th>-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

A)

B)

C)

D)
Answer Key
Testname: CHAPTER 10 FORM B

1) Answers will vary. An answer follows. 1. The predicting value might not be within the scope of the original study. If a study compared the ages of men to an index of work motivation, and the men in the study were in the range of ages from 20 to 50, it would be inappropriate to predict for a man of age 60 or for a woman of any age.

2. The data used to create the linear regression equation might be too old. If the data were from the nineties, it would be inappropriate to predict for the year 2007.

2) D

3) Although there is no significant linear correlation between the two variables, we cannot conclude that no relationship exists as there could be a nonlinear relationship between the two variables.

4) A
5) D
6) D
7) B
8) B
9) A
10) A
11) B
12) C
13) C
14) D
15) A
16) C
17) C
18) D
19) D
20) C
CHAPTER 10 FORM C

Name:_____________________________ Course Number:_______ Section Number:_____

Directions: Write your answers to the short-answer items in the spaces provided.
Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Suppose that statisticians determine that there is a significant positive correlation between the grade earned in the class College Reading Skills and the grade earned in Statistics. Does achieving a high grade in reading cause an individual to earn a high grade in Statistics? Explain your answer with reference to the term lurking variable.

2) Suppose data set are collected concerning the weight of a person in pounds and the number of calories burned in 30 minutes of walking on a treadmill at 3.5 mph. How would the value of the correlation coefficient, r, change if all of the weights were converted to kilograms?

Given the linear correlation coefficient r and the sample size n, determine the critical values of r and use your finding to state whether or not the given r represents a significant linear correlation. Use a significance level of 0.05.

3) r = -0.236, n = 90
   A) Critical values: r = ±0.207, significant linear correlation
   B) Critical value: r = 0.217, significant linear correlation
   C) Critical values: r = ±0.217, no significant linear correlation
   D) Critical values: r = ±0.207, no significant linear correlation
CHAPTER 10 FORM C

Construct a scatter diagram for the given data.

4) \[
\begin{array}{c|cccc}
    x & 3 & -1 & -1 & 7 \\
    y & 3 & -3 & -1 & -7 \\
\end{array}
\]

A) [Diagram A]

B) [Diagram B]

C) [Diagram C]

D) [Diagram D]
CHAPTER 10 FORM C

Find the value of the linear correlation coefficient $r$.

5) Two separate tests are designed to measure a student's ability to solve problems. Several students are randomly selected to take both tests and the results are shown below.

<table>
<thead>
<tr>
<th>Test A</th>
<th>48</th>
<th>52</th>
<th>58</th>
<th>44</th>
<th>43</th>
<th>43</th>
<th>40</th>
<th>51</th>
<th>59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test B</td>
<td>73</td>
<td>67</td>
<td>73</td>
<td>59</td>
<td>58</td>
<td>58</td>
<td>64</td>
<td>74</td>
<td></td>
</tr>
</tbody>
</table>

A) 0.548  B) 0.714  C) 0.109  D) 0.867

Describe the error in the stated conclusion.

6) Given: The linear correlation coefficient between scores on a math test and scores on a test of athletic ability is negative and close to zero.

Conclusion: People who score high on the math test tend to score lower on the test of athletic ability.

Find the best predicted value of $y$ corresponding to the given value of $x$.

7) The regression equation relating dexterity scores ($x$) and productivity scores ($y$) for the employees of a company is $\hat{y} = 5.50 + 1.91x$. Ten pairs of data were used to obtain the equation. The same data yield $r = 0.986$ and $\bar{y} = 56.3$. What is the best predicted productivity score for a person whose dexterity score is 20?

A) 111.91  B) 56.30  C) 58.20  D) 43.7

Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.

8) $\begin{array}{c|cccc} x & 3 & 5 & 15 & 16 \\ \hline y & 8 & 11 & 7 & 14 \\ \hline \hat{y} & 10 & 15 & 17 & 18 \\ \end{array}$

A) $\hat{y} = 4.07 + 0.850x$  B) $\hat{y} = 5.07 + 0.753x$
C) $\hat{y} = 5.07 + 0.850x$  D) $\hat{y} = 4.07 + 0.753x$

Use the given information to find the coefficient of determination.

9) A regression equation is obtained for a collection of paired data. It is found that the total variation is 21.827, the explained variation is 18.387, and the unexplained variation is 3.44. Find the coefficient of determination.

A) 0.842  B) 0.187  C) 1.187  D) 0.158
CHAPTER 10 FORM C

Is the data point, P, an outlier, an influential point, both, or neither?

10)

Use the computer display to answer the question.

11) A collection of paired data consists of the number of years that students have studied Spanish and their scores on a Spanish language proficiency test. A computer program was used to obtain the least squares linear regression line and the computer output is shown below. Along with the paired sample data, the program was also given an x value of 2 (years of study) to be used for predicting test score.

The regression equation is

\[ \text{Score} = 31.55 + 10.90 \text{ Years}. \]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>StDev</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>31.55</td>
<td>6.360</td>
<td>4.96</td>
<td>0.000</td>
</tr>
<tr>
<td>Years</td>
<td>10.90</td>
<td>1.744</td>
<td>6.25</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\[ S = 5.651 \quad \text{R-Sq} = 83.0\% \quad \text{R-Sq (Adj)} = 82.7\% \]

Predicted values

<table>
<thead>
<tr>
<th>Fit</th>
<th>StDev Fit</th>
<th>95.0% CI</th>
<th>95.0% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.35</td>
<td>3.168</td>
<td>(42.72, 63.98)</td>
<td>(31.61, 75.09)</td>
</tr>
</tbody>
</table>

If a person studies 4.5 years, what is the single value that is the best predicted test score? Assume that there is a significant linear correlation between years of study and test score.

A) 83.0 \hspace{1cm} B) 53.35 \hspace{1cm} C) 80.6 \hspace{1cm} D) 49.1
CHAPTER 10 FORM C

Find the total variation for the paired data.

12) The equation of the regression line for the paired data below is $y = 6.1826 + 4.33937x$. Find the total variation.

\[ \begin{array}{c|cccccccc} x & 9 & 7 & 2 & 3 & 4 & 22 & 17 \\ \hline y & 43 & 35 & 16 & 21 & 23 & 102 & 81 \\ \end{array} \]

A) 6,693.27    B) 6,544.86    C) 13.479    D) 6,531.37

Find the standard error of estimate for the given paired data.

13) The paired data below consists of heights and weights of 6 randomly selected adults. The equation of the regression line is $y = -181.342 + 144.46x$. Find the standard error of estimate.

\[ \begin{array}{c|cccccc} x \text{ Height (meters)} & 1.61 & 1.72 & 1.78 & 1.80 & 1.67 & 1.88 \\ \hline y \text{ Weight (kg)} & 54 & 62 & 70 & 84 & 61 & 92 \\ \end{array} \]

A) 15.648    B) 6.9205    C) 5.0015    D) 9.7944

Construct the indicated prediction interval for an individual $y$.

14) The paired data below consists of test scores and hours of preparation for 5 randomly selected students. The equation of the regression line is $y = 44.845 + 3.524x$ and the standard error of estimate is $s_e = 5.40$. Find the 99% prediction interval for the test score of a person who spent 7 hours preparing for the test.

\[ \begin{array}{c|cccccccc} x \text{ Hours of preparation} & 5 & 2 & 9 & 6 & 10 \\ \hline y \text{ Test score} & 64 & 48 & 72 & 73 & 80 \\ \end{array} \]

A) $58 < y < 82$    B) $32 < y < 107$    C) $62 < y < 78$    D) $35 < y < 104$
CHAPTER 10 FORM C

Use computer software to find the regression equation. Can the equation be used for prediction?

15) An anti-smoking group used data in the table to relate the carbon monoxide of various brands of cigarettes to their tar and nicotine content.

<table>
<thead>
<tr>
<th>CO</th>
<th>TAR</th>
<th>NICOTINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1.2</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>1.2</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>1.0</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>0.8</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.8</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>0.8</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>1.0</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>1.2</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>0.7</td>
<td>9</td>
</tr>
<tr>
<td>18</td>
<td>1.4</td>
<td>18</td>
</tr>
<tr>
<td>16</td>
<td>1.0</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>0.8</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>1.1</td>
<td>16</td>
</tr>
</tbody>
</table>

A) \( CO = 1.3 + 5.5TAR - 1.3NIC; \)  
Yes, because the adjusted \( R^2 \) is high

B) \( CO = 1.25 + 1.55TAR - 5.79NIC; \)  
Yes, because the P-value is too low

C) \( CO = 1.37 - 5.53TAR + 1.33NIC; \)  
Yes, because the \( R^2 \) is high

D) \( CO = 1.37 + 5.50TAR - 1.38NIC; \)  
Yes, because the P-value is high
CHAPTER 10 FORM C

Use computer software to obtain the regression and identify $R^2$, adjusted $R^2$, and the $P$-value.

16) A study of food consumption in the country related the level of food consumed to an index of food prices and an index of personal disposable income.

<table>
<thead>
<tr>
<th>FOODCONS</th>
<th>INCOME</th>
<th>PRICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>98.6</td>
<td>87.4</td>
<td>108.5</td>
</tr>
<tr>
<td>101.2</td>
<td>97.6</td>
<td>110.1</td>
</tr>
<tr>
<td>102.4</td>
<td>96.7</td>
<td>110.4</td>
</tr>
<tr>
<td>100.9</td>
<td>98.2</td>
<td>104.3</td>
</tr>
<tr>
<td>102.3</td>
<td>99.8</td>
<td>107.2</td>
</tr>
<tr>
<td>101.5</td>
<td>100.5</td>
<td>105.8</td>
</tr>
<tr>
<td>101.6</td>
<td>103.2</td>
<td>107.8</td>
</tr>
<tr>
<td>101.6</td>
<td>107.8</td>
<td>103.4</td>
</tr>
<tr>
<td>99.8</td>
<td>96.6</td>
<td>102.7</td>
</tr>
<tr>
<td>100.3</td>
<td>88.9</td>
<td>104.1</td>
</tr>
<tr>
<td>97.6</td>
<td>75.1</td>
<td>99.2</td>
</tr>
<tr>
<td>97.2</td>
<td>76.9</td>
<td>99.7</td>
</tr>
<tr>
<td>97.3</td>
<td>84.6</td>
<td>102.0</td>
</tr>
<tr>
<td>96.0</td>
<td>90.6</td>
<td>94.3</td>
</tr>
<tr>
<td>99.2</td>
<td>103.1</td>
<td>97.7</td>
</tr>
<tr>
<td>100.3</td>
<td>105.1</td>
<td>101.1</td>
</tr>
<tr>
<td>100.3</td>
<td>96.4</td>
<td>102.3</td>
</tr>
<tr>
<td>104.1</td>
<td>104.4</td>
<td>104.4</td>
</tr>
<tr>
<td>105.3</td>
<td>110.7</td>
<td>108.5</td>
</tr>
<tr>
<td>107.6</td>
<td>127.1</td>
<td>111.3</td>
</tr>
</tbody>
</table>

A) .855, .844, .002  B) .912, .901, .010
C) .843, .799, .005  D) .867, .852, .000
CHAPTER 10 FORM C

Use computer software to obtain the regression equation. Use the estimated equation to find the predicted value.

17) A visitor to Yellowstone National Park sat down one day and observed Old Faithful, which faithfully spurs throughout the day, day in and day out. He surmised that the height of a given spurt was caused by the pressure build-up during the interval between spurs and by the momentum build-up during the duration of the spurt. He wrote down the data to test his hypothesis. He then used the regression equation he obtained to predict the height of the geyser if the interval is 88 seconds and the duration is 224 seconds.

<table>
<thead>
<tr>
<th>HEIGHT</th>
<th>INTERVAL</th>
<th>DURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>86</td>
<td>240</td>
</tr>
<tr>
<td>154</td>
<td>86</td>
<td>237</td>
</tr>
<tr>
<td>140</td>
<td>62</td>
<td>122</td>
</tr>
<tr>
<td>140</td>
<td>104</td>
<td>267</td>
</tr>
<tr>
<td>160</td>
<td>62</td>
<td>113</td>
</tr>
<tr>
<td>140</td>
<td>95</td>
<td>258</td>
</tr>
<tr>
<td>150</td>
<td>79</td>
<td>232</td>
</tr>
<tr>
<td>150</td>
<td>62</td>
<td>105</td>
</tr>
<tr>
<td>160</td>
<td>94</td>
<td>276</td>
</tr>
<tr>
<td>155</td>
<td>79</td>
<td>248</td>
</tr>
<tr>
<td>125</td>
<td>86</td>
<td>243</td>
</tr>
<tr>
<td>136</td>
<td>85</td>
<td>241</td>
</tr>
<tr>
<td>140</td>
<td>86</td>
<td>214</td>
</tr>
<tr>
<td>155</td>
<td>58</td>
<td>114</td>
</tr>
<tr>
<td>130</td>
<td>89</td>
<td>272</td>
</tr>
<tr>
<td>125</td>
<td>79</td>
<td>227</td>
</tr>
<tr>
<td>125</td>
<td>83</td>
<td>237</td>
</tr>
<tr>
<td>139</td>
<td>82</td>
<td>238</td>
</tr>
<tr>
<td>125</td>
<td>84</td>
<td>203</td>
</tr>
<tr>
<td>140</td>
<td>82</td>
<td>270</td>
</tr>
<tr>
<td>140</td>
<td>82</td>
<td>270</td>
</tr>
<tr>
<td>140</td>
<td>78</td>
<td>218</td>
</tr>
<tr>
<td>135</td>
<td>87</td>
<td>270</td>
</tr>
<tr>
<td>140</td>
<td>70</td>
<td>241</td>
</tr>
<tr>
<td>100</td>
<td>56</td>
<td>102</td>
</tr>
<tr>
<td>105</td>
<td>81</td>
<td>271</td>
</tr>
</tbody>
</table>

A) 132  B) 153  C) 139  D) 141
CHAPTER 10 FORM C

Find the indicated multiple regression equation.

18) A fitness rating was obtained for 9 randomly selected adult women. Each person was also asked her age, weight, and the number of hours she spent exercising each week. The results are shown below.

<table>
<thead>
<tr>
<th>Age</th>
<th>39</th>
<th>27</th>
<th>41</th>
<th>48</th>
<th>56</th>
<th>59</th>
<th>22</th>
<th>64</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>140</td>
<td>129</td>
<td>137</td>
<td>125</td>
<td>162</td>
<td>152</td>
<td>118</td>
<td>142</td>
<td>126</td>
</tr>
<tr>
<td>Hours of exercise per week</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>9</td>
<td>0</td>
<td>3</td>
<td>11</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Fitness rating</td>
<td>72</td>
<td>88</td>
<td>63</td>
<td>84</td>
<td>47</td>
<td>52</td>
<td>90</td>
<td>31</td>
<td>64</td>
</tr>
</tbody>
</table>

Identify the multiple regression equation that expresses fitness in terms of age, weight, and hours of exercise per week.

A) \( \hat{y} = 52.46 - 2.14A + 1.39W + 2.48E \)

B) \( \hat{y} = 23.79 - 1.36A + 0.241W + 1.43E \)

C) \( \hat{y} = 36.07 - 1.02A + 0.429W + 3.30E \)

Construct a scatterplot and identify the mathematical model that best fits the data. Assume that the model is to be used only for the scope of the given data and consider only linear, quadratic, logarithmic, exponential, and power models. Use a calculator or computer to obtain the regression equation of the model that best fits the data. You may need to fit several models and compare the values of \( R^2 \).

19) 

<table>
<thead>
<tr>
<th>x</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>9</td>
<td>13</td>
<td>25</td>
<td>27</td>
<td>31</td>
<td>46</td>
</tr>
</tbody>
</table>

A) \( y = 3.14 + 6.59x \)

B) \( y = 8.34x^{0.88} \)

C) \( y = 1.07 + 6.89x \)

D) \( y = 4.87 + 18.5\ln x \)
CHAPTER 10 FORM C

Use computer software to find the best regression equation to explain the variation in the dependent variable, Y, in terms of the independent variables, X₁, X₂, X₃.

20)  | Y     | X₁    | X₂    | X₃    |
    |-------|-------|-------|-------|
    | 456   | 9896  | 29.1  | 1     |
    | 421   | 9680  | 42.3  | 2     |
    | 653   | 10449 | 29.8  | 3     |
    | 573   | 10811 | 26.0  | 4     |
    | 546   | 10014 | 34.3  | 5     |
    | 499   | 10293 | 22.7  | 6     |
    | 504   | 9413  | 24.2  | 7     |
    | 611   | 9860  | 31.6  | 8     |
    | 646   | 9782  | 25.6  | 9     |
    | 789   | 12139 | 37.9  | 10    |
    | 773   | 12166 | 33.9  | 11    |
    | 753   | 9976  | 37.4  | 12    |
    | 852   | 10645 | 27.0  | 13    |
    | 755   | 9738  | 31.5  | 14    |
    | 815   | 9933  | 39.9  | 15    |
    | 902   | 10132 | 25.3  | 16    |
    | 986   | 11145 | 30.4  | 17    |
    | 909   | 9775  | 32.7  | 18    |
    | 945   | 9549  | 35.0  | 19    |
    | 866   | 10077 | 33.8  | 20    |
    | 1178  | 11550 | 29.4  | 21    |
    | 1230  | 10600 | 37.1  | 22    |
    | 1207  | 11280 | 42.9  | 23    |
    | 968   | 12100 | 32.2  | 24    |
    | 1118  | 12420 | 30.5  | 25    |

CORRELATION COEFFICIENTS

<table>
<thead>
<tr>
<th></th>
<th>Y/ X₁</th>
<th>Y/ X₂</th>
<th>Y/ X₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>456</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>421</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>653</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>573</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>546</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>499</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>504</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>611</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>646</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>789</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>773</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>753</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>852</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>755</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>815</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>902</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>986</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>909</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>945</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>866</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1178</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1207</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>968</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1118</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>202</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

COEFFICIENTS OF DETERMINATION

<table>
<thead>
<tr>
<th></th>
<th>Y/ X₁, X₂, X₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>456</td>
<td>Y/ X₁ = .509</td>
</tr>
<tr>
<td>421</td>
<td>Y/ X₂ = .280</td>
</tr>
<tr>
<td>653</td>
<td>Y/ X₃ = .930</td>
</tr>
</tbody>
</table>

A) \( \hat{Y} = 308.6 + 29.9 \times X₃ \)

B) \( \hat{Y} = 57.8 + 0.036 \times X₁ + 28.1 \times X₃ \)

C) \( \hat{Y} = 201.7 + 0.40 \times X₁ + 22.3 \times X₃ \)

D) \( \hat{Y} = -21.1 + 0.36 \times X₁ + 2.62 \times X₂ + 27.6 \times X₃ \)
Answer Key
Testname: CHAPTER 10 FORM C

1) A high grade in the College Reading Skills class does not necessarily cause a high grade in Statistics. These two variables could be related by an underlying relationship. Students who earn high grades in one class tend to earn high grades in other classes perhaps because of a lurking variable -- for example, motivation to achieve.

2) The value of $r$ would remain the same as a change of scale does not affect the value of $r$.

3) A
4) C
5) D

6) Because the linear correlation coefficient is close to zero and is probably not significant, no conclusion can be reached regarding the relationship between scores on the math test and scores on the test of athletic ability.

7) D
8) B
9) A
10) A
11) C
12) B
13) C
14) D
15) C
16) D
17) D
18) C
19) C
20) B
CHAPTER 11 FORM A

Name: ____________________________ Course Number: _______ Section Number: _____

**Directions:** Write your answers to the short-answer items in the spaces provided. Circle the correct choice for the multiple-choice item.

**Provide an appropriate response.**

1) Describe a goodness-of-fit test. What assumptions are made when using a goodness-of-fit test?

2) Explain the computation of expected values for contingency tables in terms of probabilities. Refer to the assumptions of the null hypothesis as part of your explanation. You might give a brief example to illustrate.

**Perform the indicated goodness-of-fit test.**

3) In studying the occurrence of genetic characteristics, the following sample data were obtained. At the 0.05 significance level, test the claim that the characteristics occur with the same frequency.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>28</td>
<td>30</td>
<td>45</td>
<td>48</td>
<td>38</td>
<td>39</td>
</tr>
</tbody>
</table>
CHAPTER 11 FORM A

Perform the indicated goodness-of-fit test.

4) Among the four northwestern states, Washington has 51% of the total population, Oregon has 30%, Idaho has 11%, and Montana has 8%. A market researcher selects a sample of 1000 subjects, with 450 in Washington, 340 in Oregon, 150 in Idaho, and 60 in Montana. At the 0.05 significance level, test the claim that the sample of 1000 subjects has a distribution that agrees with the distribution of state populations.

Provide an appropriate response.

5) The following table shows the number of employees who called in sick at a business for different days of a particular week.

<table>
<thead>
<tr>
<th>Day</th>
<th>Sun</th>
<th>Mon</th>
<th>Tues</th>
<th>Wed</th>
<th>Thurs</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number sick</td>
<td>8</td>
<td>12</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

i) At the 0.05 level of significance, test the claim that sick days occur with equal frequency on the different days of the week.

ii) Test the claim after changing the frequency for Saturday to 152. Describe the effect of this outlier on the test.

Use a $\chi^2$ test to test the claim that in the given contingency table, the row variable and the column variable are independent.

6) 160 students who were majoring in either math or English were asked a test question, and the researcher recorded whether they answered the question correctly. The sample results are given below. At the 0.10 significance level, test the claim that response and major are independent.

<table>
<thead>
<tr>
<th></th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math</td>
<td>27</td>
<td>53</td>
</tr>
<tr>
<td>English</td>
<td>43</td>
<td>37</td>
</tr>
</tbody>
</table>
Use a $\chi^2$ test to test the claim that in the given contingency table, the row variable and the column variable are independent.

7) Responses to a survey question are broken down according to gender and the sample results are given below. At the 0.05 significance level, test the claim that response and gender are independent.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Undecided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>25</td>
<td>50</td>
<td>15</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td>30</td>
<td>10</td>
</tr>
</tbody>
</table>

Solve the problem.

8) A researcher wishes to test whether the proportion of college students who smoke is the same in four different colleges. She randomly selects 100 students from each college and records the number that smoke. The results are shown below.

<table>
<thead>
<tr>
<th></th>
<th>College A</th>
<th>College B</th>
<th>College C</th>
<th>College D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoke</td>
<td>17</td>
<td>26</td>
<td>11</td>
<td>34</td>
</tr>
<tr>
<td>Don’t smoke</td>
<td>83</td>
<td>74</td>
<td>89</td>
<td>66</td>
</tr>
</tbody>
</table>

Use a 0.01 significance level to test the claim that the proportion of students smoking is the same at all four colleges.

9) Find the value of the test statistic $\chi^2$ by McNemar’s test for the following categorical data: $a = 25$, $b = 15$, $c = 20$, $d = 10$.

A) 1.029  B) 5.600  C) 2.314  D) 0.457
Provide an appropriate response.

10) A survey conducted in a small town yielded the results shown in the table.

<table>
<thead>
<tr>
<th>Plan to vote</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan to vote</td>
<td>105</td>
<td>87</td>
</tr>
<tr>
<td>Do not plan to vote</td>
<td>312</td>
<td>246</td>
</tr>
</tbody>
</table>

i) Test the claim that the intention to vote in the next presidential election is independent of the gender of the person being surveyed. Use a 0.05 significance level.

ii) Using Yates' correction, replace \( \frac{\sum (O - E)^2}{E} \) with \( \frac{\sum |O - E - 0.5|^2}{E} \) and repeat the test.

What effect does Yates' correction have on the value of the test statistic?
1) A goodness-of-fit test is used to test the hypothesis that an observed frequency distribution fits some claimed distribution. The assumptions are 1) the sample data are randomly selected; 2) the sample data consists of frequency counts for the different categories; and 3) for each of the categories, the expected frequency is at least 5.

2) Suppose A and B are two categories in a contingency table. In probability computations, the P(A and B) would be computed as P(A) \cdot P(B), provided A and B are independent. The assumption of the null hypothesis is that A and B are in fact independent, so we use the formula P(A and B) = P(A) \cdot P(B).

Since P(A) = \frac{\text{# occurrences A}}{\text{total occurrences}} and P(B) = \frac{\text{# occurrences B}}{\text{total occurrences}}, then

P(A \text{ and } B) = \frac{\text{# occurrences A}}{\text{total occurrences}} \cdot \frac{\text{# occurrences B}}{\text{total occurrences}}.

Then the expected number of outcomes for A and B would be

P(A \text{ and } B) \cdot \text{total occurrences}

= \frac{\text{# occurrences A}}{\text{total occurrences}} \cdot \frac{\text{# occurrences B}}{\text{total occurrences}} \cdot \text{total occurrences}.

or P(A \text{ and } B) \cdot \text{total occurrences}

= \frac{\text{# occurrences A} \cdot \text{# occurrences B}}{\text{total occurrences}}.

This is also the formula for the expected frequency for each cell in a contingency table,

E = \frac{\text{# occurrences A} \cdot \text{# occurrences B}}{\text{total occurrences}}.

So the computation of the expected values is based on the assumption of independence.

Examples may be given and will vary.

Possible answer: Rolling an assumed fair die 100 times.

3) H_0: The proportions of occurrences are all equal.
H_1: Those proportions are not all equal.

Test statistic: \chi^2 = 8.263. Critical value: \chi^2 = 11.071. Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the characteristics occur with the same frequency.

[Optional: P-value: p = 0.1423 STATDISK]

4) H_0: The distribution of the sample agrees with the population distribution.
H_1: It does not agree.

Test statistic: \chi^2 = 31.938. Critical value: \chi^2 = 7.815. Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the distribution of the sample agrees with the distribution of the state populations.

[Optional: P-value: p = 0.0000 STATDISK]

5) i) H_0: p_0 = p_1 = \ldots = p_7 = 1/7. H_1: At least one of these probabilities is different from the others. The value of the test statistic is \chi^2 = 2.4, which is less than the critical value of \chi^2 = 12.59. We fail to reject the null hypothesis.

[Optional: P-value: p = 0.8795 STATDISK]

ii) The value of the test statistic is \chi^2 = 579.5, which is greater than the critical value of \chi^2 = 12.59. We reject the null hypothesis. An outlier has a significant effect on the \chi^2 test statistic.

[Optional: P-value: p = 0.0000 STATDISK]
Answer Key
Testname: CHAPTER 11 FORM A

6) H₀: Major and response are independent.
H₁: Major and response are dependent.
Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that response and major are independent.
[Optional: P-value: $p = 0.0108$ STATDISK]

7) H₀: Gender and response are independent.
H₁: Gender and response are dependent.
Test statistic: $\chi^2 = 0.579$. Critical value: $\chi^2 = 5.991$.
Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that response and gender are independent.
[Optional: P-value: $p = 0.7487$ STATDISK]

8) H₀: The proportion of students smoking is the same at all four colleges.
H₁: The proportions are different.
Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the proportion of students smoking is the same at all four colleges.
[Optional: P-value: $p = 0.0005$ STATDISK]

9) D

10) i) H₀: Gender is independent of the intention to vote.
H₁: Gender and the intention to vote are dependent.
The value of the test statistic is $\chi^2 = 0.087$, which is less than the critical value of 3.841. We fail to reject the null hypothesis.
[Optional: P-value: $p = 0.7680$ STATDISK]

ii) The value of the test statistic is $\chi^2 = 0.044$, which is less than the critical value of 3.841. We still fail to reject the null hypothesis. Yates’ correction decreases the value of the test statistic.
CHAPTER 11 FORM B

Name: __________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for the multiple-choice item.

Provide an appropriate response.

1) Define categorical data and give an example.

2) In the chi-square test of independence, the formula used is \( \chi^2 = \frac{\sum (O - E)^2}{E} \). Discuss the meaning of O and E and explain the circumstances under which the \( \chi^2 \) values will be smaller or larger. What is the relationship between a significant \( \chi^2 \) value and the values of O and E?

Perform the indicated goodness-of-fit test.

3) A company manager wishes to test a union leader's claim that absences occur on the different week days with the same frequencies. Test this claim at the 0.05 level of significance if the following sample data have been compiled.

<table>
<thead>
<tr>
<th>Day</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thur</th>
<th>Fri</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absences</td>
<td>37</td>
<td>15</td>
<td>12</td>
<td>23</td>
<td>43</td>
</tr>
</tbody>
</table>
CHAPTER 11 FORM B

Perform the indicated goodness-of-fit test.

4) In studying the responses to a multiple-choice test question, the following sample data were obtained. At the 0.05 significance level, test the claim that the responses occur with the same frequency.

<table>
<thead>
<tr>
<th>Response</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>12</td>
<td>15</td>
<td>16</td>
<td>18</td>
<td>19</td>
</tr>
</tbody>
</table>

Provide an appropriate response.

5) An observed frequency distribution is as follows:

<table>
<thead>
<tr>
<th>Number of successes</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>41</td>
<td>93</td>
<td>66</td>
</tr>
</tbody>
</table>

i) Assuming a binomial distribution with \( n = 2 \) and \( p = 1/2 \), use the binomial formula to find the probability corresponding to each category of the table.

ii) Using the probabilities found in part (i), find the expected frequency for each category.

iii) Use a 0.05 level of significance to test the claim that the observed frequencies fit a binomial distribution for which \( n = 2 \) and \( p = 1/2 \).

Use a \( \chi^2 \) test to test the claim that in the given contingency table, the row variable and the column variable are independent.

6) The table below shows the age and favorite type of music of 668 randomly selected people.

<table>
<thead>
<tr>
<th></th>
<th>Rock</th>
<th>Pop</th>
<th>Classical</th>
</tr>
</thead>
<tbody>
<tr>
<td>15–25</td>
<td>50</td>
<td>85</td>
<td>73</td>
</tr>
<tr>
<td>25–35</td>
<td>68</td>
<td>91</td>
<td>60</td>
</tr>
<tr>
<td>35–45</td>
<td>90</td>
<td>74</td>
<td>77</td>
</tr>
</tbody>
</table>

Use a 5 percent level of significance to test the null hypothesis that age and preferred music type are independent.
CHAPTER 11 FORM B

Use a $\chi^2$ test to test the claim that in the given contingency table, the row variable and the column variable are independent.

7) Use the sample data below to test whether car color affects the likelihood of being in an accident. Use a significance level of 0.01.

<table>
<thead>
<tr>
<th>Car has been in accident</th>
<th>Red</th>
<th>Blue</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car has not been in accident</td>
<td>28</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>22</td>
<td>30</td>
</tr>
</tbody>
</table>

Solve the problem.

8) Use a 0.01 significance level to test the claim that the proportion of men who plan to vote in the next election is the same as the proportion of women who plan to vote. 300 men and 300 women were randomly selected and asked whether they planned to vote in the next election. The results are shown below.

<table>
<thead>
<tr>
<th>Plan to vote</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan to vote</td>
<td>170</td>
<td>185</td>
</tr>
<tr>
<td>Do not plan to vote</td>
<td>130</td>
<td>115</td>
</tr>
</tbody>
</table>

9) Find the test statistic $\chi^2$ by McNemar's test for the following categorical data: $a = 50$, $b = 60$, $c = 85$, $d = 15$.

A) 3.972    B) 4.303    C) 4.267    D) 5.673
Provide an appropriate response.

10) A survey conducted in a small business yielded the results shown in the table.

<table>
<thead>
<tr>
<th>Health insurance</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health insurance</td>
<td>41</td>
<td>22</td>
</tr>
<tr>
<td>No health insurance</td>
<td>34</td>
<td>24</td>
</tr>
</tbody>
</table>

i) Test the claim that health care coverage is independent of gender. Use a 0.05 significance level.

ii) Using Yates’ correction, replace \( \sum \frac{(O - E)^2}{E} \) with \( \sum \frac{(O - E - 0.5)^2}{E} \) and repeat the test.

What effect does Yates’ correction have on the value of the test statistic?
1) Categorical data are data that can be separated into different nonnumeric categories. Examples will vary. Possible answer: Colors of recently manufactured GMC trucks.

2) The O represents the observed frequencies. The E represents the expected frequencies based on the assumption of independence. The $\chi^2$ value will be smaller when the difference between observed and expected frequencies is small and will be larger when the difference between observed and expected values is large. The $\chi^2$ value will be significant when there is a significant difference between the observed and expected values.

3) H0: The proportions of absences are all the same.
   
   H1: The proportions of absences are not all the same.

   Test statistic: $\chi^2 = 28.308$. Critical value: $\chi^2 = 9.488$. Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that absences occur on the different week days with the same frequency.

   [Optional: P-value: p = 0.0000 STATDISK]

4) H0: The proportions of responses are all equal.
   
   H1: The proportions of responses are not all equal.

   Test statistic: $\chi^2 = 1.875$. Critical value: $\chi^2 = 9.488$. Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the responses occur with the same frequency.

   [Optional: P-value: p = 0.7587 STATDISK]

5) i) Number of successes 0 1 2
   Probability 0.25 0.50 0.25

   ii) Number of successes 0 1 2
   Expected frequency 50 100 50

   iii) H0: $p_0 = 0.25$, $p_1 = 0.50$, $p_3 = 0.25$; H1: At least one of the above probabilities is different from the claimed value. The value of the test statistic is $\chi^2 = 7.23$, which is greater than the critical value of $\chi^2 = 5.991$. We reject the null hypothesis that the observed frequency distribution is binomially distributed.

6) H0: Age and preferred music type are independent.
   
   H1: Age and preferred music type are dependent.

   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that age and preferred music type are independent.

   [Optional: P-value: p = 0.0115 STATDISK]

7) H0: Car color and being in an accident are independent.
   
   H1: Car color and being in an accident are dependent.

   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that car color and being in an accident are independent.

   [Optional: P-value: p = 0.8071 STATDISK]
8) H₀: The proportion of men who plan to vote in the next election is the same as the proportion of women who plan to vote.  
   H₁: The proportions are different.  
   Test statistic: $\chi^2 = 1.552$. Critical value: $\chi^2 = 6.635$.  
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the proportion of men who plan to vote in the next election is the same as the proportion of women who plan to vote.  
   [Optional: P-value: p = 0.2128 STATDISK]

9) A

10) i) H₀: Gender is independent of the health care coverage.  
    H₁: Gender and health care coverage are dependent.  
    The value of the test statistic is $\chi^2 = 0.5346$, which is less than the critical value of 3.841. We fail to reject the null hypothesis.  
    [Optional: P-value: p = 0.4647 STATDISK]  
    ii) The value of the test statistic is $\chi^2 = 0.2955$, which is still less than the critical value of 3.841. We still fail to reject the null hypothesis. Yates’ correction decreases the value of the test statistic.
CHAPTER 11 FORM C

Name:________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for the multiple-choice item.

Provide an appropriate response.

1) Describe the null hypothesis for the test of independence. List the assumptions for the $\chi^2$ test of independence. What is the major difference between the assumptions for this test and the assumptions for the previous tests we have studied?

2) Describe the test of homogeneity. What characteristic distinguishes a test of homogeneity from a test of independence?

Perform the indicated goodness-of-fit test.

3) Using the data below and a 0.05 significance level, test the claim that the responses occur with percentages of 15%, 20%, 25%, 25%, and 15% respectively.

<table>
<thead>
<tr>
<th>Response</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>12</td>
<td>15</td>
<td>16</td>
<td>18</td>
<td>19</td>
</tr>
</tbody>
</table>
CHAPTER 11 FORM C

Perform the indicated goodness-of-fit test.

4) Use a significance level of 0.01 to test the claim that workplace accidents are distributed on workdays as follows: Monday 25%, Tuesday: 15%, Wednesday: 15%, Thursday: 15%, and Friday: 30%.

In a study of 100 workplace accidents, 24 occurred on a Monday, 13 occurred on a Tuesday, 16 occurred on a Wednesday, 16 occurred on a Thursday, and 31 occurred on a Friday.

Provide an appropriate response.

5) An observed frequency distribution is as follows:

<table>
<thead>
<tr>
<th>Number of successes</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>47</td>
<td>98</td>
<td>55</td>
</tr>
</tbody>
</table>

i) Assuming a binomial distribution with \( n = 2 \) and \( p = 1/2 \), use the binomial formula to find the probability corresponding to each category of the table.

ii) Using the probabilities found in part (i), find the expected frequency for each category.

iii) Use a 0.05 level of significance to test the claim that the observed frequencies fit a binomial distribution for which \( n = 2 \) and \( p = 1/2 \).

Use a \( \chi^2 \) test to test the claim that in the given contingency table, the row variable and the column variable are independent.

6) Tests for adverse reactions to a new drug yielded the results given in the table. At the 0.05 significance level, test the claim that the treatment (drug or placebo) is independent of the reaction (whether or not headaches were experienced).

<table>
<thead>
<tr>
<th>Headaches</th>
<th>Drug</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>91</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 11 FORM C

Use a $\chi^2$ test to test the claim that in the given contingency table, the row variable and the column variable are independent.

7) Responses to a survey question are broken down according to employment status and the sample results are given below. At the 0.10 significance level, test the claim that response and employment status are independent.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Undecided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>30</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Unemployed</td>
<td>20</td>
<td>25</td>
<td>10</td>
</tr>
</tbody>
</table>

Solve the problem.

8) A researcher wishes to test the effectiveness of a flu vaccination. 150 people are vaccinated, 180 people are vaccinated with a placebo, and 100 people are not vaccinated. The number in each group who later caught the flu was recorded. The results are shown below.

<table>
<thead>
<tr>
<th></th>
<th>Vaccinated</th>
<th>Placebo</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caught the flu</td>
<td>8</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>Did not catch the flu</td>
<td>142</td>
<td>161</td>
<td>79</td>
</tr>
</tbody>
</table>

Use a 0.05 significance level to test the claim that the proportion of people catching the flu is the same in all three groups.

9) Find the test statistic $\chi^2$ by McNemar’s test for the following categorical data: $a = 72$, $b = 68$, $c = 110$, $d = 50$.

A) 3.615  B) 7.934  C) 9.444  D) 10.388
CHAPTER 11 FORM C

Provide an appropriate response.

10) A survey conducted in a small business yielded the results shown in the table.

<table>
<thead>
<tr>
<th>Health insurance</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health insurance</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>No health insurance</td>
<td>30</td>
<td>10</td>
</tr>
</tbody>
</table>

i) Test the claim that health care coverage is independent of gender. Use a 0.05 significance level.

ii) Using Yates’ correction, replace $\sum \frac{(O - E)^2}{E}$ with $\sum \frac{(|O - E| - 0.5)^2}{E}$ and repeat the test.

What effect does Yates' correction have on the value of the test statistic?
1) The null hypothesis is that the row and column variables in a contingency table are independent; that is, they are not related. The assumptions are 1) the null hypothesis is that the row and column variables are independent while the alternate hypothesis is that the row and column variables are dependent; 2) the sample data are randomly selected; and 3) each cell in the contingency table has an expected frequency E of at least 5. The major difference is that these assumptions do not require that the parent population be normally distributed.

2) The test of homogeneity tests the claim that different populations have the same proportions of some characteristics. In the test of homogeneity, there are predetermined totals for either the rows or columns of the contingency table. In the test of independence, there is one big sample drawn so that the row and column totals are determined randomly. In the test of homogeneity, predetermined sample sizes are used for each population.

3) H₀: The responses occur according to the stated percentages.
   H₁: The responses do not occur according to the stated percentages.
   Test statistic: \( \chi^2 = 5.146 \). Critical value: \( \chi^2 = 9.488 \). Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the responses occur according to the stated percentages.
   [Optional: P-value: p = 0.2727 STATDISK]

4) H₀: Workplace accidents occur according to the stated percentages.
   H₁: Workplace accidents do not occur according to the stated percentages.
   Test statistic: \( \chi^2 = 0.473 \). Critical value: \( \chi^2 = 13.277 \). Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that workplace accidents occur according to the stated percentages.
   [Optional: P-value: p = 0.9760 STATDISK]

5) i) Number of successes | 0 | 1 | 2
   Probability | 0.25 | 0.50 | 0.25
   ii) Number of successes | 0 | 1 | 2
   Expected frequency | 50 | 100 | 50
   iii) H₀: \( p_0 = 0.25 \), \( p_1 = 0.50 \), \( p_3 = 0.25 \); H₁: At least one of the above probabilities is different from the claimed value. The value of the test statistic is \( \chi^2 = 0.72 \), which is less than the critical value of \( \chi^2 = 5.991 \). We fail to reject the null hypothesis.

6) H₀: Treatment and reaction are independent.
   H₁: Treatment and reaction are dependent.
   Test statistic: \( \chi^2 = 1.798 \). Critical value: \( \chi^2 = 3.841 \).
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that treatment and reaction are independent.
   [Optional: P-value: p = 0.1799 STATDISK]

7) H₀: Employment status and response are independent.
   H₁: Employment status and response are dependent.
   Test statistic: \( \chi^2 = 5.942 \). Critical value: \( \chi^2 = 4.605 \).
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that response and employment status are independent.
   [Optional: P-value: p = 0.0513 STATDISK]
Answer Key
Testname: CHAPTER 11 FORM C

8) H₀: The proportion of people catching the flu is the same in all three groups.
   H₁: The proportions are different.

   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the proportion of people catching the flu is the same in all three groups.
   [Optional: P-value: $p = 0.0006$ STATDISK]

9) C

10) i) H₀: Gender is independent of the health care coverage.
    H₁: Gender and health care coverage are dependent.

    The value of the test statistic is $\chi^2 = 0.1637$, which is less than the critical value of 3.841. We fail to reject the null hypothesis.
    [Optional: P-value: $p = 0.6858$ STATDISK]

   ii) The value of the test statistic is $\chi^2 = 0.0333$, which is still less than the critical value of 3.841. We still fail to reject the null hypothesis. Yates’ correction decreases the value of the test statistic.
CHAPTER 12 FORM A

Name:__________________________ Course Number: ________ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) List the assumptions for testing hypotheses that three or more means are equivalent.

2) Define the term "treatment". What other term means the same thing? Give an example.

Given below are the analysis of variance results from a Minitab display. Assume that you want to use a 0.05 significance level in testing the null hypothesis that the different samples come from populations with the same mean.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>30</td>
<td>10.00</td>
<td>1.6</td>
<td>0.264</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>50</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Identify the value of the test statistic.

A) 0.264 B) 1.6 C) 10.00 D) 30

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>13.500</td>
<td>4.500</td>
<td>5.17</td>
<td>0.011</td>
</tr>
<tr>
<td>Error</td>
<td>16</td>
<td>13.925</td>
<td>0.870</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>27.425</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What can you conclude about the equality of the population means?

A) Reject the null hypothesis since the p-value is less than the significance level.
B) Accept the null hypothesis since the p-value is greater than the significance level.
C) Reject the null hypothesis since the p-value is greater than the significance level.
D) Accept the null hypothesis since the p-value is less than the significance level.

222
CHAPTER 12 FORM A

Test the claim that the samples come from populations with the same mean. Assume that the populations are normally distributed with the same variance.

5) Given the sample data below, test the claim that the populations have the same mean. Use a significance level of 0.05.

<table>
<thead>
<tr>
<th></th>
<th>Brand A</th>
<th>Brand B</th>
<th>Brand C</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( \bar{x} )</td>
<td>31.5</td>
<td>32.1</td>
<td>27.3</td>
</tr>
<tr>
<td>( s^2 )</td>
<td>3.29</td>
<td>4.34</td>
<td>3.32</td>
</tr>
</tbody>
</table>

6) A consumer magazine wants to compare the lifetimes of ballpoint pens of three different types. The magazine takes a random sample of pens of each type in the following table.

<table>
<thead>
<tr>
<th></th>
<th>Brand 1</th>
<th>Brand 2</th>
<th>Brand 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>260</td>
<td>181</td>
<td>238</td>
<td></td>
</tr>
<tr>
<td>218</td>
<td>240</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td>184</td>
<td>162</td>
<td>241</td>
<td></td>
</tr>
<tr>
<td>219</td>
<td>218</td>
<td>213</td>
<td></td>
</tr>
</tbody>
</table>

Do the data indicate that there is a difference in mean lifetime for the three brands of ballpoint pens? Use \( \alpha = 0.01 \).

7) At the 0.025 significance level, test the claim that the four brands have the same mean if the following sample results have been obtained.

<table>
<thead>
<tr>
<th></th>
<th>Brand A</th>
<th>Brand B</th>
<th>Brand C</th>
<th>Brand D</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>20</td>
<td>21</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>17</td>
<td>22</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>20</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>23</td>
<td>19</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>18</td>
<td>22</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>28</td>
<td>28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Provide an appropriate response.

8) Fill in the missing entries in the following partially completed one-way ANOVA table.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS=SS/df</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3</td>
<td>22.97</td>
<td>7.66</td>
<td>11.16</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>13.72</td>
<td>0.686</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>36.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A) | Source | df | SS   | MS=SS/df | F-statistic |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3</td>
<td>22.97</td>
<td>7.66</td>
<td>11.16</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>13.72</td>
<td>0.686</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>36.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B) | Source | df | SS   | MS=SS/df | F-statistic |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3</td>
<td>0.184</td>
<td>0.061</td>
<td>11.16</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>13.72</td>
<td>0.686</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>13.90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C) | Source | df | SS   | MS=SS/df | F-statistic |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3</td>
<td>2.55</td>
<td>7.66</td>
<td>11.16</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>13.72</td>
<td>0.686</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>16.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D) | Source | df | SS   | MS=SS/df | F-statistic |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3</td>
<td>48.80</td>
<td>16.27</td>
<td>11.16</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>13.72</td>
<td>0.686</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>62.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 12 FORM A

Solve the problem.

9) Use the data given below to verify that the t test for independent samples and the
ANOVA method are equivalent.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>74</td>
</tr>
<tr>
<td>81</td>
<td>72</td>
</tr>
<tr>
<td>73</td>
<td>65</td>
</tr>
<tr>
<td>91</td>
<td>83</td>
</tr>
<tr>
<td>64</td>
<td>59</td>
</tr>
</tbody>
</table>

i) Use a t test with a 0.05 significance level to test the claim that the two samples come
from populations with the same means. (Assume the standard deviations of the
populations are equal.)
ii) Use the ANOVA method with a 0.05 significance level to test the same claim.
iii) Verify that the squares of the t test statistic and the critical value are equal to the F
test statistic and critical value.

Use the Minitab display to test the indicated claim.

10) A manager records the production output of three employees who each work on three
different machines for three different days. The sample results are given below and the
Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>23,</td>
<td>27,</td>
<td>30,</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>18,</td>
<td>20,</td>
<td>22</td>
</tr>
<tr>
<td>Machine II</td>
<td>25,</td>
<td>26,</td>
<td>24,</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>24</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>19,</td>
<td>16,</td>
<td>14</td>
</tr>
<tr>
<td>Machine III</td>
<td>28,</td>
<td>25,</td>
<td>25,</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>15,</td>
<td>11,</td>
<td>17</td>
</tr>
</tbody>
</table>

ANALYSIS OF VARIANCE ITEMS

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>34.67</td>
<td>17.33</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>504.67</td>
<td>252.33</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>26.67</td>
<td>6.67</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>98.00</td>
<td>5.44</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>664.00</td>
<td></td>
</tr>
</tbody>
</table>

Assume that the number of items produced is not affected by an interaction between
employee and machine. Using a 0.05 significance level, test the claim that the machine
has no effect on the number of items produced.
CHAPTER 12 FORM A

Use the Minitab display to test the indicated claim.

11) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>16, 18, 19</td>
<td>15, 17, 20</td>
<td>14, 18, 16</td>
</tr>
<tr>
<td>Machine</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20, 27, 29</td>
<td>25, 28, 27</td>
<td>29, 28, 26</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>15, 18, 17</td>
<td>16, 16, 19</td>
</tr>
</tbody>
</table>

**ANALYSIS OF VARIANCE ITEMS**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>588.74</td>
<td>294.37</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>2.07</td>
<td>1.04</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>15.48</td>
<td>3.87</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>98.67</td>
<td>5.48</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>704.96</td>
<td></td>
</tr>
</tbody>
</table>

Assume that the number of items produced is not affected by an interaction between employee and machine. Using a 0.05 significance level, test the claim that the machine has no effect on the number of items produced.

12) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>31, 34, 32</td>
<td>29, 23, 22</td>
<td>21, 20, 24</td>
</tr>
<tr>
<td>Machine</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19, 26, 22</td>
<td>35, 33, 30</td>
<td>25, 19, 23</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>21, 18, 26</td>
<td>20, 23, 24</td>
</tr>
</tbody>
</table>

**ANALYSIS OF VARIANCE ITEMS**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>1.19</td>
<td>.59</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>5.85</td>
<td>2.93</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>710.81</td>
<td>177.70</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>160.00</td>
<td>8.89</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>877.85</td>
<td></td>
</tr>
</tbody>
</table>

Using a 0.05 significance level, test the claim that the interaction between employee and machine has no effect on the number of items produced.
CHAPTER 12 FORM A

Use the data in the given table and the corresponding Minitab display to test the hypothesis.

13) The following table entries are the times in seconds for three different drivers racing on four different tracks. Assuming no effect from the interaction between driver and track, test the claim that the track has no effect on the time. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>Track 1</th>
<th>Track 2</th>
<th>Track 3</th>
<th>Track 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver 1</td>
<td>72</td>
<td>70</td>
<td>68</td>
</tr>
<tr>
<td>Driver 2</td>
<td>74</td>
<td>71</td>
<td>66</td>
</tr>
<tr>
<td>Driver 3</td>
<td>76</td>
<td>69</td>
<td>64</td>
</tr>
</tbody>
</table>

Source | DF | SS | MS | F | p |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.33</td>
<td>0.729</td>
</tr>
<tr>
<td>Track</td>
<td>3</td>
<td>98.25</td>
<td>32.75</td>
<td>10.92</td>
<td>0.00763</td>
</tr>
<tr>
<td>Error</td>
<td>6</td>
<td>18</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>118.25</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14) The following table entries are test scores for males and females at different times of day. Assuming no effect from the interaction between gender and test time, test the claim that males and females perform the same on the test. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>6 a.m. – 9 a.m.</th>
<th>9 a.m. – 12 p.m.</th>
<th>12 p.m. – 3 p.m.</th>
<th>3 p.m. – 6 p.m.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>87</td>
<td>89</td>
<td>92</td>
</tr>
<tr>
<td>Female</td>
<td>72</td>
<td>84</td>
<td>94</td>
</tr>
</tbody>
</table>

Source | DF | SS | MS  | F  | p  |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>24.5</td>
<td>24.5</td>
<td>0.6652</td>
<td>0.4745</td>
</tr>
<tr>
<td>Time</td>
<td>3</td>
<td>183</td>
<td>61</td>
<td>1.6561</td>
<td>0.3444</td>
</tr>
<tr>
<td>Error</td>
<td>3</td>
<td>110.5</td>
<td>36.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>318</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Provide an appropriate response.

15) The following results are from a statistics software package in which all of the F values and P-values are given. Is there a significant effect from the interaction? Should you test to see if there is a significant effect due to either A or B? If the answer is yes, is there a significant effect due to either A or B?

ANOVA Table

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum squares</th>
<th>Mean square</th>
<th>F test</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>164.020</td>
<td>82.010</td>
<td>25.010</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>230.786</td>
<td>57.697</td>
<td>18.002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Interaction</td>
<td>8</td>
<td>80.879</td>
<td>10.110</td>
<td>3.154</td>
<td>.0031</td>
</tr>
<tr>
<td>Error</td>
<td>101</td>
<td>323.708</td>
<td>3.205</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
<td>799.393</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 12 FORM A

Solve the problem.

16) The following data shows the yield, in bushels per acre, categorized according to three varieties of corn and three different soil conditions. Test the null hypothesis of no interaction between variety and soil conditions at a significance level of 0.05.

<table>
<thead>
<tr>
<th></th>
<th>Plot 1</th>
<th>Plot 2</th>
<th>Plot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety 1</td>
<td>156, 167, 160, 145, 151, 170, 162, 169, 168, 148, 155</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety 2</td>
<td>172, 176, 179, 186, 161, 162, 166, 179, 160, 176, 165, 170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety 3</td>
<td>175, 157, 178, 170, 169, 165, 179, 178, 172, 174, 170, 169</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17) The following data contains task completion times, in minutes, categorized according to the gender of the machine operator and the machine used.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

Assume that two-way ANOVA is used to analyze the data. How are the ANOVA results affected if the times are converted to hours?
CHAPTER 12 FORM A

Solve the problem.

18) The following data contains task completion times, in minutes, categorized according to the gender of the machine operator and the machine used.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

The ANOVA results lead us to conclude that the completion times are not affected by an interaction between machine and gender, and the times are not affected by gender, but they are affected by the machine. Change the table entries so that there is no effect from the interaction between machine and gender, but there is an effect from the gender of the operator.

Provide an appropriate response.

19) What is a Bonferroni test? Why do we use it?

20) Explain why in two-way ANOVA there cannot be an interaction for sample data with one observation per cell.
1) 1) The populations have normal distributions.
   2) The populations have the same variance $\sigma^2$ (or standard deviation $\sigma$).
   3) The samples are simple random.
   4) The samples are independent of each other.
   5) The different samples are from populations that are categorized in only one way.
   (The requirements of normality and equal variances are somewhat relaxed.)
2) A treatment (also known as a factor) is a property or characteristic that allows us to distinguish
   the different populations from one another. Examples will vary.
   Possible example: Engine type distinguished according to three models is a factor, whose
   associated experiment is to determine fuel efficiency.
3) B
4) A
   Reject the claim of equal means. The different brands do not appear to have the same mean.
   [Optional: $P$-value: $p < 0.001$.]
   Fail to reject the claim of equal means. The data do not provide sufficient evidence to conclude
   that there is a difference in the mean lifetimes of the three brands of ballpoint pen.
   [Optional: $P$-value: $p = 0.2507$ STATDISK]
7) $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$. $H_1 :$ The means are not all equal.
   Test statistic: $F = 0.0555$. Critical value: $F = 3.9539$.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the
   claim that the four brands have the same mean.
   [Optional: $P$-value: $p = 0.9822$ STATDISK]
8) A
9) i) Since $t = 1.315 < t_{0.025}(8) = 2.306$, accept the claim that the two samples come from
   populations with the same means.
   [Optional: $P$-value: $p = 0.2258$ STATDISK, TI-83/84 Plus]
   ii) Since $F = 1.73 < F_{0.05}(1, 8) = 5.32$, accept the claim that the two samples come from
   populations with the same means.
   iii) $t^2 = 1.315^2 = 1.73 = F$ and $t_{0.025}(8) = 2.306^2 = 5.32 = F_{0.05}(1, 8)$.
10) $H_0$: There is no machine effect.
    $H_1$: There is a machine effect.
    Fail to reject the null hypothesis. The type of machine does not appear to have an effect on the
    number of items produced.
    [Optional: Test statistic $F = 3.1837$, $P$-value: $p = 0.0655$ STATDISK]
11) $H_0$: There is no machine effect.
    $H_1$: There is a machine effect.
    Reject the null hypothesis. There does appear to be a machine effect.
    [Optional: Test statistic $F = 53.7027$, $P$-value: $p = 0.0000$ STATDISK]
12) $H_0$: There is no interaction effect.
    $H_1$: There is an interaction effect.
    Reject the null hypothesis. There does appear to be an interaction effect.
    [Optional: $P$-value: $p = 0.0000$ STATDISK]
13) \( H_0 \): There is no track effect. \( H_1 \): There is a track effect. The \( P \)-value is 0.00763, which is less than 0.05. We reject the null hypothesis; it appears that the track does effect the racing times.

14) \( H_0 \): There is no gender effect. \( H_1 \): There is a gender effect. The \( P \)-value is 0.4745, which is greater than 0.05. We fail to reject the null hypothesis; it appears that the scores are not affected by gender.

15) Since \( P = 0.0031 \) for the interaction, you reject the null hypothesis that there is no effect due to the interaction. No, it is not appropriate to see if there is a significant effect due to either A or B. Do not consider the effects of either factor without considering the effects of the other.

16) \( H_0 \): There is no interaction between variety and soil conditions. \( H_1 \): There is an interaction between variety and soil conditions. The test statistic is \( F = 0.803973 \), and the corresponding \( P \)-value is 0.533402. Because the \( P \)-value is greater than 0.05, we fail to reject the null hypothesis of no interaction between variety and soil conditions.

17) The ANOVA results are not affected by converting the times to hours. The null hypothesis of no interaction between machine and gender is not rejected since the \( P \)-value is 0.946. The null hypothesis of no effect from machine is rejected since the \( P \)-value is 0.013. The null hypothesis of no effect from gender is not rejected since the \( P \)-value is 0.382.

18) The following table is one example of entries that produce an effect from the gender of the operator.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>14, 14</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 18</td>
<td>10, 10</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 16</td>
<td>12, 12</td>
</tr>
</tbody>
</table>

19) The Bonferroni test is one of several tests of multiple comparisons designed to determine which population mean or means is different from the others. This test is used after a significant \( F \) ratio results from an ANOVA procedure.

20) Variation within each cell of a two-way design does not exist when there is only one observation per cell. Therefore, the measures dependent on cell variation, such as the mean sum of squares, cannot be calculated for an interaction between the two variables.
CHAPTER 12 FORM B

Name:_________________________ Course Number: _________ Section Number: ______

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Describe the null and alternative hypotheses for one-way ANOVA. Give an example.

2) The test statistic for one-way ANOVA is \( F = \frac{\text{variance between samples}}{\text{variance within samples}} \). Describe variance within samples and variance between samples. What relationship between variance within samples and variance between samples would result in the conclusion that the value of \( F \) is significant?

Given below are the analysis of variance results from a Minitab display. Assume that you want to use a 0.05 significance level in testing the null hypothesis that the different samples come from populations with the same mean.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>13.500</td>
<td>4.500</td>
<td>5.17</td>
<td>0.011</td>
</tr>
<tr>
<td>Error</td>
<td>16</td>
<td>13.925</td>
<td>0.870</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>27.425</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Identify the p-value.

A) 5.17  B) 4.500  C) 0.870  D) 0.011

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>30</td>
<td>10.00</td>
<td>1.6</td>
<td>0.264</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>50</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What can you conclude about the equality of the population means?

A) Accept the null hypothesis since the p-value is greater than the significance level.
B) Reject the null hypothesis since the p-value is less than the significance level.
C) Reject the null hypothesis since the p-value is greater than the significance level.
D) Accept the null hypothesis since the p-value is less than the significance level.
CHAPTER 12 FORM B

Test the claim that the samples come from populations with the same mean. Assume that the populations are normally distributed with the same variance.

5) Given the sample data below, test the claim that the populations have the same mean. Use a significance level of 0.05.

<table>
<thead>
<tr>
<th></th>
<th>Brand A</th>
<th>Brand B</th>
<th>Brand C</th>
<th>Brand D</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n)</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>(\bar{x})</td>
<td>2.09</td>
<td>3.48</td>
<td>1.86</td>
<td>2.84</td>
</tr>
<tr>
<td>(s)</td>
<td>0.37</td>
<td>0.61</td>
<td>0.45</td>
<td>0.53</td>
</tr>
</tbody>
</table>

6) The data below represent the weight losses for people on three different exercise programs.

<table>
<thead>
<tr>
<th>Exercise A</th>
<th>Exercise B</th>
<th>Exercise C</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>5.8</td>
<td>4.3</td>
</tr>
<tr>
<td>8.8</td>
<td>4.9</td>
<td>6.2</td>
</tr>
<tr>
<td>7.3</td>
<td>1.1</td>
<td>5.8</td>
</tr>
<tr>
<td>9.8</td>
<td>7.8</td>
<td>8.1</td>
</tr>
<tr>
<td>5.1</td>
<td>1.2</td>
<td>7.9</td>
</tr>
</tbody>
</table>

At the 1% significance level, does it appear that a difference exists in the true mean weight loss produced by the three exercise programs?

7) At the 0.025 significance level, test the claim that the three brands have the same mean if the following sample results have been obtained.

<table>
<thead>
<tr>
<th></th>
<th>Brand A</th>
<th>Brand B</th>
<th>Brand C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x)</td>
<td>32</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>(x)</td>
<td>34</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>(x)</td>
<td>37</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>(x)</td>
<td>33</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>(x)</td>
<td>36</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>(x)</td>
<td>39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

233
Provide an appropriate response.

8) Fill in the missing entries in the following partially completed one-way ANOVA table.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS=SS/df</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>54</td>
<td>28.7</td>
<td>0.53</td>
<td>369.00</td>
</tr>
<tr>
<td>Error</td>
<td>25</td>
<td>105.0</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>133.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS=SS/df</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>4</td>
<td>28.7</td>
<td>7.17</td>
<td>1.71</td>
</tr>
<tr>
<td>Error</td>
<td>25</td>
<td>105.0</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>28.87</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS=SS/df</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>4</td>
<td>28.7</td>
<td>7.17</td>
<td>0.59</td>
</tr>
<tr>
<td>Error</td>
<td>25</td>
<td>105.0</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>133.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS=SS/df</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>4</td>
<td>28.7</td>
<td>7.17</td>
<td>1.71</td>
</tr>
<tr>
<td>Error</td>
<td>25</td>
<td>105.0</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>133.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D)
CHAPTER 12 FORM B

Solve the problem.

9) Use the data given below to verify that the t test for independent samples and the ANOVA method are equivalent.

\[
\begin{array}{c|c}
A & B \\
10 & 19 \\
29 & 18 \\
11 & 27 \\
19 & 30 \\
15 & 18 \\
16 & 21 \\
\end{array}
\]

i) Use a t test with a 0.05 significance level to test the claim that the two samples come from populations with the same means.

ii) Use the ANOVA method with a 0.05 significance level to test the same claim.

iii) Verify that the squares of the t test statistic and the critical value are equal to the F test statistic and critical value.

Use the Minitab display to test the indicated claim.

10) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

\[
\text{Employee} \\
\begin{array}{c|c|c|c|c|c|c|c}
 & A & B & C \\
I & 23, 27, 29 & 30, 27, 25 & 18, 20, 22 \\
II & 25, 26, 24 & 24, 29, 26 & 19, 16, 14 \\
III & 28, 25, 26 & 25, 27, 23 & 15, 11, 17 \\
\end{array}
\]

ANALYSIS OF VARIANCE ITEMS

\[
\begin{array}{llll}
\text{SOURCE} & \text{DF} & \text{SS} & \text{MS} \\
\text{MACHINE} & 2 & 34.67 & 17.33 \\
\text{EMPLOYEE} & 2 & 504.67 & 252.33 \\
\text{INTERACTION} & 4 & 26.67 & 6.67 \\
\text{ERROR} & 18 & 98.00 & 5.44 \\
\text{TOTAL} & 26 & 664.00 & \\
\end{array}
\]

Using a 0.05 significance level, test the claim that the interaction between employee and machine has no effect on the number of items produced.
CHAPTER 12 FORM B

Use the Minitab display to test the indicated claim.

11) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>16, 18, 19</td>
<td>15, 17, 20</td>
<td>14, 18, 16</td>
</tr>
<tr>
<td>Machine</td>
<td>II</td>
<td>20, 27, 29</td>
<td>25, 28, 27</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>15, 18, 17</td>
<td>16, 16, 19</td>
</tr>
</tbody>
</table>

ANALYSIS OF VARIANCE ITEMS

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>588.74</td>
<td>294.37</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>2.07</td>
<td>1.04</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>15.48</td>
<td>3.87</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>98.67</td>
<td>5.48</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>704.96</td>
<td></td>
</tr>
</tbody>
</table>

Assume that the number of items produced is not affected by an interaction between employee and machine. Using a 0.05 significance level, test the claim that the choice of employee has no effect on the number of items produced.

12) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>31, 34, 32</td>
<td>29, 23, 22</td>
<td>21, 20, 24</td>
</tr>
<tr>
<td>Machine</td>
<td>II</td>
<td>19, 26, 22</td>
<td>35, 33, 30</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>21, 18, 26</td>
<td>20, 23, 24</td>
</tr>
</tbody>
</table>

ANALYSIS OF VARIANCE ITEMS

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>1.19</td>
<td>.59</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>5.85</td>
<td>2.93</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>710.81</td>
<td>177.70</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>160.00</td>
<td>8.89</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>877.85</td>
<td></td>
</tr>
</tbody>
</table>

Assume that the number of items produced is not affected by an interaction between employee and machine. Using a 0.05 significance level, test the claim that the machine has no effect on the number of items produced.
CHAPTER 12 FORM B

Use the data in the given table and the corresponding Minitab display to test the hypothesis.

13) The following table entries are the times in seconds for three different drivers racing on four different tracks. Assuming no effect from the interaction between driver and track, test the claim that the three drivers have the same mean time. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>Track 1</th>
<th>Track 2</th>
<th>Track 3</th>
<th>Track 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver 1</td>
<td>72</td>
<td>70</td>
<td>68</td>
</tr>
<tr>
<td>Driver 2</td>
<td>74</td>
<td>71</td>
<td>66</td>
</tr>
<tr>
<td>Driver 3</td>
<td>76</td>
<td>69</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.33</td>
<td>0.729</td>
</tr>
<tr>
<td>Track</td>
<td>3</td>
<td>98.25</td>
<td>32.75</td>
<td>10.92</td>
<td>0.00763</td>
</tr>
<tr>
<td>Error</td>
<td>6</td>
<td>18</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>118.25</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14) The following Minitab display results from a study in which three different teachers taught calculus classes of five different sizes. The class average was recorded for each class. Assuming no effect from the interaction between teacher and class size, test the claim that the class size has no effect on the class average. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>2</td>
<td>56.93</td>
<td>28.47</td>
<td>1.018</td>
<td>0.404</td>
</tr>
<tr>
<td>Class Size</td>
<td>4</td>
<td>672.67</td>
<td>168.17</td>
<td>6.013</td>
<td>0.016</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>223.73</td>
<td>27.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>953.33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Provide an appropriate response.

15) The following results are from a statistics package in which all of the F values and P-values are given. Is there a significant effect from the interaction? Should you test to see if there is a significant effect due to either A or B? If the answer is yes, is there a significant effect due to either A or B?

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum squares</th>
<th>Mean square</th>
<th>F test</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>415.87305</td>
<td>207.93652</td>
<td>1.88259</td>
<td>.1637</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>2997.47186</td>
<td>999.15729</td>
<td>9.04603</td>
<td>.0001</td>
</tr>
<tr>
<td>Interaction</td>
<td>6</td>
<td>707.26626</td>
<td>117.87771</td>
<td>1.06723</td>
<td>.3958</td>
</tr>
<tr>
<td>Error</td>
<td>46</td>
<td>5080.81667</td>
<td>110.45254</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>9201.42784</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 12 FORM B

Solve the problem.

16) The following data shows the yield, in bushels per acre, categorized according to three varieties of corn and three different soil conditions. Assume that yields are not affected by an interaction between variety and soil conditions, and test the null hypothesis that variety has no effect on yield. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th></th>
<th>Plot 1</th>
<th>Plot 2</th>
<th>Plot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety 1</td>
<td>156, 167, 162, 160, 145, 151, 170, 162, 169, 168, 148, 155</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety 2</td>
<td>172, 176, 179, 186, 161, 162, 166, 179, 160, 176, 165, 170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety 3</td>
<td>175, 157, 178, 170, 169, 165, 179, 178, 172, 174, 170, 169</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17) The following data contains task completion times, in minutes, categorized according to the gender of the machine operator and the machine used.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

Assume that two-way ANOVA is used to analyze the data. How are the ANOVA results affected if 5 minutes is added to each completion time?
CHAPTER 12 FORM B

Solve the problem.

18) The following data contains task completion times, in minutes, categorized according to the gender of the machine operator and the machine used.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

The ANOVA results lead us to conclude that the completion times are not affected by an interaction between machine and gender, and the times are not affected by gender, but they are affected by the machine. Change the table entries so that there is an effect from the interaction between machine and gender.

Provide an appropriate response.

19) Why is it unnecessary to conduct multiple comparison tests after a nonsignificant F test statistic results?

20) What is a Tukey HSD test? Why is it used?
Answer Key
Testname: CHAPTER 12 FORM B

1) The null hypothesis for one-way ANOVA is that three or more means are equal. The alternative hypothesis is that the means are not all equal. Examples will vary.
   Possible example: Null hypothesis: mean mpg of three different car engines are the same.
   Alternative hypothesis: means are not all the same.
2) Variance between samples measures the variation between the sample means, that is the variation due to the treatment. The variance within the samples depends solely on the sample variances and is a measure of pooled variation. The F ratio compares the two. If the F ratio is relatively close to 1, the two variances are about the same, and we conclude that there are no significant differences among the sample means. When the value of F is excessively large (that is, greater than 1), we conclude that the variation among the samples is not the same and that the means are not equal.
3) D
4) A
   Reject the claim of equal means. The different brands do not appear to have the same mean.
   [Optional: $P$-value: $p < 0.0001$]
   Fail to reject the claim of equal means. The data do not provide sufficient evidence to conclude that there is a difference in the true mean weight loss produced by the three exercise programs.
   [Optional: $P$-value: $p = 0.264$ STATDISK]
7) $H_0 : \mu_1 = \mu_2 = \mu_3$. $H_1 :$ The means are not all equal.
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the three brands have the same mean.
   [Optional: Test statistic $F = 12.123$, $P$-value: $p = 0.0013$ STATDISK]
8) D
9) i) Since $t = 1.570 < t_{0.025}(10) = 2.228$, accept the claim that the two samples come from populations with the same means.
   [Optional: $P$-value: $p = 0.1500$ STATDISK, TI-83/84 Plus]
   ii) Since $F = 2.465 < F_{0.05}(1, 10) = 4.96$, accept the claim that the two samples come from populations with the same means.
   iii) $t^2 = 1.570^2 = 2.465 = F$ and $t_{0.025}(10) = 2 = 2.228^2 = 4.96 = F_{0.05}(1, 10)$.
10) $H_0$: There is no interaction effect.
    $H_1$: There is an interaction effect.
    Fail to reject the null hypothesis. There does not appear to be an interaction effect.
    [Optional: Test statistic $F = 1.22$, $P$-value: $p = 0.3353$ STATDISK]
11) $H_0$: There is no employee effect.
    $H_1$: There is an employee effect.
    Test statistic: $F = 0.1898$. Critical value: $F = 3.5546$.
    Fail to reject the null hypothesis. There does not appear to be an employee effect.
    [Optional: Test statistic $F = 0.1892$, $P$-value: $p = 0.8293$ STATDISK]
12) $H_0$: There is no machine effect.
    $H_1$: There is a machine effect.
    Test statistic: $F = 0.0664$. Critical value: $F = 3.5546$.
    Fail to reject the null hypothesis. The type of machine does not appear to have an effect on the number of items produced.
    [Optional: Test statistic $F = 0.0667$, $P$-value: $p = 0.9357$ STATDISK]

240
13) H₀: There is no driver effect. H₁: There is a driver effect. The P-value is 0.729, which is greater than 0.05. We fail to reject the null hypothesis; it appears that the driver does not affect the racing times.

14) H₀: There is no effect due to class size. H₁: There is an effect due to class size. The P-value is 0.016, which is less than 0.05. We reject the null hypothesis; it appears that the class size does affect the class average.

15) Since P = 0.3958 for the interaction, you fail to reject the null hypothesis that there is no effect due to the interaction. Yes, it is appropriate to see if there is a significant effect due to either A or B. The P-value for B is P = 0.0001, which rejects the null hypothesis that there is no effect due to B. The means for B are not all the same.

16) H₀: Variety has no effect on yield
H₁: Variety has an effect on yield
The test statistic is F = 13.54801, and the corresponding P-value is 0.000843. Because the P-value is less than 0.05, we reject the null hypothesis that variety has no effect on yield. It appears that the variety of corn does affect the yield.

17) The ANOVA results are not affected by adding 5 minutes to each completion time.
The null hypothesis of no interaction between machine and gender is not rejected since the p-value is 0.946. The null hypothesis of no effect from machine is rejected since the P-value is 0.013. The null hypothesis of no effect from gender is not rejected since the P-value is 0.382.

18) The following table is one example of entries that produce an effect from the interaction between machine and gender.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>12, 10</td>
</tr>
<tr>
<td>Machine 3</td>
<td>12, 10</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

19) A nonsignificant F ratio suggests that the population means are equal. Consequently, there is no need to determine which mean or means are different from the others.

20) The Tukey Honestly Significant Difference test is one of several multiple comparisons procedures designed to determine which mean or means are different from the others in ANOVA. It is used to find the source of a significant F test statistic.
CHAPTER 12 FORM C

Name: __________________________ Course Number: ________ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) When using statistical software packages, the critical value is typically not given. What method is used to determine whether you reject or fail to reject the null hypothesis?

2) Suppose you are to test for equality of four different means, with $H_0: \mu_A = \mu_B = \mu_C = \mu_D$. Write the hypotheses for the paired tests. Use methods of probability to explain why the process of ANOVA has a higher degree of confidence than testing each of the pairs separately.

3) Draw an example of an F distribution and list the characteristics of the F distribution.
CHAPTER 12 FORM C

Given below are the analysis of variance results from a Minitab display. Assume that you want to use a 0.05 significance level in testing the null hypothesis that the different samples come from populations with the same mean.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>13.500</td>
<td>4.500</td>
<td>5.17</td>
<td>0.011</td>
</tr>
<tr>
<td>Error</td>
<td>16</td>
<td>13.925</td>
<td>0.870</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>27.425</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Identify the value of the test statistic.
A) 4.500  B) 5.17  C) 0.011  D) 13.500

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>30</td>
<td>10.00</td>
<td>1.6</td>
<td>0.264</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>50</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Find the critical value.
A) 1.6  B) 8.85  C) 7.59  D) 4.07

Test the claim that the samples come from populations with the same mean. Assume that the populations are normally distributed with the same variance.

6) Random samples of four different models of cars were selected and the gas mileage of each car was measured. The results are shown below.

<table>
<thead>
<tr>
<th>Model A</th>
<th>Model B</th>
<th>Model C</th>
<th>Model D</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>28</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>24</td>
<td>29</td>
<td>32</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>30</td>
<td>27</td>
<td>28</td>
</tr>
</tbody>
</table>

Test the claim that the four different models have the same population mean. Use a significance level of 0.05.
CHAPTER 12 FORM C

Test the claim that the samples come from populations with the same mean. Assume that the populations are normally distributed with the same variance.

7) At the 0.025 significance level, test the claim that the four brands have the same mean if the following sample results have been obtained.

<table>
<thead>
<tr>
<th>Brand A</th>
<th>Brand B</th>
<th>Brand C</th>
<th>Brand D</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>18</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>21</td>
<td>23</td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>22</td>
<td>25</td>
<td>26</td>
<td>29</td>
</tr>
<tr>
<td>21</td>
<td>26</td>
<td>29</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37</td>
</tr>
</tbody>
</table>

Provide an appropriate response.

8) Four independent samples of 100 values each are randomly drawn from populations that are normally distributed with equal variances. You wish to test the claim that $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

i) If you test the individual claims $\mu_1 = \mu_2$, $\mu_1 = \mu_3$, $\mu_1 = \mu_4$, ..., $\mu_3 = \mu_4$, how many ways can you pair off the 4 means?

ii) Assume that the tests are independent and that for each test of equality between two means, there is a 0.99 probability of not making a type I error. If all possible pairs of means are tested for equality, what is the probability of making no type I errors?

iii) If you use analysis of variance to test the claim that $\mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 0.01 level of significance, what is the probability of not making a type I error?
CHAPTER 12 FORM C

Solve the problem.

9) At the same time each day, a researcher records the temperature in each of three greenhouses. The table shows the temperatures in degrees Fahrenheit recorded for one week.

<table>
<thead>
<tr>
<th>Greenhouse #1</th>
<th>Greenhouse #2</th>
<th>Greenhouse #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>71</td>
<td>67</td>
</tr>
<tr>
<td>72</td>
<td>69</td>
<td>63</td>
</tr>
<tr>
<td>73</td>
<td>72</td>
<td>62</td>
</tr>
<tr>
<td>66</td>
<td>72</td>
<td>61</td>
</tr>
<tr>
<td>68</td>
<td>65</td>
<td>60</td>
</tr>
<tr>
<td>71</td>
<td>73</td>
<td>62</td>
</tr>
<tr>
<td>72</td>
<td>71</td>
<td>59</td>
</tr>
</tbody>
</table>

i) Use a 0.05 significance level to test the claim that the average temperature is the same in each greenhouse.

ii) How are the analysis of variance results affected if the same constant is added to every one of the original sample values?

Use the Minitab display to test the indicated claim.

10) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>Machine I</th>
<th>Machine II</th>
<th>Machine III</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>23, 27, 29</td>
<td>25, 26, 24</td>
<td>28, 25, 26</td>
</tr>
<tr>
<td>B</td>
<td>30, 27, 25</td>
<td>24, 29, 26</td>
<td>25, 27, 23</td>
</tr>
<tr>
<td>C</td>
<td>18, 20, 22</td>
<td>19, 16, 14</td>
<td>15, 11, 17</td>
</tr>
</tbody>
</table>

ANALYSIS OF VARIANCE ITEMS

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine</td>
<td>2</td>
<td>34.67</td>
<td>17.33</td>
</tr>
<tr>
<td>Employee</td>
<td>2</td>
<td>504.67</td>
<td>252.33</td>
</tr>
<tr>
<td>Interaction</td>
<td>4</td>
<td>26.67</td>
<td>6.67</td>
</tr>
<tr>
<td>Error</td>
<td>18</td>
<td>98.00</td>
<td>5.44</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>26</td>
<td>664.00</td>
<td></td>
</tr>
</tbody>
</table>

Assume that the number of items produced is not affected by an interaction between employee and machine. Using a 0.05 significance level, test the claim that the choice of employee has no effect on the number of items produced.
CHAPTER 12 FORM C

Use the Minitab display to test the indicated claim.

11) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>16, 18, 19</td>
<td>15, 17, 20</td>
<td>14, 18, 16</td>
</tr>
<tr>
<td>Machine</td>
<td>II</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>20, 27, 29</td>
<td>25, 28, 27</td>
<td>29, 28, 26</td>
</tr>
<tr>
<td>III</td>
<td>15, 18, 17</td>
<td>16, 16, 19</td>
<td>13, 17, 16</td>
</tr>
</tbody>
</table>

**ANALYSIS OF VARIANCE ITEMS**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>588.74</td>
<td>294.37</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>2.07</td>
<td>1.04</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>15.48</td>
<td>3.87</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>98.67</td>
<td>5.48</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>704.96</td>
<td></td>
</tr>
</tbody>
</table>

Using a 0.05 significance level, test the claim that the interaction between employee and machine has no effect on the number of items produced.

12) A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and the Minitab results follow.

<table>
<thead>
<tr>
<th>Employee</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>31, 34, 32</td>
<td>29, 23, 22</td>
<td>21, 20, 24</td>
</tr>
<tr>
<td>Machine</td>
<td>II</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>19, 26, 22</td>
<td>35, 33, 30</td>
<td>25, 19, 23</td>
</tr>
<tr>
<td>III</td>
<td>21, 18, 26</td>
<td>20, 23, 24</td>
<td>36, 37, 31</td>
</tr>
</tbody>
</table>

**ANALYSIS OF VARIANCE ITEMS**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACHINE</td>
<td>2</td>
<td>1.19</td>
<td>.59</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>2</td>
<td>5.85</td>
<td>2.93</td>
</tr>
<tr>
<td>INTERACTION</td>
<td>4</td>
<td>710.81</td>
<td>177.70</td>
</tr>
<tr>
<td>ERROR</td>
<td>18</td>
<td>160.00</td>
<td>8.89</td>
</tr>
<tr>
<td>TOTAL</td>
<td>26</td>
<td>877.85</td>
<td></td>
</tr>
</tbody>
</table>

Assume that the number of items produced is not affected by an interaction between employee and machine. Using a 0.05 significance level, test the claim that the choice of employee has no effect on the number of items produced.
CHAPTER 12 FORM C

Use the data in the given table and the corresponding Minitab display to test the hypothesis.

13) The following Minitab display results from a study in which three different teachers taught calculus classes of five different sizes. The class average was recorded for each class. Assuming no effect from the interaction between teacher and class size, test the claim that the teacher has no effect on the class average. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>2</td>
<td>56.93</td>
<td>28.47</td>
<td>1.018</td>
<td>0.404</td>
</tr>
<tr>
<td>Class Size</td>
<td>4</td>
<td>672.67</td>
<td>168.17</td>
<td>6.013</td>
<td>0.016</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>223.73</td>
<td>27.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>953.33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14) The following table entries are test scores for males and females at different times of day. Assuming no effect from the interaction between gender and test time, test the claim that time of day does not affect test scores. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>Time</th>
<th>6 a.m. - 9 a.m.</th>
<th>9 a.m. - 12 p.m.</th>
<th>12 p.m. - 3 p.m.</th>
<th>3 p.m. - 6 p.m.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>87</td>
<td>89</td>
<td>92</td>
<td>85</td>
</tr>
<tr>
<td>Female</td>
<td>72</td>
<td>84</td>
<td>94</td>
<td>89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>24.5</td>
<td>24.5</td>
<td>0.6652</td>
<td>0.4745</td>
</tr>
<tr>
<td>Time</td>
<td>3</td>
<td>183</td>
<td>61</td>
<td>1.6561</td>
<td>0.3444</td>
</tr>
<tr>
<td>Error</td>
<td>3</td>
<td>110.5</td>
<td>36.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>318</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Provide an appropriate response.

15) The following data show annual income, in thousands of dollars, categorized according to the two factors of gender and level of education. Test the null hypothesis of no interaction between gender and level of education at a significance level of 0.05.

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th></th>
<th>Male</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High school</td>
<td>23, 27, 24, 26</td>
<td>25, 26, 22, 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>28, 36, 31, 33</td>
<td>35, 32, 39, 28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced degree</td>
<td>41, 38, 43, 49</td>
<td>35, 50, 47, 44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 12 FORM C

Solve the problem.

16) The following data shows the yield, in bushels per acre, categorized according to three varieties of corn and three different soil conditions. Assume that yields are not affected by an interaction between variety and soil conditions, and test the null hypothesis that soil conditions have no effect on yield. Use a 0.05 significance level.

<table>
<thead>
<tr>
<th>Variety 1</th>
<th>Plot 1</th>
<th>Plot 2</th>
<th>Plot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>156, 167</td>
<td>162, 160, 145, 151, 170, 162</td>
<td></td>
<td></td>
</tr>
<tr>
<td>169, 168</td>
<td>148, 155</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variety 2</th>
<th>Plot 1</th>
<th>Plot 2</th>
<th>Plot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>172, 176</td>
<td>179, 186, 161, 162, 166, 179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>160, 176</td>
<td>165, 170</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variety 3</th>
<th>Plot 1</th>
<th>Plot 2</th>
<th>Plot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>175, 157</td>
<td>178, 170, 169, 165, 179, 178</td>
<td></td>
<td></td>
</tr>
<tr>
<td>172, 174</td>
<td>170, 169</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17) The following data contains task completion times, in minutes, categorized according to the gender of the machine operator and the machine used.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

Assume that two-way ANOVA is used to analyze the data. How are the ANOVA results affected if the times are all doubled?

18) The following data contains task completion times, in minutes, categorized according to the gender of the machine operator and the machine used.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>14, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

The ANOVA results lead us to conclude that the completion times are not affected by an interaction between machine and gender, and the times are not affected by gender, but they are affected by the machine. Change the table entries so that there is no effect from the interaction between machine and gender, there is no effect from the machine used, and there is no effect from the gender of the operator.
CHAPTER 12 FORM C

Provide an appropriate response.

19) What two conditions are likely to result in a significant F test statistic in a one-way ANOVA experiment?

20) Why do researchers concentrate on explaining an interaction in two-way ANOVA rather than the effects of each factor separately?
1) The decision to reject or fail to reject is based on P-values. If the P-value is less than or equal to the significance level, you reject the null hypothesis. Otherwise you fail to reject.

2) The six paired hypotheses are \( \mu_A = \mu_B, \mu_A = \mu_C, \mu_A = \mu_D, \mu_B = \mu_C, \mu_B = \mu_D, \mu_C = \mu_D \).

Suppose we test each with a 5\% significance level (95\% confidence level). Then, the degree of confidence for all six would be 0.956 or 0.735, yielding an excessively high risk of a type I error.

ANOVA maintains the 5\% significance level while testing equivalence of all four means.

3) 1) The F distribution is not symmetric, it is skewed to the right.

2) The values of F can be 0 or positive, but they cannot be negative.

3) There is a different F distribution for each pair of degrees of freedom for the numerator and denominator.

General form of sketch: Smooth curve with peak up from 1 on the baseline and skewed to the right.

4) B

5) D

6) Test statistic: \( F = 6.435 \). Critical value: \( F = 3.4903 \).

Reject the claim of equal means. The different models do not appear to have the same mean.

[Optional: P-value: \( p = 0.0076 \) STATDISK]

7) \( H_0 \): \( \mu_1 = \mu_2 = \mu_3 = \mu_4 \). \( H_1 \): The means are not all equal.

Test statistic: \( F = 6.6983 \). Critical value: \( F = 3.9034 \).

Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the four brands have the same mean.

[Optional: Test statistic \( F = 6.699 \), P-value: \( p = 0.0028 \) STATDISK]

8) i) 6

ii) 0.996 = 0.9415

iii) 0.99

9) i) Reject the claim that the average temperature is the same in each greenhouse since

\[ F = 24.29 > F_{0.05}(2, 18) = 3.5546 \].

[Optional: P-value: \( p = 0.0000 \) STATDISK]

ii) The analysis of variance results are not affected.

10) \( H_0 \): There is no employee effect.

\( H_1 \): There is an employee effect.

Test statistic: \( F = 46.3842 \). Critical value: \( F = 3.5546 \).

Reject the null hypothesis. The employee does appear to have an effect on the number of items produced.

[Optional: Test statistic \( F = 46.3469 \), P-value: \( p = 0.0000 \) STATDISK]

11) \( H_0 \): There is no interaction effect.

\( H_1 \): There is an interaction effect.

Test statistic: \( F = 0.7062 \). Critical value: \( F = 2.9277 \).

Fail to reject the null hypothesis. There does not appear to be an interaction effect.

[Optional: Test statistic \( F = 0.71 \), P-value: \( p = 0.5981 \)]

12) \( H_0 \): There is no employee effect.

\( H_1 \): There is an employee effect.

Test statistic: \( F = 0.3296 \). Critical value: \( F = 3.5546 \).

Fail to reject the null hypothesis. There does not appear to be an employee effect.

[Optional: Test statistic \( F = 0.3292 \), P-value: \( p = 0.7238 \) STATDISK]

13) \( H_0 \): There is no teacher effect. \( H_1 \): There is a teacher effect. The P-value is 0.404, which is greater than 0.05. We fail to reject the null hypothesis; it appears that the teacher does not affect the class average.
14) \( H_0: \) There is no effect due to the time of day. \( H_1: \) There is an effect due to the time of day. The P-value is 0.3444, which is greater than 0.05. We fail to reject the null hypothesis; it appears that the scores are not affected by time of day.

15) \( H_0: \) There is no interaction between gender and level of education. \( H_1: \) There is an interaction between gender and level of education. The test statistic is \( F = 0.177472 \), and the corresponding P-value is 0.838832. Because the P-value is greater than 0.05, we fail to reject the null hypothesis of no interaction between gender and level of education.

16) \( H_0: \) Soil conditions have no effect on yield. \( H_1: \) Soil conditions have an effect on yield. The test statistic is \( F = 9.232917 \), and the corresponding P-value is 0.00088. Because the P-value is less than 0.05, we reject the null hypothesis that soil condition has no effect on yield. It appears that the soil condition does affect the yield.

17) The ANOVA results are not affected by doubling the completion times. The null hypothesis of no interaction between machine and gender is not rejected since the P-value is 0.946. The null hypothesis of no effect from machine is rejected since the P-value is 0.013. The null hypothesis of no effect from gender is not rejected since the P-value is 0.382.

18) The following table is one example of entries that produce no effect from interaction between gender and machine, no effect from machine, and no effect from gender.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine 1</td>
<td>15, 17</td>
<td>16, 17</td>
</tr>
<tr>
<td>Machine 2</td>
<td>17, 13</td>
<td>15, 13</td>
</tr>
<tr>
<td>Machine 3</td>
<td>16, 18</td>
<td>17, 19</td>
</tr>
</tbody>
</table>

19) Large sample sizes coupled with substantial variation between sample means could cause a significant F ratio, especially if the pooled variation within samples is not exceedingly high.

20) When an interaction occurs, the effect of one of the factors changes for the different categories of the other factor. Consequently, it makes sense to address how the two factors interact.
CHAPTER 13 FORM A

Name: ____________________________  Course Number: _________  Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Describe parametric and nonparametric tests. Explain why nonparametric tests are important.

2) Describe the Kruskal–Wallis test. What types of hypotheses is it used to test? What assumptions are made for this test?

3) Describe the rank correlation test. What types of hypotheses is it used to test? How does the rank correlation coefficient \( r_s \) differ from the Pearson correlation coefficient \( r \)?

Use the sign test to test the indicated claim.

4) An instructor gives a test before and after a lesson and results from randomly selected students are given below. At the 0.05 level of significance, test the claim that the lesson has no effect on the grade. Use the sign test.

<table>
<thead>
<tr>
<th>Before</th>
<th>54</th>
<th>61</th>
<th>56</th>
<th>41</th>
<th>38</th>
<th>57</th>
<th>42</th>
<th>71</th>
<th>88</th>
<th>42</th>
<th>36</th>
<th>23</th>
<th>46</th>
<th>51</th>
</tr>
</thead>
<tbody>
<tr>
<td>After</td>
<td>82</td>
<td>87</td>
<td>84</td>
<td>76</td>
<td>79</td>
<td>87</td>
<td>42</td>
<td>97</td>
<td>99</td>
<td>74</td>
<td>85</td>
<td>96</td>
<td>69</td>
<td>84</td>
</tr>
</tbody>
</table>
CHAPTER 13 FORM A

Use the sign test to test the indicated claim.

5) Fourteen people rated two brands of soda on a scale of 1 to 5.

| Brand A | 2 | 3 | 2 | 4 | 3 | 1 | 2 |
| Brand B | 1 | 4 | 5 | 5 | 1 | 2 | 3 |

| Brand A | 5 | 4 | 2 | 1 | 1 | 4 | 3 |
| Brand B | 4 | 5 | 5 | 2 | 4 | 5 | 4 |

At the 5 percent level, test the null hypothesis that the two brands of soda are equally popular.

6) A researcher wishes to study whether music has any effect on the ability to memorize information. 84 randomly selected adults are given a memory test in a quiet room. They are then given a second memory test while listening to classical music. 64 people received a higher score on the second test, 19 a lower score, and 1 received the same score. At the 0.05 significance level, test the claim that the music has no effect on memorization skills.

Use the Wilcoxon signed-ranks test to test the claim that both samples come from populations having the same distribution.

7) Use the Wilcoxon signed-ranks test and the sample data below. At the 0.05 significance level, test the claim that math and verbal scores are the same.

| Mathematics | 347 440 327 456 427 349 377 398 425 |
| Verbal      | 285 378 243 371 340 271 294 322 385 |

8) In a study of the effectiveness of physical exercise in weight reduction, 12 subjects followed a program of physical exercise for two months. Their weights (in pounds) before and after this program are shown in the table. Use a significance level of 0.05 to test the claim that the exercise program has no effect on weight.

| Before  | 162 190 188 152 148 127 195 164 175 156 180 136 |
| After   | 157 194 179 149 135 130 133 168 168 148 170 138 |
CHAPTER 13 FORM A

Use the Wilcoxon rank–sum test to test the claim that the two independent samples come from the same distribution.

9) A person who commutes to work is choosing between two different routes. He tries the first route 11 times and the second route 12 times and records the time of each trip. The results (in minutes) are shown below. Use a significance level of 0.01 to test the claim that the times for both routes have the same distribution.

<table>
<thead>
<tr>
<th>Route 1</th>
<th>Route 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>35 42 41 41 46 38</td>
<td></td>
</tr>
<tr>
<td>33 48 40 49 53 36</td>
<td></td>
</tr>
<tr>
<td>39 50 46 51 57 53</td>
<td></td>
</tr>
<tr>
<td>36 40 45 50 55</td>
<td></td>
</tr>
</tbody>
</table>

Solve the problem.

10) The Mann–Whitney U test is equivalent to the Wilcoxon rank–sum test for independent samples in the sense that they both apply to the same situations and always lead to the same conclusions. In the Mann–Whitney U test we calculate

\[ z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \]

where

\[ U = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R \]

For the sample data below, use the Mann–Whitney U test to test the null hypothesis that the two independent samples come from populations with the same distribution. State the hypotheses, the value of the test statistic, the critical values, and your conclusion.

Test scores (men): 70, 96, 77, 90, 81, 45, 55, 68, 74, 99, 88
Test scores (women): 89, 92, 60, 78, 84, 96, 51, 67, 85, 94

254
CHAPTER 13 FORM A

Use a Kruskal–Wallis test to test the claim that the samples come from identical populations.

11) Listed below are grade averages for randomly selected students with three different categories of high-school background. At the 0.05 level of significance, test the claim that the three groups come from identical populations.

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.21</td>
<td>2.87</td>
<td>2.01</td>
</tr>
<tr>
<td>2</td>
<td>3.65</td>
<td>3.05</td>
<td>2.31</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>2.00</td>
<td>2.98</td>
</tr>
<tr>
<td>4</td>
<td>3.12</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>5</td>
<td>2.75</td>
<td>1.98</td>
<td>2.36</td>
</tr>
</tbody>
</table>

Use the rank correlation coefficient to test the claim of no correlation between the two variables.

12) Given that the rank correlation coefficient, \( r_s \), for 71 pairs of data is -0.894, test the claim of no correlation between the two variables. Use a significance level of 0.05.

13) Ten luxury cars were ranked according to their comfort levels and their prices.

<table>
<thead>
<tr>
<th>Make</th>
<th>Comfort</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>I</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>J</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

Find the rank correlation coefficient and test the claim of no correlation between comfort and price. Use a significance level of 0.05.
CHAPTER 13 FORM A

Use the rank correlation coefficient to test the claim of no correlation between the two variables.

14) A placement test is required for students desiring to take a finite mathematics course at a university. The instructor of the course studies the relationship between students’ placement test score and final course score. A random sample of eight students yields the following data.

<table>
<thead>
<tr>
<th>Placement Score</th>
<th>Final Course Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>63</td>
</tr>
<tr>
<td>90</td>
<td>41</td>
</tr>
<tr>
<td>95</td>
<td>54</td>
</tr>
<tr>
<td>51</td>
<td>32</td>
</tr>
<tr>
<td>86</td>
<td>93</td>
</tr>
<tr>
<td>74</td>
<td>60</td>
</tr>
<tr>
<td>60</td>
<td>61</td>
</tr>
<tr>
<td>57</td>
<td>89</td>
</tr>
</tbody>
</table>

Compute the rank correlation coefficient, $r_s$, of the data and test the claim of no correlation between placement score and final course score. Use a significance level of 0.05.

15) A college administrator collected information on first-semester night-school students. A random sample taken of 12 students yielded the following data on age and GPA during the first semester.

<table>
<thead>
<tr>
<th>Age</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$y$</td>
</tr>
<tr>
<td>18</td>
<td>1.2</td>
</tr>
<tr>
<td>26</td>
<td>3.8</td>
</tr>
<tr>
<td>27</td>
<td>2.0</td>
</tr>
<tr>
<td>37</td>
<td>3.3</td>
</tr>
<tr>
<td>33</td>
<td>2.5</td>
</tr>
<tr>
<td>47</td>
<td>1.6</td>
</tr>
<tr>
<td>20</td>
<td>1.4</td>
</tr>
<tr>
<td>48</td>
<td>3.6</td>
</tr>
<tr>
<td>50</td>
<td>3.7</td>
</tr>
<tr>
<td>38</td>
<td>3.4</td>
</tr>
<tr>
<td>34</td>
<td>2.7</td>
</tr>
<tr>
<td>22</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Do the data provide sufficient evidence to conclude that the variables age, $x$, and GPA, $y$, are correlated? Apply a rank-correlation test. Use $\alpha = 0.05$. 


CHAPTER 13 FORM A

Use the runs test to determine whether the given sequence is random. Use a significance level of 0.05.

16) The outcomes (odd number or even number) of a roulette wheel are shown below. Test for randomness of odd (O) and even (E) numbers.

O E O E O E O E O E E
E E O E O O E O E O E E

17) A sample of 15 clock radios is selected in sequence from an assembly line. Each radio is examined and judged to be acceptable (A) or defective (D). The results are shown below. Test for randomness.

D D A A A
A A A A A
A A D D D

18) Test the sequence of digits below for randomness of odd and even digits.

0 4 7 3 6 0 9 7 4 8
7 2 8 5 7 3 9 6 4 6
4 7 9 1 6 1 9 5 8 3
7 8 5 7 3 5 2 9 3 8

Provide an appropriate response.

19) Which of the following methods could lead to stronger evidence for the outcome of a nonparametric test?

A) ensure that the distributions are normal
B) ensure that the populations have equal variances
C) increase sample size substantially
D) take multistage random samples

20) Which of the following distribution-free tests has no parametric counterpart?

A) sign test
B) runs test
C) Kruskal–Wallis test
D) rank correlation test
1) Parametric tests require assumptions about the nature or shape of the populations involved. Most of the tests we have worked with have required that the populations be normal. Nonparametric tests do not have requirements as to parent population.

2) The Kruskal–Wallis test is used to test claims about the differences in means among several samples, as opposed to the Wilcoxon rank-sum test which looks at claims for two independent samples. The assumptions include: there are at least three random samples; we want to test the null hypothesis that the samples come from the same or identical populations; and each sample has at least five observations. The Kruskal–Wallis test also sums ranks for the sample data ranked as a whole.

3) The rank correlation test uses ranks to measure the strength of the relation between two variables. The rank correlation procedure is used to test the null hypothesis that there is no correlation between the two variables. The Pearson correlation coefficient r detects linear relationships between two variables. The rank correlation rs, also known as Spearman's rank correlation coefficient, detects relationships which are non-linear as well as linear.

4) H0: There is no difference between before and after grades.
   H1: There is a difference between before and after grades.
   Test statistic: \( t = 0 \). Critical value: \( t = 2 \).
   Reject the null hypothesis of no difference. There is sufficient evidence to warrant rejection of the claim that the lesson has no effect on grade.

5) H0: The two brands of soda are equally popular.
   H1: The two brands of soda are not equally popular.
   Test statistic: \( x = 3 \). Critical value: \( x = 2 \).
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the two brands are equally popular.

6) H0: the music has no effect on memorization skills.
   H1: the music has an effect on memorization skills.
   Convert \( x = 19 \) to the test statistic \( z = -4.83 \). Critical values: \( z = \pm 1.96 \).
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that music has no effect on memorization skills.

7) Test statistic \( T = 0 \). Critical value: \( T = 6 \).
   Reject the null hypothesis that the samples of math and verbal scores come from populations having the same distribution.

8) Test statistic \( T = 12.5 \). Critical value: \( T = 14 \).
   Reject the null hypothesis that the samples of before and after weights come from the same population distribution.

9) \( \mu_R = 132, \sigma_R = 16.2481, R = 94, z = 2.34 \).
   Test statistic: \( z = 2.34 \). Critical values \( z = \pm 2.575 \).
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the samples of times for commuting come from the same distribution.

10) H0: The two samples come from populations with the same distribution.
    H1: The two samples come from populations with different distributions.
    Critical values \( z = \pm 1.96 \), \( R = 115.5 \).
    Test statistic: \( z = 0.39 \)
    Do not reject the null hypothesis. There is not sufficient evidence to reject the claim that the two samples of test scores come from populations with the same distribution.
11) H_0: The three populations are identical.
   H_1: The three populations are not identical.
   Test statistic: H = 2.9600. Critical value is 5.991.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the three groups of grade averages come from identical populations.
12) r_s = -0.894. Critical values: r_s = ±0.234.
   Significant correlation. There appears to be a correlation between the two variables.
13) r_s = -0.285. Critical values: r_s = ±0.648.
   No significant correlation. There does not appear to be a correlation between comfort and price.
14) r_s = -0.167. Critical values: r_s = ±0.738.
   No significant correlation. There does not appear to be a correlation between placement score and final course score.
15) r_s = 0.532. Critical values: r_s = ±0.587.
   No significant correlation. The data do not provide sufficient evidence to indicate that age and GPA are correlated.
16) n_1 = 10, n_2 = 14, G = 18, 5% cutoff values: 7, 18.
   Reject the null hypothesis of randomness.
17) n_1 = 5, n_2 = 10, G = 3, 5% cutoff values: 3, 12.
   Reject the null hypothesis of randomness.
18) n_1 = 17, n_2 = 23, G = 19, \( \mu_G = 20.550 \), \( \sigma_G = 3.0494 \).
   Test statistic: z = -0.51. Critical values: z = ±1.96.
   Fail to reject the null hypothesis of randomness.
19) C
20) B
CHAPTER 13 FORM B

Name: ___________________________ Course Number: ________ Section Number: ________

Directions: Write your answers to the short-answer items in the spaces provided. Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) List the advantages and disadvantages of nonparametric tests.

2) Describe the Wilcoxon signed-ranks test. What type of hypotheses is it used to test? What assumptions are made for this test?

3) Describe the Wilcoxon rank-sum test. What type of hypotheses is it used to test? What assumptions are made for this test? What is the underlying concept?

Use the sign test to test the indicated claim.

4) A standard aptitude test is given to several randomly selected programmers, and the scores are given below for the mathematics and verbal portions of the test. Use the sign test to test the claim that programmers do better on the mathematics portion of the test. Use a 0.05 level of significance.

| Mathematics  | 347 440 327 456 427 349 377 398 425 |
| Verbal       | 285 378 243 371 340 271 294 322 385 |
CHAPTER 13 FORM B

Use the sign test to test the indicated claim.

5) A researcher wishes to test whether a particular diet has an effect on blood pressure. The blood pressure of 25 randomly selected adults is measured. After one month on the diet, each person’s blood pressure is again measured. For 19 people, the second blood pressure reading was lower than the first, and for 6 people, the second blood pressure reading was higher than the first. At the 0.01 significance level, test the claim that the diet has an effect on blood pressure.

6) The waiting times (in minutes) of 28 randomly selected customers in a bank are given below. Use a significance level of 0.05 to test the claim that the population median is equal to 5.3 minutes.

| 8.2 | 8.0 | 10.5 | 3.8 | 6.4 | 5.3 | 7.8 |
| 2.9 | 6.0 | 7.7 | 6.1 | 5.9 | 1.2 | 10.4 |
| 7.3 | 6.9 | 5.8 | 5.1 | 6.2 | 3.1 | 5.8 |
| 11.7 | 4.5 | 6.5 | 9.8 | 7.4 | 2.3 | 7.8 |

Use the Wilcoxon signed-ranks test to test the claim that both samples come from populations having the same distribution.

7) An instructor gives a test before and after a lesson and results from randomly selected students are given below. At the 0.05 level of significance, test the claim that the lesson has no effect on the grade. Use Wilcoxon’s signed-ranks test.

| Before | 54 | 61 | 56 | 41 | 38 | 57 | 42 | 71 | 88 | 42 | 36 | 23 | 22 | 46 | 51 |
| After | 82 | 87 | 84 | 76 | 79 | 87 | 42 | 97 | 99 | 74 | 85 | 96 | 67 | 84 | 79 |

8) 11 runners are timed at the 100-meter dash and are timed again one month later after following a new training program. The times (in seconds) are shown in the table. Use a significance level of 0.05 to test the claim that the training has no effect on the times.

| Before | 12.1 | 12.4 | 11.7 | 11.5 | 11.0 | 11.8 | 12.3 | 10.8 | 12.6 | 12.7 | 10.7 |
| After | 11.9 | 12.4 | 11.8 | 11.4 | 11.2 | 11.5 | 12.0 | 10.9 | 12.0 | 12.2 | 11.1 |
CHAPTER 13 FORM B

Use the Wilcoxon rank-sum test to test the claim that the two independent samples come from the same distribution.

9) 11 female employees and 11 male employees are randomly selected from one company and their weekly salaries are recorded. The salaries (in dollars) are shown below. Use a significance level of 0.10 to test the claim that salaries for female and male employees of the company have the same distribution.

<table>
<thead>
<tr>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>350</td>
<td>410</td>
</tr>
<tr>
<td>420</td>
<td>460</td>
</tr>
<tr>
<td>470</td>
<td>650</td>
</tr>
<tr>
<td>385</td>
<td>545</td>
</tr>
<tr>
<td>675</td>
<td>720</td>
</tr>
<tr>
<td>520</td>
<td>810</td>
</tr>
<tr>
<td>540</td>
<td>660</td>
</tr>
<tr>
<td>400</td>
<td>500</td>
</tr>
<tr>
<td>550</td>
<td>880</td>
</tr>
<tr>
<td>450</td>
<td>700</td>
</tr>
<tr>
<td>640</td>
<td>750</td>
</tr>
</tbody>
</table>

Solve the problem.

10) When performing a rank correlation test, one alternative to using Table A-9 to find critical values is to compute them using this approximation:

\[ r_s = \pm \sqrt{\frac{t^2}{t^2 + n - 2}} \]

where \( t \) is the \( t \)-score from Table A-3 corresponding to \( n - 2 \) degrees of freedom. Use this approximation to find critical values of \( r_s \) for the case where \( n = 40 \) and \( \alpha = 0.10 \).

A) \( \pm 0.202 \)  B) \( \pm 0.304 \)  C) \( \pm 0.264 \)  D) \( \pm 0.312 \)

Use a Kruskal–Wallis test to test the claim that the samples come from identical populations.

11) A fire-science specialist tests three different brands of flares for their burning times (in minutes) and the results are given below for the sample data. At the 0.05 significance level, test the claim that the three brands have the same mean burn time. Use the Kruskal–Wallis test.

<table>
<thead>
<tr>
<th>Brand X</th>
<th>16.4</th>
<th>17.6</th>
<th>18.3</th>
<th>17.0</th>
<th>17.1</th>
<th>17.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand Y</td>
<td>17.9</td>
<td>18.0</td>
<td>17.8</td>
<td>18.4</td>
<td>17.6</td>
<td>19.0</td>
</tr>
<tr>
<td>Brand Z</td>
<td>17.3</td>
<td>16.4</td>
<td>16.5</td>
<td>16.0</td>
<td>15.8</td>
<td>16.3</td>
</tr>
</tbody>
</table>

262
CHAPTER 13 FORM B

Use a Kruskal–Wallis test to test the claim that the samples come from identical populations.

12) SAT scores for students selected randomly from three different schools are shown below. Use a significance level of 0.05 to test the claim that the samples come from identical populations.

<table>
<thead>
<tr>
<th>School A</th>
<th>School C</th>
<th>School B</th>
</tr>
</thead>
<tbody>
<tr>
<td>550 480 670</td>
<td>500 620 700</td>
<td>460 580 620</td>
</tr>
<tr>
<td>400 600 520</td>
<td>550 760</td>
<td>380 600 470</td>
</tr>
<tr>
<td></td>
<td></td>
<td>450</td>
</tr>
</tbody>
</table>

Use the rank correlation coefficient to test the claim of no correlation between the two variables.

13) Given that the rank correlation coefficient, \( r_s \), for 20 pairs of data is 0.635, test the claim of no correlation between the two variables. Use a significance level of 0.05.

14) Ten trucks were ranked according to their comfort levels and their prices.

<table>
<thead>
<tr>
<th>Make</th>
<th>Comfort</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>G</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>H</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>I</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>J</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

Find the rank correlation coefficient and test the claim of no correlation between comfort and price. Use a significance level of 0.05.

15) Use the sample data below to find the rank correlation coefficient and test the claim of no correlation between math and verbal scores. Use a significance level of 0.05.

<table>
<thead>
<tr>
<th>Mathematics</th>
<th>347 440 327 456 427 349 377 398 425</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal</td>
<td>285 378 243 371 340 271 294 322 385</td>
</tr>
</tbody>
</table>

263
CHAPTER 13 FORM B

Use the runs test to determine whether the given sequence is random. Use a significance level of 0.05.

16) A true–false test had the following answer sequence.

\[
\begin{align*}
T & T T F T F T F T T F T \\
T & T F F F F F F T T F T F
\end{align*}
\]

Test the null hypothesis that the sequence was random.

17) Answers to a questionnaire were in the following sequence. Test for randomness.

\[
\begin{align*}
Y & Y N Y N N N N Y Y \\
N & N N N N Y Y Y N N N
\end{align*}
\]

18) A sample of 30 clock radios is selected in sequence from an assembly line. Each radio is examined and judged to be acceptable (A) or defective (D). The results are shown below. Test for randomness.

\[
\begin{align*}
A & A D A A A D A A D \\
A & D A A A D A A A A \\
A & A D D A A A A D A
\end{align*}
\]

Provide an appropriate response.

19) Which of the following nonparametric tests reaches a conclusion equivalent to the Mann–Whitney U test?

A) sign test \hspace{1cm} B) Wilcoxon rank–sum test
C) Wilcoxon signed–ranks test \hspace{1cm} D) Kruskal–Wallis test

20) Which of the following distribution–free tests has the lowest efficiency rating compared to its parametric counterpart?

A) Kruskal–Wallis test \hspace{1cm} B) Wilcoxon signed–ranks test
C) Wilcoxon rank–sum test \hspace{1cm} D) rank correlation test
1) Advantages: 1) Nonparametric methods can be applied to a wide variety of situations because they do not have the rigid requirements of their parametric counterparts. In particular, nonparametric tests do not require normally distributed populations. 2) Nonparametric tests can often be applied to nonnumerical data. 3) Nonparametric methods usually involve simpler computations than the corresponding parametric methods. Disadvantages: 1) Nonparametric methods tend to waste information because exact numerical data are reduced to a qualitative form. 2) Nonparametric tests are not as efficient as parametric tests so we generally need stronger evidence (such as a larger sample or a greater difference) before we reject a null hypothesis.

2) The Wilcoxon signed-ranks test is similar to the sign test, but it looks at the magnitude as well as the signs of the differences, and thus has a higher efficiency level than the signs test. The test is used to test claims that matched pairs have differences that come from a population whose median is 0. The Wilcoxon signed-ranks test assumes that the population of the differences (found from the pairs of data) has a distribution that is approximately symmetric.

3) This test is used to test claims about equal medians of two independent populations. The assumptions include: two independent samples; testing the null hypothesis that the two independent samples come from the same distribution; and more than 10 scores in each of the samples. The underlying principle is that if two samples are drawn from identical populations and the individual scores are all ranked as one combined collection of values, then the high and low ranks should fall evenly between the two samples. For example, if low ranks are found predominantly in one sample with the high ranks in the other, then we suspect that the two samples are from populations with different medians.

4) H₀: The math scores are equal to or less than the verbal scores.
   H₁: The math scores are greater than the verbal scores.
   Test statistic: x = 0. Critical value: x = 1.
   Reject the null hypothesis. There is sufficient evidence to support the claim that the math scores are greater than the verbal scores.

5) H₀: The diet does not have an effect on blood pressure.
   H₁: The diet has an effect on blood pressure.
   Test statistic: x = 6. Critical value: x = 5.
   Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the diet has an effect on blood pressure.

6) H₀: median is equal to 5.3 minutes.
   H₁: median is not equal to 5.3 minutes.
   Convert x = 7 to the test statistic z = -2.31. Critical values: z = ±1.96.
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the population median is equal to 5.3 minutes.

7) Test statistic T = 0. Critical value: T = 21.
   Reject the null hypothesis that both samples of test scores come from the same population distribution.

   Fail to reject the null hypothesis that both samples of running times come from the same population distribution.

9) \( \mu_R = 126.5, \sigma_R = 15.2288, R = 90, z = -2.40. \)
   Test statistic: z = -2.40. Critical values z = ±1.645.
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the two populations of salaries are identical.

10) C
Answer Key
Testname: CHAPTER 13 FORM B

11) \( H_0 \): The three brands have identical populations.
   \( H_1 \): The three populations are not identical.
   Test statistic: \( H = 12.7979 \). The critical value is 5.991.
   Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the
   three brands have the same mean burn time.
12) Test statistic: \( H = 3.6586 \). Critical value is 5.991.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the
   claim that the samples of SAT scores come from identical populations.
13) \( r_s = 0.635 \). Critical values: \( r_s = \pm 0.447 \).
   Significant correlation. There appears to be a correlation between the two variables.
14) \( r_s = 0.382 \). Critical values: \( r_s = \pm 0.648 \).
   No significant correlation. There does not appear to be a correlation between comfort and price.
15) \( r_s = 0.867 \). Critical values: \( r_s = \pm 0.700 \).
   Significant correlation. There appears to be a correlation between the two variables.
16) \( n_1 = 15, n_2 = 15, G = 18, 5\% \) cutoff values: 10, 22.
   Fail to reject the null hypothesis of randomness.
17) \( n_1 = 8, n_2 = 12, G = 8, 5\% \) cutoff values: 6, 16.
   Fail to reject the null hypothesis of randomness.
18) \( n_1 = 22, n_2 = 8, G = 15, \mu_G = 12.73, \sigma_G = 2.0839 \).
   Test statistic: \( z = 1.09 \). Critical values: \( z = \pm 1.96 \).
   Fail to reject the null hypothesis of randomness.
19) B
20) D
CHAPTER 13 FORM C

Name:__________________________ Course Number: ________ Section Number: ______

Directions: Write your answers to the short-answer items in the spaces provided.
Circle the correct choice for multiple-choice items.

Provide an appropriate response.

1) Define rank. Explain how to find the rank for data which repeats (for example, the data set: 4, 5, 5, 7, 8, 12, 12, 15, 18).

2) Describe the sign test. What types of hypotheses is it used to test? What is the underlying concept?

3) Describe the runs test for randomness. What types of hypotheses is it used to test? Does the runs test measure frequency? What is the underlying concept?

Use the sign test to test the indicated claim.

4) The heights of 16 randomly selected women are given below. Use a significance level of 0.05 to test the claim that the population median is equal to 64.0 inches.

<table>
<thead>
<tr>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>62.9</td>
<td>61.9</td>
<td>66.4</td>
<td>68.5</td>
<td>63.7</td>
<td>64.0</td>
<td>65.2</td>
<td>67.0</td>
</tr>
<tr>
<td>70.2</td>
<td>65.3</td>
<td>64.0</td>
<td>60.3</td>
<td>64.3</td>
<td>66.9</td>
<td>65.0</td>
<td>63.8</td>
</tr>
</tbody>
</table>
CHAPTER 13 FORM C

Use the sign test to test the indicated claim.

5) A researcher wishes to study whether a particular diet is effective in helping people to lose weight. 85 randomly selected adults were weighed before starting the diet and again after following the diet for one month. 47 people lost weight, 36 gained weight, and 2 observed no change in their weight. At the 0.01 significance level, test the claim that the diet is effective.

Use the Wilcoxon signed-ranks test to test the claim that both samples come from populations having the same distribution.

6) The systolic blood pressure readings of ten subjects before and after following a particular diet for a month are shown in the table. Use a significance level of 0.01 to test the claim that the diet has no effect on systolic blood pressure.

<table>
<thead>
<tr>
<th>Subject</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>175</td>
<td>192</td>
<td>167</td>
<td>180</td>
<td>161</td>
<td>203</td>
<td>185</td>
<td>176</td>
<td>204</td>
<td>146</td>
</tr>
<tr>
<td>After</td>
<td>160</td>
<td>190</td>
<td>170</td>
<td>180</td>
<td>153</td>
<td>197</td>
<td>191</td>
<td>174</td>
<td>192</td>
<td>150</td>
</tr>
</tbody>
</table>

Use the Wilcoxon rank-sum test to test the claim that the two independent samples come from the same distribution.

7) SAT scores for students selected randomly from two different schools are shown below. Use a significance level of 0.05 to test the claim that the scores for the two schools have the same distribution.

<table>
<thead>
<tr>
<th>School A</th>
<th>550 480 670</th>
<th>School B</th>
<th>460 580 620</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 700 520</td>
<td>380 680 570</td>
<td></td>
<td></td>
</tr>
<tr>
<td>540 740 560</td>
<td>660 500 480</td>
<td></td>
<td></td>
</tr>
<tr>
<td>360 560 650</td>
<td>600 550</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 13 FORM C

Use the Wilcoxon rank–sum test to test the claim that the two independent samples come from the same distribution.

8) Use the Wilcoxon rank–sum approach to test the claim that students at two colleges achieve the same distribution of grade averages. The sample data is listed below. Use a 0.05 level of significance.

College A 3.2 4.0 2.4 2.6 2.0 1.8 1.3 0.0 0.5 1.4 2.9
College B 2.4 1.9 0.3 0.8 2.8 3.0 3.1 3.1 3.5 3.5

Solve the problem.

9) The Wilcoxon signed–ranks test can be used to test the claim that a sample comes from a population with a specified median. The procedure used is the same as the one described in this section except that the differences are obtained by subtracting the value of the hypothesized median from each value.

The sample data below represent the weights (in pounds) of 12 women aged 20–30. Use a Wilcoxon signed–ranks test to test the claim that the median weight of women aged 20–30 is equal to 130 pounds. Use a significance level of 0.05. Be sure to state the hypotheses, the value of the test statistic, the critical values, and your conclusion.

140 116 125 120 153 140
111 127 133 137 132 160

Use a Kruskal–Wallis test to test the claim that the samples come from identical populations.

10) The table below shows the lifetimes (in hours) of random samples of light bulbs of three different brands. Use a 0.01 significance level to test the claim that the samples come from identical populations.

<table>
<thead>
<tr>
<th>Brand</th>
<th>Brand B</th>
<th>Brand C</th>
</tr>
</thead>
<tbody>
<tr>
<td>190</td>
<td>182</td>
<td>203</td>
</tr>
<tr>
<td>220</td>
<td>170</td>
<td>210</td>
</tr>
<tr>
<td>230</td>
<td>203</td>
<td>199</td>
</tr>
<tr>
<td>215</td>
<td>175</td>
<td>200</td>
</tr>
<tr>
<td>224</td>
<td>178</td>
<td>196</td>
</tr>
<tr>
<td>231</td>
<td>181</td>
<td>197</td>
</tr>
</tbody>
</table>
CHAPTER 13 FORM C

Use a Kruskal–Wallis test to test the claim that the samples come from identical populations.

11) The table below shows the weights (in pounds) of 6 randomly selected women in each of three different age groups. Use a 0.01 significance level to test the claim that the 3 age-group populations of weights are identical.

<table>
<thead>
<tr>
<th>18–34</th>
<th>35–55</th>
<th>56 and older</th>
</tr>
</thead>
<tbody>
<tr>
<td>119</td>
<td>123</td>
<td>140</td>
</tr>
<tr>
<td>134</td>
<td>147</td>
<td>128</td>
</tr>
<tr>
<td>114</td>
<td>135</td>
<td>59</td>
</tr>
<tr>
<td>125</td>
<td>110</td>
<td>134</td>
</tr>
<tr>
<td>153</td>
<td>154</td>
<td>120</td>
</tr>
<tr>
<td>138</td>
<td>163</td>
<td>116</td>
</tr>
</tbody>
</table>

Use the rank correlation coefficient to test the claim of no correlation between the two variables.

12) Given that the rank correlation coefficient, $r_s$, for 15 pairs of data is $-0.506$, test the claim of no correlation between the two variables. Use a significance level of 0.01.

13) Given that the rank correlation coefficient, $r_s$, for 31 pairs of data is 0.348, test the claim of no correlation between the two variables. Use a significance level of 0.01.

14) The scores of twelve students on the midterm exam and the final exam were as follows.

<table>
<thead>
<tr>
<th>Student</th>
<th>Midterm</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navarro</td>
<td>93</td>
<td>91</td>
</tr>
<tr>
<td>Reaves</td>
<td>89</td>
<td>85</td>
</tr>
<tr>
<td>Hurlburt</td>
<td>71</td>
<td>73</td>
</tr>
<tr>
<td>Knuth</td>
<td>65</td>
<td>77</td>
</tr>
<tr>
<td>Lengyel</td>
<td>62</td>
<td>67</td>
</tr>
<tr>
<td>Mommekean</td>
<td>74</td>
<td>79</td>
</tr>
<tr>
<td>Bolker</td>
<td>77</td>
<td>65</td>
</tr>
<tr>
<td>Ammato</td>
<td>87</td>
<td>83</td>
</tr>
<tr>
<td>Pothakos</td>
<td>82</td>
<td>89</td>
</tr>
<tr>
<td>Sullivan</td>
<td>81</td>
<td>71</td>
</tr>
<tr>
<td>Wahl</td>
<td>91</td>
<td>81</td>
</tr>
<tr>
<td>Zurfluh</td>
<td>83</td>
<td>94</td>
</tr>
</tbody>
</table>

Find the rank correlation coefficient and test the claim of no correlation between midterm score and final exam score. Use a significance level of 0.05.
CHAPTER 13 FORM C

Use the runs test to determine whether the given sequence is random. Use a significance level of 0.05.

15) Use a 0.05 level of significance to test the claim that the sequence of computer-generated numbers is random. Test for randomness above and below the mean.

8 7 5 7 3 9 1 8 0 4 3 8 4 6 2 3 9 7 5

16) The sequence of numbers below represents the maximum temperature (in degrees Fahrenheit) in July in one U.S. town for 30 consecutive years. Test the sequence for randomness above and below the median.

94 96 97 99 95 90 97 98 100 100
92 95 98 99 102 97 97 101 99 100
98 95 93 99 101 99 101 100 99 103

17) Test the sequence of digits below for randomness above and below the value of 4.5.

0 4 7 3 6 0 9 7 4 8
7 2 8 5 7 3 9 6 4 6
4 7 9 1 6 1 9 5 8 3
7 8 5 7 3 5 2 9 3 8

Solve the problem.

18) When performing a rank correlation test, one alternative to using Table A−9 to find critical values is to compute them using this approximation:

\[ r_s = \pm \sqrt{\frac{t^2}{t^2 + n - 2}} \]

where \( t \) is the \( t \)-score from Table A−3 corresponding to \( n - 2 \) degrees of freedom. Use this approximation to find critical values of \( r_s \) for the case where \( n = 17 \) and \( \alpha = 0.05 \).

A) \( \pm 0.411 \)    B) \( \pm 0.311 \)    C) \( \pm 0.480 \)    D) \( \pm 0.482 \)
CHAPTER 13 FORM C

Provide an appropriate response.

19) Which of the following tests could detect some nonlinear relationships between two variables?
   A) Wilcoxon signed-ranks test       B) Wilcoxon rank-sum test
   C) sign test                        D) rank correlation test

20) Which of the following tests can lead to the same conclusion as that of the Mann–Whitney U test?
    A) Wilcoxon rank-sum test          B) Kruskal–Wallis test
    C) Spearman correlation test       D) Wilcoxon signed-ranks test
1) A rank is a number assigned to an individual sample item according to its order in the ranked list. (Ranked lists are arranged in order by some criterion such as smallest to largest.) For repeating data points, you find the mean of the ranks involved. For the data point 5 in the list above, the rank for each would be 3 (which is the average of ranks 2, 3, and 4). For the data point 12 in the list above, the rank for each would be 7.5 (which is the average of 7 and 8).

2) The sign test compares the signs (negative or positive) of the differences for data sets, ignoring any ties resulting in a difference of zero. The sign test can be used to test claims involving two dependent samples, claims involving nominal data, and claims about the median of a single population. The underlying concept is that if two sets of data have equal medians, the number of positive signs should be approximately equal to the number of negative signs.

3) The runs test for randomness is a procedure for testing the randomness of data (with only two characteristics) using the concepts of runs. A run is a sequence of data that exhibit the same characteristic. For example, the data set M M M M F F F F F F F has four runs. The null hypothesis is that the sequence is random and the alternate hypothesis is that the sequence is not random. The runs test is based only on the order in which the data occur; it does not test the frequency of the data. The underlying concept is that if the number of runs is very low or very high, randomness is lacking.

4) $H_0$: median is 64.0 inches. $H_1$: median is not 64.0 inches.
   Test statistic: $x = 5$. Critical value: $x = 2$.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the population median is 64.0 inches.

5) $H_0$: the diet is not effective. $H_1$: the diet is effective.
   Convert $x = 36$ to the test statistic $z = -1.10$. Critical value: $z = -2.33$.
   Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the diet is effective.

6) Test statistic $T = 12.5$. Critical value: $T = 2$.
   Fail to reject the null hypothesis that both samples of systolic blood pressure readings come from the same population distribution.

7) $\mu_R = 144$, $\sigma_R = 16.2481$, $R = 145$, $z = 0.06$.
   Test statistic: $z = 0.06$. Critical values $z = \pm 1.96$.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the two populations of SAT scores are identical.

8) $\mu_R = 126.5$, $\sigma_R = 15.2288$, $R = 108.5$, $z = -1.18$.
   Test statistic: $z = -1.18$. Critical values $z = \pm 1.96$.
   Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the two populations of grade averages are identical.

9) $H_0$: The sample comes from a population with a median of 130 pounds.
   $H_1$: The sample comes from a population with a median different from 130 pounds.
   Test statistic: $T = 32.5$
   Critical value: 14
   Do not reject the null hypothesis. There is not sufficient evidence to reject the hypothesis that the sample comes from a population with a median of 130 pounds.

    Reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the samples of lifetimes of light bulbs come from identical populations.

    Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the samples of weights come from identical populations.
Answer Key
Testname: CHAPTER 13 FORM C

12) \( r_s = -0.506 \). Critical values: \( r_{0.05} = \pm 0.654 \).
   No significant correlation. There does not appear to be a correlation between the two variables.
13) \( r_s = 0.348 \). Critical values: \( r_{0.05} = \pm 0.470 \).
   No significant correlation. There does not appear to be a correlation between the two variables.
14) \( r_s = 0.706 \). Critical values: \( r_{0.05} = \pm 0.587 \).
   Significant correlation. There appears to be a correlation between midterm score and final exam score.
15) \( n_1 = 9, n_2 = 10, G = 14 \).
   Test statistic: \( G = 14 \). Critical values: 5, 16.
   Fail to reject the null hypothesis of randomness.
16) \( n_1 = 15, n_2 = 15, G = 10, 5\% \) cutoff values: 10, 22.
   Reject the null hypothesis of randomness.
17) \( n_1 = 15, n_2 = 25, G = 28, \mu_G = 19.75, \sigma_G = 2.9212 \).
   Test statistic: \( z = 2.82 \). Critical values: \( z = \pm 1.96 \).
   Reject the null hypothesis of randomness.
18) D
19) D
20) A
CHAPTER 14 FORM A

Name: ___________________________ Course Number: ________ Section Number: ______

**Directions:** Write your answers to the short-answer items in the spaces provided. Construct process control charts on the given axes. Label the axes specifically to the problem.

**Provide an appropriate response.**

1) Describe what process data are. Why are process data important to businesses? What is a common goal of businesses using quality control?

2) Describe the three criteria used to determine if a control chart indicates a process which is not statistically stable.

3) Draw a control chart that illustrates a process which is statistically stable and one which illustrates a process which is not statistically stable. Discuss the results.
CHAPTER 14 FORM A

Construct a run chart for individual values corresponding to the given data.

4) A machine that is supposed to produce ball bearings with a diameter of 7 mm yields the following data from a test of 5 ball bearings every 20 minutes.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.3</td>
<td>6.8</td>
<td>6.9</td>
</tr>
<tr>
<td>2</td>
<td>6.3</td>
<td>6.6</td>
<td>6.3</td>
</tr>
<tr>
<td>3</td>
<td>6.8</td>
<td>7.0</td>
<td>6.5</td>
</tr>
<tr>
<td>4</td>
<td>7.0</td>
<td>6.7</td>
<td>6.8</td>
</tr>
<tr>
<td>5</td>
<td>6.8</td>
<td>6.6</td>
<td>6.4</td>
</tr>
<tr>
<td>6</td>
<td>6.8</td>
<td>6.7</td>
<td>6.3</td>
</tr>
<tr>
<td>7</td>
<td>7.3</td>
<td>7.4</td>
<td>7.0</td>
</tr>
<tr>
<td>8</td>
<td>9.2</td>
<td>7.0</td>
<td>6.9</td>
</tr>
<tr>
<td>9</td>
<td>7.3</td>
<td>7.2</td>
<td>7.0</td>
</tr>
<tr>
<td>10</td>
<td>7.2</td>
<td>7.6</td>
<td>7.1</td>
</tr>
<tr>
<td>11</td>
<td>9.2</td>
<td>7.4</td>
<td>7.0</td>
</tr>
<tr>
<td>12</td>
<td>7.5</td>
<td>7.4</td>
<td>7.1</td>
</tr>
</tbody>
</table>
CHAPTER 14 FORM A

Construct an R chart and determine whether the process variation is within statistical control.

5) A machine that is supposed to fill small bottles to contain 20 ml yields the following data from a test of 4 bottles every hour.

<table>
<thead>
<tr>
<th>S</th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.9</td>
<td>20.1</td>
<td>20.2</td>
</tr>
<tr>
<td>2</td>
<td>20.4</td>
<td>20.0</td>
<td>20.3</td>
</tr>
<tr>
<td>3</td>
<td>20.0</td>
<td>20.7</td>
<td>20.4</td>
</tr>
<tr>
<td>4</td>
<td>20.4</td>
<td>20.2</td>
<td>20.1</td>
</tr>
<tr>
<td>5</td>
<td>19.9</td>
<td>19.8</td>
<td>19.6</td>
</tr>
<tr>
<td>6</td>
<td>19.4</td>
<td>19.4</td>
<td>19.6</td>
</tr>
<tr>
<td>7</td>
<td>19.8</td>
<td>19.4</td>
<td>19.6</td>
</tr>
<tr>
<td>8</td>
<td>19.9</td>
<td>19.8</td>
<td>20.0</td>
</tr>
<tr>
<td>9</td>
<td>20.2</td>
<td>20.3</td>
<td>20.1</td>
</tr>
<tr>
<td>10</td>
<td>20.0</td>
<td>20.3</td>
<td>20.0</td>
</tr>
<tr>
<td>11</td>
<td>20.3</td>
<td>20.5</td>
<td>20.1</td>
</tr>
<tr>
<td>12</td>
<td>20.1</td>
<td>19.9</td>
<td>19.8</td>
</tr>
<tr>
<td>13</td>
<td>19.5</td>
<td>19.8</td>
<td>19.7</td>
</tr>
<tr>
<td>14</td>
<td>19.4</td>
<td>19.8</td>
<td>19.8</td>
</tr>
<tr>
<td>15</td>
<td>19.5</td>
<td>19.6</td>
<td>19.6</td>
</tr>
</tbody>
</table>
CHAPTER 14 FORM A

Construct a control chart for $\bar{x}$.

6) A machine is supposed to fill cans that contain 12 oz. Each hour, a sample of four cans is tested; the results of 15 consecutive hours are given below.

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.2</td>
<td>11.7</td>
<td>11.8</td>
<td>11.7</td>
</tr>
<tr>
<td>2</td>
<td>11.8</td>
<td>11.3</td>
<td>11.6</td>
<td>11.6</td>
</tr>
<tr>
<td>3</td>
<td>11.3</td>
<td>12.0</td>
<td>11.8</td>
<td>11.7</td>
</tr>
<tr>
<td>4</td>
<td>12.1</td>
<td>11.6</td>
<td>12.1</td>
<td>12.0</td>
</tr>
<tr>
<td>5</td>
<td>11.8</td>
<td>11.8</td>
<td>11.9</td>
<td>11.9</td>
</tr>
<tr>
<td>6</td>
<td>12.0</td>
<td>11.9</td>
<td>11.8</td>
<td>11.7</td>
</tr>
<tr>
<td>7</td>
<td>11.6</td>
<td>12.0</td>
<td>11.9</td>
<td>11.8</td>
</tr>
<tr>
<td>8</td>
<td>11.5</td>
<td>12.1</td>
<td>11.9</td>
<td>12.0</td>
</tr>
<tr>
<td>9</td>
<td>12.1</td>
<td>12.1</td>
<td>11.7</td>
<td>11.9</td>
</tr>
<tr>
<td>10</td>
<td>11.7</td>
<td>12.0</td>
<td>11.6</td>
<td>11.9</td>
</tr>
<tr>
<td>11</td>
<td>12.1</td>
<td>12.4</td>
<td>11.9</td>
<td>12.2</td>
</tr>
<tr>
<td>12</td>
<td>12.5</td>
<td>12.0</td>
<td>12.4</td>
<td>12.3</td>
</tr>
<tr>
<td>13</td>
<td>12.5</td>
<td>12.0</td>
<td>12.1</td>
<td>12.1</td>
</tr>
<tr>
<td>14</td>
<td>12.4</td>
<td>12.0</td>
<td>12.0</td>
<td>12.4</td>
</tr>
<tr>
<td>15</td>
<td>12.4</td>
<td>12.4</td>
<td>12.6</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

7) A control chart for $\bar{x}$ is shown below. Determine whether the process mean is within statistical control. If it is not, identify which of the three out-of-control criteria lead to rejection of a statistically stable mean.

![X Chart for Can Amt]

- **UCL** = 13.00
- **$\bar{x}$** = 11.95
- **LCL** = 11.62

Sample Number

278
CHAPTER 14 FORM A

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

8) A control chart for R is shown below. Determine whether the process variation is within statistical control. If it is not, identify which of the three out-of-control criteria lead to rejection of statistically stable variation.

![R Chart for BB Size](image)

9) A run chart for individual values is shown below. Does there appear to be a pattern suggesting that the process is not within statistical control? If so, describe the pattern.

![Individual Value Chart](image)

10)
CHAPTER 14 FORM A

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

11)

Use the given process data to construct a control chart for p.

12) If the weight of cereal in a particular packet is less than 14 oz, the packet is considered nonconforming. Each week, the manufacturer randomly selects 1,000 cereal packets and determines the number that are nonconforming. The results for 12 consecutive weeks are shown below.

46 32 21 30 47 31 32 52 48 45 62 58
Answer Key
Testname: CHAPTER 14 FORM A

1) Process data are data arranged according to some time sequence. Process data are important to businesses because important characteristics of process data can change over time, and businesses are interested in ensuring quality by controlling these important characteristics. Businesses using quality control are interested in reducing variability in characteristics of their product.

2) 1) There is a pattern, trend, or cycle that is obviously not random. There is a point lying beyond the upper or lower control limits. 3) There are 8 consecutive points all above or all below the center line. Additional criteria mentioned include: 6 consecutive points all increasing or decreasing; 14 consecutive points all alternating between up and down; 2 out of 3 consecutive points beyond control limits 1 standard deviation away from the center line; 4 out of 5 consecutive points beyond control limits 2 standard deviations away from the center line.

3) Examples will vary. The control chart which is statistically stable should show only random variation. The control chart which is not statistically stable will display at least one of the following: 1) There is a pattern, trend, or cycle that is obviously not random. There is a point lying beyond the upper or lower control limits. 3) There are 8 consecutive points all above or all below the center line.

4)

I Chart for BB Size

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millimeter</td>
<td>6</td>
<td>6.5</td>
<td>7</td>
<td>7.5</td>
<td>8</td>
<td>6.5</td>
<td>7</td>
</tr>
</tbody>
</table>

5) The process appears to be within statistical control.

R Chart for C1

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Range</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

UCL = 0.8672

R = 0.38

LCL = 0.000
6) X Chart for Can Ant

7) Process mean is not within statistical control. There are points above and below the control limits. There is an upward trend.

8) Process variation appears to be in statistical control.

9) Process appears to be out of statistical control. There is a cyclical pattern.

10) Process appears to be out of statistical control. There are points that lie above the upper control limit. There are 8 consecutive points below the center line. There is increasing variation.

11) Process appears to be out of statistical control. There is an upward shift. There are 8 consecutive points below the center line.

12) Prop Portion

Week Number
CHAPTER 14 FORM B

Name:_________________________ Course Number: _________ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided.
Construct process control charts on the given axes. Label the axes specifically to the problem.

Provide an appropriate response.

1) A common goal of quality control is to reduce variation in a product or service. List and describe the two types of variability. Give an example of each.

2) Describe a control chart. Complete the table to identify the important parts of different types of control charts.

<table>
<thead>
<tr>
<th>Control chart</th>
<th>Center line and Upper control limit</th>
<th>Lower control limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control chart for R</td>
<td>Points plotted</td>
<td>how to compute</td>
</tr>
<tr>
<td>Control chart for X</td>
<td></td>
<td>control limit</td>
</tr>
<tr>
<td>Control chart for p</td>
<td></td>
<td>control limit</td>
</tr>
</tbody>
</table>

3) Relate the concept of control charts to the concept of confidence intervals.
CHAPTER 14 FORM B

Construct a run chart for individual values corresponding to the given data.

4) A machine is supposed to fill boxes to a weight of 50 lbs. Every 30 minutes a sample of four boxes is tested; the results are given below.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>51</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>60</td>
<td>32</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
<td>59</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>61</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>50</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>52</td>
<td>51</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>40</td>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td>9</td>
<td>48</td>
<td>67</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>43</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>48</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>58</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>47</td>
<td>52</td>
<td>47</td>
</tr>
<tr>
<td>15</td>
<td>52</td>
<td>57</td>
<td>58</td>
</tr>
</tbody>
</table>
CHAPTER 14 FORM B

Construct an R chart and determine whether the process variation is within statistical control.

5) A machine is supposed to fill cans that contain 12 oz. Each hour, a sample of four cans is tested; the results of 15 consecutive hours are given below.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.2</td>
<td>11.7</td>
<td>11.6</td>
</tr>
<tr>
<td>2</td>
<td>11.8</td>
<td>11.3</td>
<td>11.6</td>
</tr>
<tr>
<td>3</td>
<td>11.3</td>
<td>12.0</td>
<td>11.8</td>
</tr>
<tr>
<td>4</td>
<td>12.1</td>
<td>11.6</td>
<td>12.1</td>
</tr>
<tr>
<td>5</td>
<td>11.8</td>
<td>11.8</td>
<td>11.9</td>
</tr>
<tr>
<td>6</td>
<td>12.0</td>
<td>11.9</td>
<td>11.8</td>
</tr>
<tr>
<td>7</td>
<td>11.6</td>
<td>12.0</td>
<td>11.9</td>
</tr>
<tr>
<td>8</td>
<td>11.5</td>
<td>12.1</td>
<td>11.9</td>
</tr>
<tr>
<td>9</td>
<td>12.1</td>
<td>12.1</td>
<td>11.7</td>
</tr>
<tr>
<td>10</td>
<td>11.7</td>
<td>12.0</td>
<td>11.6</td>
</tr>
<tr>
<td>11</td>
<td>12.1</td>
<td>12.4</td>
<td>11.9</td>
</tr>
<tr>
<td>12</td>
<td>12.5</td>
<td>12.0</td>
<td>12.4</td>
</tr>
<tr>
<td>13</td>
<td>12.5</td>
<td>12.0</td>
<td>12.1</td>
</tr>
<tr>
<td>14</td>
<td>12.4</td>
<td>12.0</td>
<td>12.4</td>
</tr>
<tr>
<td>15</td>
<td>12.4</td>
<td>12.4</td>
<td>12.6</td>
</tr>
</tbody>
</table>
CHAPTER 14 FORM B

Construct a control chart for \( \bar{x} \).

6) A machine that is supposed to produce ball bearings with a diameter of 7 mm yields the following data from a test of 5 ball bearings every 20 minutes.

<table>
<thead>
<tr>
<th>S</th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.3</td>
<td>6.8</td>
<td>6.9</td>
</tr>
<tr>
<td>2</td>
<td>6.3</td>
<td>6.6</td>
<td>6.3</td>
</tr>
<tr>
<td>3</td>
<td>6.8</td>
<td>6.7</td>
<td>7.0</td>
</tr>
<tr>
<td>4</td>
<td>7.0</td>
<td>6.7</td>
<td>6.8</td>
</tr>
<tr>
<td>5</td>
<td>6.8</td>
<td>6.8</td>
<td>6.5</td>
</tr>
<tr>
<td>6</td>
<td>6.8</td>
<td>6.7</td>
<td>6.6</td>
</tr>
<tr>
<td>7</td>
<td>7.3</td>
<td>7.3</td>
<td>7.4</td>
</tr>
<tr>
<td>8</td>
<td>7.2</td>
<td>7.0</td>
<td>7.2</td>
</tr>
<tr>
<td>9</td>
<td>7.3</td>
<td>7.6</td>
<td>7.1</td>
</tr>
<tr>
<td>10</td>
<td>7.2</td>
<td>7.6</td>
<td>7.5</td>
</tr>
<tr>
<td>11</td>
<td>7.2</td>
<td>7.4</td>
<td>7.0</td>
</tr>
<tr>
<td>12</td>
<td>7.5</td>
<td>7.4</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

7) A control chart for \( \bar{x} \) is shown below. Determine whether the process mean is within statistical control. If it is not, identify which of the three out-of-control criteria lead to rejection of a statistically stable mean.

![X Chart for C1](chart.png)

UCL = 20.80

\[ \bar{x} = 19.93 \]

LCL = 19.67
CHAPTER 14 FORM B

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

8) A control chart for R is shown below. Determine whether the process variation is within statistical control. If it is not, identify which of the three out-of-control criteria lead to rejection of statistically stable variation.

9) A run chart for individual values is shown below. Does there appear to be a pattern suggesting that the process is not within statistical control? If so, describe the pattern.

10)
CHAPTER 14 FORM B

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

11)

![Control Chart Image]

Use the given process data to construct a control chart for p.

12) A drugstore considers a wait of more than 5 minutes to be a defect. Each week 100 customers are randomly selected and timed at the checkout line. The numbers of defects for 20 consecutive weeks are given below.

4 4 5 5 5 5 5 6 6 6 12 6 6 7 6 7 8 7
1) Random variation is due to chance, the variation inherent in any process that is not capable of producing each good or service exactly the same way every time. Assignable variation results from causes that can be identified. Examples may vary.

2) A control chart of a process characteristic (such as a mean or range) consists of values plotted sequentially over time, and it includes a center line, representing a central value of the characteristic measurement, as well as lower and upper control limits, representing boundaries used to separate and identify any points considered to be unusual.

<table>
<thead>
<tr>
<th>Points plotted</th>
<th>Center line and how to compute</th>
<th>Upper control limit</th>
<th>Lower control limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control chart for R</td>
<td>$\bar{R}$, the average of the sample ranges</td>
<td>$D_4 \bar{R}$</td>
<td>$D_3 \bar{R}$</td>
</tr>
<tr>
<td>Control chart for $\bar{X}$</td>
<td>$\bar{X}$, the average of the sample means</td>
<td>$\bar{X} + A_2 \bar{R}$</td>
<td>$\bar{X} - A_2 \bar{R}$</td>
</tr>
<tr>
<td>Control chart for p</td>
<td>$\bar{p}$, the pooled estimate of the proportion for all items sampled</td>
<td>$p + 3\sqrt{\frac{p \cdot (1 - p)}{n}}$</td>
<td>If the lower control limit is negative, use 0.</td>
</tr>
</tbody>
</table>

3) Control charts have upper control limits and lower control limits found by processes similar to those for finding confidence intervals. Control charts allow us to examine processes to see if they remain within control, that is, within the confidence intervals.

4)
5) The process appears to be within statistical control.

6) 

7) Process mean is not within statistical control. There are points above and below the control limits. There is a cyclical pattern.

8) Process variation appears to be out of statistical control. There is an upward trend indicating that variation is increasing. There are points above the upper control limit. There are more than 8 consecutive points below the center line.

9) Process appears to be out of statistical control. The variation is increasing over time.

10) Process is out of statistical control. There is a downward trend.

11) Process appears to be out of statistical control. There is a point that lies above the upper control limit.

12) 

290
CHAPTER 14 FORM C

Name: __________________________ Course Number: _______ Section Number: _____

Directions: Write your answers to the short-answer items in the spaces provided. Construct process control charts on the given axes. Label the axes specifically to the problem.

Provide an appropriate response.

1) Describe a run chart and give an example. Refer to the values on each of the axes as you describe the run chart.

2) Define statistically stable (or "within statistical control"). Show examples of run charts which illustrate processes which are not statistically controlled. Discuss the pattern which indicates the process is not statistically controlled for each example.

3) Describe how the term Six Sigma relates to statistical process control.
CHAPTER 14 FORM C

Construct a run chart for individual values corresponding to the given data.

4) A machine that is supposed to fill small bottles to contain 20 ml yields the following data from a test of 4 bottles every hour.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.9</td>
<td>20.1</td>
<td>20.2</td>
</tr>
<tr>
<td>2</td>
<td>20.4</td>
<td>20.0</td>
<td>20.3</td>
</tr>
<tr>
<td>3</td>
<td>20.0</td>
<td>20.7</td>
<td>20.4</td>
</tr>
<tr>
<td>4</td>
<td>20.4</td>
<td>20.1</td>
<td>19.9</td>
</tr>
<tr>
<td>5</td>
<td>19.9</td>
<td>19.8</td>
<td>19.6</td>
</tr>
<tr>
<td>6</td>
<td>19.4</td>
<td>19.4</td>
<td>19.6</td>
</tr>
<tr>
<td>7</td>
<td>19.8</td>
<td>19.4</td>
<td>19.6</td>
</tr>
<tr>
<td>8</td>
<td>19.9</td>
<td>19.8</td>
<td>20.0</td>
</tr>
<tr>
<td>9</td>
<td>20.2</td>
<td>20.3</td>
<td>20.1</td>
</tr>
<tr>
<td>10</td>
<td>20.0</td>
<td>20.3</td>
<td>20.2</td>
</tr>
<tr>
<td>11</td>
<td>20.3</td>
<td>20.5</td>
<td>20.1</td>
</tr>
<tr>
<td>12</td>
<td>20.1</td>
<td>19.9</td>
<td>19.8</td>
</tr>
<tr>
<td>13</td>
<td>19.5</td>
<td>19.8</td>
<td>19.7</td>
</tr>
<tr>
<td>14</td>
<td>19.4</td>
<td>19.8</td>
<td>19.4</td>
</tr>
<tr>
<td>15</td>
<td>19.5</td>
<td>19.6</td>
<td>19.9</td>
</tr>
</tbody>
</table>
Construct an R chart and determine whether the process variation is within statistical control.

5) A machine is supposed to fill boxes to a weight of 50 lbs. Every 30 minutes a sample of four boxes is tested; the results are given below.

<table>
<thead>
<tr>
<th>S</th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>51</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>60</td>
<td>32</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
<td>59</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>61</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>50</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>52</td>
<td>51</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>40</td>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td>9</td>
<td>48</td>
<td>67</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>43</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>48</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>58</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>47</td>
<td>52</td>
<td>47</td>
</tr>
<tr>
<td>15</td>
<td>52</td>
<td>57</td>
<td>58</td>
</tr>
</tbody>
</table>
CHAPTER 14 FORM C

Construct a control chart for $\bar{x}$.

6) A machine is supposed to fill cans that contain 12 oz. Each hour, a sample of four cans is tested; the results of 15 consecutive hours are given below.

<table>
<thead>
<tr>
<th>S</th>
<th>A</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.2</td>
<td>11.6</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>11.8</td>
<td>11.2</td>
<td>11.67</td>
</tr>
<tr>
<td>3</td>
<td>11.3</td>
<td>11.2</td>
<td>11.97</td>
</tr>
<tr>
<td>4</td>
<td>12.1</td>
<td>11.6</td>
<td>12.1</td>
</tr>
<tr>
<td>5</td>
<td>11.8</td>
<td>11.8</td>
<td>11.9</td>
</tr>
<tr>
<td>6</td>
<td>12.0</td>
<td>11.9</td>
<td>11.7</td>
</tr>
<tr>
<td>7</td>
<td>11.6</td>
<td>12.0</td>
<td>11.9</td>
</tr>
<tr>
<td>8</td>
<td>11.5</td>
<td>12.1</td>
<td>11.9</td>
</tr>
<tr>
<td>9</td>
<td>12.1</td>
<td>12.1</td>
<td>11.7</td>
</tr>
<tr>
<td>10</td>
<td>11.7</td>
<td>12.0</td>
<td>11.6</td>
</tr>
<tr>
<td>11</td>
<td>12.1</td>
<td>12.4</td>
<td>11.9</td>
</tr>
<tr>
<td>12</td>
<td>12.5</td>
<td>12.0</td>
<td>12.4</td>
</tr>
<tr>
<td>13</td>
<td>12.5</td>
<td>12.0</td>
<td>12.1</td>
</tr>
<tr>
<td>14</td>
<td>12.4</td>
<td>12.0</td>
<td>12.0</td>
</tr>
<tr>
<td>15</td>
<td>12.4</td>
<td>12.4</td>
<td>12.6</td>
</tr>
</tbody>
</table>

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

7) A control chart for $\bar{x}$ is shown below. Determine whether the process mean is within statistical control. If it is not, identify which of the three out-of-control criteria lead to rejection of a statistically stable mean.
CHAPTER 14 FORM C

Examine the given run chart or control chart and determine whether the process is within statistical control. If it is not, identify which of the three out-of-control criteria apply.

8) A control chart for $\bar{x}$ is shown below. Determine whether the process mean is within statistical control. If it is not, identify which of the three out-of-control criteria lead to rejection of a statistically stable mean.

**X Chart for BB Size**

![X Chart for BB Size](image)

- **UCL** = 7.990
- **$\bar{x}$** = 6.992
- **LCL** = 6.717

9) A run chart for individual values is shown below. Does there appear to be a pattern suggesting that the process is not within statistical control? If so, describe the pattern. (Assume the center line is at 0.03, UCL at 0.04, and LCL at 0.02.)

![Run Chart for Individual Values](image)

10)
CHAPTER 14 FORM C

Use the given process data to construct a control chart for p.

11) A manufacturer monitors the level of defects in the television sets that it produces. Each week, 200 television sets are randomly selected and tested and the number of defects is recorded. The results for 12 consecutive weeks are shown below.

4 7 5 6 8 3 12 4 4 5 6 2

Solve the problem.

12) A control chart for attributes is to be constructed. Which process would have wider control limits, a process which has been having a 5% rate of nonconforming items, or a process which has been having a 10% of nonconforming items? Assume that both processes have the same sample sizes. For a given sample size, would it be easier to detect a shift from 5% to 10% or a shift from 10% to 15%? Explain your reasoning.
Answer Key
Testname: CHAPTER 14 FORM C

1) A run chart is a sequential plot of individual data values over time. The horizontal axis typically is used for the time sequence, and the vertical axis is used for the values of the data. Examples will vary.

2) A process is statistically stable if it has only natural variation, with no patterns, cycles, or any unusual points. A process is statistically unstable, given the following conditions: an obvious upward or downward trend, graphs with an upward or downward shift (relatively stable values for the first few, a shift, relatively stable values at the end), graphs with one exceptionally high or low value, graphs with cyclical behavior, or graphs whose variation is increasing over time.

3) Answers will vary. Possible answer: Six Sigma represents reduced variation in a manufacturing process, which is the goal of statistical process control.

4) ![I Chart for C1](chart1.png)

5) The process appears to be within statistical control.

![R Chart for C1](chart2.png)

6) ![X Chart for Can Qty](chart3.png)
7) Process mean is not within statistical control. One of the points lies above the upper control limit.
8) Process mean is not within statistical control. There are points above and below the control limits. There is a shift upward.
9) Process appears to be in statistical control.
10) Process appears to be within statistical control.

11) 

```
Week Number
```

```
Propotion
```

```
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
```

```
UCL = 0.0622
```

```
LCL = 0.000
```

```
P = 0.0275
```

12) The process which has been having a 10% of nonconforming items would have wider control limits. It would be easier to detect a shift from 5% to 10% than a shift from 10% to 15%, because at a 5% rate of nonconforming items, the control limits are narrower, and it thus takes a smaller shift before the proportion falls outside the control limits.