MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Provide an appropriate response.

1) A company held a cholesterol screening clinic for its employees. The results are summarized in the table below by age group and cholesterol level. A graphical display of the conditional distributions of cholesterol levels by age is given.

<table>
<thead>
<tr>
<th>Cholesterol Level</th>
<th>Under 30</th>
<th>30-49</th>
<th>Over 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>31</td>
<td>38</td>
<td>24</td>
</tr>
<tr>
<td>Borderline</td>
<td>50</td>
<td>96</td>
<td>111</td>
</tr>
<tr>
<td>High</td>
<td>18</td>
<td>46</td>
<td>63</td>
</tr>
</tbody>
</table>

Does this prove that aging causes increased cholesterol levels? Explain.

A) Yes. There is evidence of an association between age and cholesterol level. The percentage of employees with high cholesterol increased with age. Likewise, the percentage of employees with low cholesterol decreased with age.

B) No. Generally, an association between two variables does not imply a cause-and-effect relationship. Specifically, these data come from only one company and cannot be applied to all people. Only a controlled experiment can isolate the relationship between age and cholesterol level.

C) No. There is sufficient evidence of an association between age and cholesterol level. However, the percentage of employees with low cholesterol increased with age. Likewise, the percentage of employees with high cholesterol decreased with age.

D) Yes. Generally, an association between two variables does imply a cause-and-effect relationship. However, these data come from only one company and cannot necessarily be applied to all people. A controlled experiment that can isolate the relationship between age and cholesterol level would make this a stronger claim.
E) No. Generally, an association between two variables does imply a cause-and-effect relationship. Specifically, these data come from only one company and cannot be applied to all people. However, if more companies provided cholesterol screenings and these data were compiled, then the claim could be proven.

Describe the distribution (shape, center, spread, unusual features).

2) A dotplot of the number of tornadoes each year in a certain county from 1948 to 2004 is given. Each dot represents a year in which there were that many tornadoes.

A) The distribution of the number of tornadoes per year is unimodal and symmetric, with a center around 5 tornadoes per year. The number of tornadoes per year ranges from 0 to 7.

B) The distribution of the number of tornadoes per year is unimodal and skewed to the right, with a center around 5 tornadoes per year. The number of tornadoes per year ranges from 0 to 7.

C) The distribution of the number of tornadoes per year is unimodal and skewed to the left, with a center around 5 tornadoes per year. The number of tornadoes per year ranges from 0 to 7.

D) The distribution of the number of tornadoes per year is unimodal and symmetric, with a center around 3.5 tornadoes per year. The number of tornadoes per year ranges from 0 to 7.

E) The distribution of the number of tornadoes per year is unimodal and skewed to the left, with a center around 3.5 tornadoes per year. The number of tornadoes per year ranges from 0 to 7.

3) A student at a local university took a total of 20 exams during freshman year. The student recorded the exam scores as percentages and created the following stem-and-leaf display. The lower stem contains leaves with the digits 0-4 and the upper stem contains leaves with digits 5-9. In addition to describing the distribution, give a reason to account for the shape of this distribution.

Exam Grades
9 6 6 6 6 6 7 8 9
9 0 1 2 2
8 5 6 7 8
8 7 9
7 7
6 6 2 4
5
Compare the distributions (shape, center, spread, unusual features).

4) The back-to-back stem-and-leaf display compares the percent growth in sales for a retail chain's stores located in two regions of the United States. The lower stem contains leaves with the digits 0-4 and the upper stem contains leaves with digits 5-9.

<table>
<thead>
<tr>
<th>NE/MW Stores</th>
<th>S/W Stores</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>211000</td>
<td></td>
</tr>
<tr>
<td>998876555555</td>
<td>4</td>
</tr>
<tr>
<td>444433</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>556</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
</tr>
<tr>
<td>2</td>
<td>67</td>
</tr>
<tr>
<td>2</td>
<td>233</td>
</tr>
<tr>
<td>1</td>
<td>55567</td>
</tr>
<tr>
<td>1</td>
<td>024444</td>
</tr>
<tr>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td>0</td>
<td>444</td>
</tr>
</tbody>
</table>

Key:
3 | 5 =35% sales growth
A) The distribution of sales growth in NE/MW stores is unimodal, symmetric and tightly clustered around 45% growth. The distribution of sales growth in S/W stores is much more spread out, with most stores having sales growth between 5% and 35%. A typical S/W store had about 15% growth. There were two outliers, one store with 59% growth and another with 67% growth. Generally, the sales growth rates in the NE/MW stores were higher and more variable than the rates in the S/W stores.

B) The distribution of sales growth in NE/MW stores is unimodal, symmetric and tightly clustered around 45% growth. The distribution of sales growth in S/W stores is much more spread out, with most stores having sales growth between 5% and 35%. A typical S/W store had about 15% growth. There were two outliers, one store with 59% growth and another with 67% growth. Generally, the sales growth rates in the S/W stores were higher and more variable than the rates in the NE/MW stores.

C) The distribution of sales growth in the NE/MW stores is unimodal, symmetric and tightly clustered around 45% growth. The distribution of sales growth in S/W stores is much more spread out, with most stores having sales growth between 5% and 35%. A typical S/W store had about 15% growth. There were two outliers, one store with 59% growth and another with 67% growth. Generally, the sales growth rates in the NE/MW stores were higher and less variable than the rates in the S/W stores.

D) The distribution of sales growth in NE/MW stores is unimodal, symmetric and tightly clustered around 55% growth. The distribution of sales growth in S/W stores is much more spread out, with most stores having sales growth between 5% and 35%. A typical S/W store had about 25% growth. There were two outliers, one store with 59% growth and another with 67% growth. Generally, the sales growth rates in the S/W stores were higher and more variable than the rates in the NE/MW stores.

E) The distribution of sales growth in NE/MW stores is unimodal, symmetric and tightly clustered around 55% growth. The distribution of sales growth in S/W stores is much more spread out, with most stores having sales growth between 5% and 35%. A typical S/W store had about 25% growth. There were two outliers, one store with 59% growth and another with 67% growth. Generally, the sales growth rates in the NE/MW stores were higher and less variable than the rates in the S/W stores.

5) The back- to- back dotplot shows the number of fatalities per year caused by tornadoes in a certain state for two periods: 1950-1974 and 1975-1999. In addition to comparing these distributions, state a reason explaining any differences.

<table>
<thead>
<tr>
<th>Number of Fatalities per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fatalities 1950-1974</strong></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>
<tr>
<td>5</td>
</tr>
<tr>
<td><strong>Fatalities 1975-1999</strong></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>
<tr>
<td>5</td>
</tr>
</tbody>
</table>
A) The distribution of the number of fatalities per year for the period 1950-1974 is unimodal and approximately symmetric. The center of the distribution is about 2 fatalities per year. The number of fatalities per year ranges from 0 to 5 deaths. For the period 1975-1999, the distribution of the number of fatalities per year is also unimodal, but skewed to the right. A typical number of fatalities for this distribution is 0 fatalities, with a range of 0 to 5 deaths. Before 1975, there were more fatalities as a result of tornadoes. Higher construction standards, better warning systems, or medical advancements could all account for this difference.

B) The distribution of the number of fatalities per year for the period 1950-1974 is unimodal and skewed to the right. The center of the distribution is about 3 fatalities per year. The number of fatalities per year ranges from 0 to 5 deaths. For the period 1975-1999, the distribution of the number of fatalities per year is also unimodal and skewed to the right. A typical number of fatalities for this distribution is 0 fatalities, with a range of 0 to 5 deaths.

C) The distribution of the number of fatalities per year for the period 1950-1974 is unimodal and approximately symmetric. The center of the distribution is about 2 fatalities per year. The number of fatalities per year ranges from 0 to 5 deaths. For the period 1975-1999, the distribution of the number of fatalities per year is also unimodal, but skewed to the left. A typical number of fatalities for this distribution is 0 fatalities, with a range of 0 to 5 deaths.

D) The distribution of the number of fatalities per year for the period 1950-1974 is unimodal and approximately symmetric. The center of the distribution is about 2 fatalities per year. The number of fatalities per year ranges from 0 to 5 deaths. For the period 1975-1999, the distribution of the number of fatalities per year is also unimodal, but skewed to the left. A typical number of fatalities for this distribution is 0 fatalities, with a range of 0 to 5 deaths. Before 1975, there were more fatalities as a result of tornadoes. Higher construction standards, better warning systems, or medical advancements could all account for this difference.

E) The distribution of the number of fatalities per year for the period 1950-1974 is unimodal and skewed to the right. The center of the distribution is about 3 fatalities per year. The number of fatalities per year ranges from 0 to 5 deaths. For the period 1975-1999, the distribution of the number of fatalities per year is also unimodal and skewed to the right. A typical number of fatalities for this distribution is 0 fatalities, with a range of 0 to 5 deaths. Before 1975, there were more fatalities as a result of tornadoes. Higher construction standards, better warning systems, or medical advancements could all account for this difference.
Solve the problem.

6) The weekly salaries (in dollars) of 24 randomly selected employees of a company are shown below, and displayed on the following boxplot.

<table>
<thead>
<tr>
<th>310</th>
<th>320</th>
<th>450</th>
<th>460</th>
<th>470</th>
<th>500</th>
<th>520</th>
<th>540</th>
</tr>
</thead>
<tbody>
<tr>
<td>580</td>
<td>600</td>
<td>650</td>
<td>700</td>
<td>710</td>
<td>840</td>
<td>870</td>
<td>900</td>
</tr>
<tr>
<td>1000</td>
<td>1200</td>
<td>1250</td>
<td>1300</td>
<td>1400</td>
<td>1720</td>
<td>2500</td>
<td>3700</td>
</tr>
</tbody>
</table>

Find the mean and the median. Which is the most appropriate measure of center?

A) Mean: $979; Median: $705
   The mean is the most appropriate measure of center, because the outliers are not significant in the distribution.

B) Mean: $979; Median: $705
   The median is the most appropriate measure of center, because of the outliers.

C) Mean: $979; Median: $705
   The mean is the most appropriate measure of center, because of the outliers.

D) Mean: $979; Median: $705
   The median is the most appropriate measure of center, because the outliers are not significant in the distribution.

E) Mean: $942; Median: $700
   The median is the most appropriate measure of center, because of the outliers.
7) Here are 3 boxplots of weekly gas prices at a service station in the United States (price in $ per gallon). Compare the distribution of prices over the three years.

A) Gas price have been increasing on average over the 3-year period, and the spread has been increasing as well. The distribution has been skewed to the left, and there were 3 high outliers in 2005.

B) Gas price have been increasing on average over the 3-year period, and the spread has been increasing as well. The distribution has been skewed to the right, and there were 3 high outliers in 2005.

C) Gas price have been increasing on average over the 3-year period, but the spread has been decreasing. The distribution has been skewed to the left, and there were 3 high outliers in 2005.

D) Gas price have been decreasing on average over the 3-year period, and the spread has been decreasing. The distribution has been skewed to the left, and there were 3 high outliers in 2005.

E) Gas price have been decreasing on average over the 3-year period, but the spread has been increasing. The distribution has been skewed to the right, and there were 3 high outliers in 2005.
8) Here are boxplots of the points scored during the first 10 games of the basketball season for both Caroline and Alexandra. The coach can take only one player to the state championship. Which one should she take knowing that she would like a safe player?

![Boxplots of points scored](image)

A) Both of them, because both girls have a median score of about 18 points per game.
B) Alexandra, because she is the more consistent player.
C) Caroline, because the IQR is the largest.
D) Alexandra, because the IQR is the largest.
E) Caroline, because she is the more consistent player.

9) The boxplots display case prices (in dollars) of white wines produced by three vineyards in the western United States. In which vineyard are the wines generally more expensive?

![Boxplots of case prices](image)

A) Vineyards A and B, because they have a similar average price, and roughly the same spread.
B) Vineyard C, because it has the smallest range.
C) Vineyard A, because it has one outlier at about $145.
D) Vineyard C, because it has the highest average price and the smallest spread.
E) Vineyard B, because it has the highest case price at about $150.
10) Shown below are the boxplot, the histogram and summary statistics for the weekly salaries (in dollars) of 24 randomly selected employees of a company:

<table>
<thead>
<tr>
<th>Count</th>
<th>Mean</th>
<th>Median</th>
<th>StdDev</th>
<th>Min</th>
<th>Max</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>978.8</td>
<td>705</td>
<td>765.7</td>
<td>310</td>
<td>3700</td>
<td>510</td>
<td>1225</td>
</tr>
</tbody>
</table>

Write a few sentences describing the distribution.

A) The distribution is bimodal and skewed to the right. As shown in the boxplot, there are two outliers, weekly salaries of $2500 and about $3700. The median was 705, while the mean was 978.8, above the median score. The middle 50% of the weekly salaries were between $705 and $1225 for an IQR of $520.

B) The distribution is unimodal and skewed to the left. As shown in the boxplot, there are two outliers, weekly salaries of $2500 and about $3700. The median was 705, while the mean was 978.8, above the median score. The middle 50% of the weekly salaries were between $705 and $1225 for an IQR of $520.

C) The distribution is unimodal and skewed to the right. As shown in the boxplot, there are two outliers, weekly salaries of $2500 and about $3700. The median was 705, above the median score. The middle 50% of the weekly salaries were between $510 and $1225 for an IQR of $715.

D) The distribution is unimodal and skewed to the right. As shown in the boxplot, there are two outliers, weekly salaries of $2500 and about $3700. The median was 705, while the mean was 978.8, above the median score. The middle 50% of the weekly salaries were between $510 and $1225 for an IQR of $715.

E) The distribution is unimodal and skewed to the left. As shown in the boxplot, there are two outliers, weekly salaries of $2500 and about $3700. The median was 705, while the mean was 978.8, above the median score. The middle 50% of the weekly salaries were between $510 and $1225 for an IQR of $715.
Provide an appropriate response.

11) The ages of people attending the opening show of a new movie are summarized in the ogive shown. Estimate the IQR of the ages.

\[
\begin{array}{c|c|c|c|c|c}
\text{Age} & 20 & 30 & 40 & 50 & 60 \\
\hline
\text{Cumulative %} & 0 & 20 & 40 & 60 & 100 \\
\end{array}
\]

A) 13  
B) 5  
C) 30  
D) 21  
E) 37

Suppose you were to collect data for the pair of given variables in order to make a scatterplot. Determine for each variable if it is the explanatory variable, the response variable, or whether it could be both.

12) Hot chocolate sales, heater sales

A) Hot chocolate sales: both
   Heater sales: explanatory

B) Hot chocolate sales: both
   Heater sales: both

C) Hot chocolate sales: response
   Heater sales: explanatory

D) Hot chocolate sales: explanatory
   Heater sales: both

E) Hot chocolate sales: explanatory
   Heater sales: response
Determine whether the scatterplot shows little or no association, a negative association, a positive association, a linear association, a moderately strong association, or a very strong association (multiple associations are possible).

13) 
A) Little or no association  
B) Moderately strong association  
C) Negative association  
D) Linear association, moderately strong association  
E) Negative association, linear association

14) 
A) Positive association, linear association  
B) Moderately strong association  
C) Little or no association  
D) Negative association, linear association  
E) Positive association, moderately strong association
Several scatterplots are given with calculated correlations. Which is which?

15) 1) 2) 3) 4)

a) - 0.956, b) - 0.061, c) 0.716, d) 0.961

A) 1a, 2c, 3b, 4d  
B) 1d, 2b, 3a, 4c  
C) 1a, 2c, 3d, 4b  
D) 1c, 2d, 3b, 4a  
E) 1b, 2d, 3c, 4d
Solve the problem.

16) Soda is often considered unhealthy because its content is high in both caffeine and refined sugar. But are the two related? Caffeine and refined sugar contents (in milligrams) of 12 ounces of several brands of soda are shown on the following scatterplot. The correlation between caffeine and refined sugar is 0.145. Describe the association.

A) Weak curved relation in a positive direction
B) Weak linear relation in a positive direction
C) Weak nonlinear relation in a positive direction
D) No evidence of relation
E) Strong linear relation in a positive direction

Tell what the residual plot indicates about the appropriateness of the linear model that was fit to the data.

A) Model is appropriate.
B) Model is not appropriate. The relationship is nonlinear.
C) Model may not be appropriate. The spread is changing.
A random sample of records of electricity usage of homes gives the amount of electricity used in July and size (in square feet) of 135 homes. A regression was done to predict the amount of electricity used (in kilowatt-hours) from size. The residuals plot indicated that a linear model is appropriate. The model is usage $= 1298 + 0.4$ size. Explain what the slope of the line says about the electricity usage and home size.

A) On average, the size of the house increases by 1298 feet for every kilowatt-hour used.
B) On average, the amount of electricity used increases by 1298 kilowatt-hours when the size of the house is increased by a square foot.
C) On average, the size of the house increases by 0.4 feet for every kilowatt-hour used.
D) On average, the amount of electricity used increases by 0.4 kilowatt-hours when the size of the house is increased by a square foot.
E) On average, the amount of electricity used is 0.4 kilowatt hours less than the size of the house.

A random sample of records of electricity usage of homes gives the amount of electricity used and size (in square feet) of 135 homes. A regression to predict the amount of electricity used (in kilowatt-hours) from size has an $R^2$ of 71.8%. The residuals plot indicated that a linear model is appropriate. Write a sentence summarizing what $R^2$ says about this regression.

A) Differences in electricity usage explain 28.2% of the variation in the number of house.
B) Size differences explain 71.8% of the variation in the number of homes.
C) Differences in electricity usage explain 71.8% of the variation in the size of house.
D) Size differences explain 71.8% of the variation in electricity usage.
E) Size differences explain 28.2% of the variation in electricity usage.
21) The relationship between the number of games won by a minor league baseball team and the average attendance at their home games is analyzed. A regression to predict the average attendance from the number of games won has an $R^2 = 29.9\%$. The residuals plot indicated that a linear model is appropriate. What is the correlation between the average attendance and the number of games won.

A) 0.837  
B) 0.701  
C) 0.299  
D) 0.089  
E) 0.547

Use the model to make the appropriate prediction.

22) A golf ball is dropped from 15 different heights (in inches) and the height of the bounce is recorded (in inches.) The regression analysis gives the model bounce $= -0.2 + 0.72$ drop. Predict the height of the bounce if dropped from 81 inches.

A) 58.12 inches  
B) 58.32 inches  
C) 81.52 inches  
D) 58.52 inches  
E) 112.78 inches

23) The relationship between the selling price (in dollars) of used Ford Escorts and their age (in years) is analyzed. A regression analysis to predict the price from the age gives the model price $= 14,396 - 1448$ age. You want to sell a 15 year old Escort. Use the model to determine an appropriate price. Explain any problems.

A) $21,720 There is no way the car is worth this much.  
B) - $7324 You won’t sell a car for a negative amount. The model doesn’t give meaningful prices for Escorts this old.  
C) - $21,720 You won’t sell a car for a negative amount. The model doesn’t give meaningful prices for Escorts this old.  
D) $10 The car should be worth more than this.  
E) - $36,116 There are no problems with this prediction.

Answer the question appropriately.

24) A random sample of records of electricity usage of homes in the month of July gives the amount of electricity used and size (in square feet) of 135 homes. A regression was done to predict the amount of electricity used (in kilowatt-hours) from size. The residuals plot indicated that a linear model is appropriate. The model is usage $= 1254 + 0.7$ size. What would a negative residual mean for people living in a house that is 2284 square feet?

A) They are using more electricity than expected.  
B) Their house is smaller than expected.  
C) Their house is bigger than expected.  
D) They are using the least amount of electricity of all of the houses sampled.  
E) They are using less electricity than expected.
Explain what is wrong with each interpretation. Assume calculations are done correctly.

25) A psychologist does an experiment to determine whether an outgoing person can be identified by his or her handwriting. She claims that the $R^2$ of 89% shows that this linear model is appropriate.
   A) An $R^2$ this high means there is a very weak linear association and the model is probably inappropriate.
   B) There is nothing wrong with the interpretation.
   C) $R^2$ does not tell whether the model is appropriate, but measures the strength of the linear relationship. High $R^2$ could also be due to an outlier.
   D) This $R^2$ means that 89% of the dependent values will fall within one standard deviation of the mean and tells nothing about the appropriateness of the model.
   E) $R^2$ does not tell whether the model is appropriate, but gives the percentage of data points that are close to the model. You can sometimes have a high $R^2$ with a nonlinear relationship.

Answer the question appropriately.

26) A correlation of zero between two quantitative variables means that
   A) None of the above.
   B) we have done something wrong in our calculation of r.
   C) re-expressing the data will guarantee a linear association between the two variables.
   D) there is no association between the two variables.
   E) there is no linear association between the two variables.

27) A residuals plot is useful because
   I. it will help us to see whether our model is appropriate.
   II. it might show a pattern in the data that was hard to see in the original scatterplot.
   III. it will clearly identify influential points.
   A) I and III only
   B) I and II only
   C) I only
   D) I, II, and III
   E) II only
Solve the problem.

28) The figure below shows the life expectancy for persons living in the United States.

![Life Expectancy Graph]

The regression analysis of the data yields the following values:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-353.87</td>
</tr>
<tr>
<td>Year</td>
<td>0.2157</td>
</tr>
</tbody>
</table>

$R^2 = 0.9539$

Use the regression model to predict the life expectancy (to the nearest year) in 2007.

A) 82 years  B) 78 years  C) 79 years  D) 81 years  E) 77 years
29) The figure below shows the association between life expectancy and infant mortality for several different countries. Also shown is the equation and correlation from a regression analysis. What is the correct conclusion to draw from the figure?

A) High infant mortality is causing reduced life expectancy, probably because of the increased emotional stress exerted on parents who have lost a child at birth.

B) The association must be coincidental. I would expect the association to have a positive slope, not the negative one illustrated above.

C) Those countries with low life expectancies clearly have no regard for children or expectant mothers.

D) Countries that have low life expectancies and high infant mortality rates seem to have less regard for the sanctity of human life.

E) While there appears to be a very strong association, there is probably not a cause-and-effect relationship between infant mortality and life expectancy. Access to basic health care is probably a lurking variable that drives both life expectancy and infant mortality.

30) Which statement about influential points is true?

I. Removal of an influential point changes the regression line.

II. Data points that are outliers in the horizontal direction are more likely to be influential than points that are outliers in the vertical direction.

III. Influential points have large residuals.

A) I and III

B) I and II

C) I only

D) I, II, and III

E) II and III
31) Two variables that are actually not related to each other may nonetheless have a very high correlation because they both result from some other, possibly hidden, factor. This is an example of

A) leverage.
B) an outlier.
C) extrapolation.
D) regression.
E) a lurking variable.

32) If the point in the upper left corner of the scatterplot is removed, what will happen to the correlation (r) and the slope of the line of best fit (b)?

A) They will not change.
B) Both will decrease.
C) r will decrease and b will increase
D) r will increase and b will decrease.
E) Both will increase.

Provide an appropriate response.

33) At a college there are 120 freshmen, 90 sophomores, 110 juniors, and 80 seniors. A school administrator selects a random sample of 12 of the freshmen, a random sample of 9 of the sophomores, a random sample of 11 of the juniors, and a random sample of 8 of the seniors. She then interviews all the students selected. Identify the type of sampling used in this example.

A) Attempted census
B) Stratified sampling
C) Voluntary response sampling
D) Systematic sampling
E) Cluster sampling
34) A blockbuster movie had just premiered on exactly 1859 screens throughout the country. Within the first week, a few theaters were chosen at random, and then theater management interviewed approximately 5% of the theatergoers at those selected theaters. All these theatergoers loved the movie, therefore management decided to increase the number of showings per week. Identify the type of sampling used in this example.

A) Voluntary response sampling  
B) Convenience sampling  
C) Simple random sampling  
D) Multistage sampling  
E) Systematic sampling

35) A newly-premiered play just ended that evening at a local theater. Theater management briefly interviews every fifth person leaving the theater to see if that person will recommend the play at that theater to other people. Identify the type of sampling used in this example.

A) Stratified sampling  
B) Cluster sampling  
C) Systematic sampling  
D) Convenience sampling  
E) Voluntary response sampling

Identify the bias.

36) Inside the boxes of a new brand of cereal, is a short survey that can be mailed back for free to the manufacturer. The survey asks the consumer if he or she likes the cereal or not. What, if any, will be the most noticeable bias for this survey?

A) Voluntary response bias  
B) Undercoverage of the population  
C) Nonresponse bias  
D) Response bias  
E) There does not seem to be any bias

Answer the question.

37) The human resources department of a large, well-known telecommunications firm would like to know the job satisfaction of the employees working at its company, but the HR department is restricted by company policy to use only email to distribute and collect any type of employee questionnaire or survey. What sampling method is HR being forced to rely upon?

A) Attempted census  
B) Convenience sampling  
C) Stratified sampling  
D) Voluntary response sampling  
E) Cluster sampling
38) The human resources department of a large, well-known telecommunications firm is behind schedule in writing a report about the job satisfaction of the company’s employees. In an effort to quickly sample some people, the HR manager orders his subordinates to give out paper surveys "to anyone you see sitting at their desk", and the survey, completed or not, will be picked up in ten minutes. Only the completed surveys will be used in the report. What sampling method best describes what this HR manager is doing?

A) Convenience sampling
B) Stratified sampling
C) Cluster sampling
D) Voluntary response sampling
E) Attempted census

39) The human resources department of a large, well-known telecommunications firm is behind schedule in sampling the job satisfaction of the company’s employees. In an effort to catch up, the HR manager quickly goes down an alphabetical list of employees and e-mails a survey to every tenth employee. An neutral third party collects all surveys and ensures all of the selected employees respond to the survey. What sampling method best describes what the HR manager is doing?

A) Convenience sampling
B) Stratified sampling
C) Voluntary response sampling
D) Systematic sampling
E) Cluster sampling

40) This state’s largest university is comprised of several different colleges, institutes, and schools of study. These colleges, institutes, and schools are spread out throughout the city and nearby suburbs. The president of the university is curious about the average cost for a full-time, undergraduate student to attend one semester. All relevant costs to the student will be counted, including tuition, room and board, transportation, books, and so on. Tuition varies greatly within this university, and there are significant population differences among the colleges, institutes, and schools. What would be the most appropriate sampling method to use in order to estimate an average cost of attending this university for one semester?

A) Simple random sampling
B) Stratified sampling
C) Attempted census
D) Voluntary response sampling
E) Convenience sampling
Provide an appropriate response.

41) Hoping to get information that would allow them to negotiate new rates with their advertisers, Natural Health magazine phoned a random sample of 600 subscribers. 64% of those polled said they use nutritional supplements. Which is true?
   I. The population of interest is the people who read this magazine.
   II. "64%" is not a statistic; it's the parameter of interest.
   III. This sampling design should provide the company with a reasonably accurate estimate of the percentage of all subscribers who use supplements.

A) I, II, and III
B) II and III only
C) I and III only
D) I only
E) I and II only

A designed experiment is described. Identify the specified element.

42) A researcher wants to investigate whether different forms of exercise can be used to help hyperactive children. A group of 90 children is divided into two groups according to age - those aged 9-12 and those aged 5-8. Within each age group the children are randomly assigned to one of three groups. The first group will just do their normal exercise. The second group will be given an additional exercise routine (moderate). The third group will be given an additional exercise routine (strenuous). At the end of a four month period parents will be asked to evaluate their children’s progress. Identify the factor(s) and any blocking variables and the number of levels for each.

A) Factor: exercise (3 levels), Blocking variable: age (2 levels)
B) Normal exercise and age 5-8, additional exercise (moderate) and age 5-8, additional exercise (strenuous) and age 5-8, normal exercise and age 9-12, additional exercise (moderate) and age 9-12, additional exercise (strenuous) and age 9-12 (6 levels)
C) Blocking variable: exercise (3 levels), Factor: age (2 levels)
D) Factor: Normal exercise, additional exercise (moderate), additional exercise (strenuous) (3 levels)
E) Factor: exercise (3 levels), Factor: age (2 levels)
A new type of pain reliever is administered to 30 consenting post-operative patients in various hospitals. Although the pain reliever has already been tested for safety and effectiveness, this experiment is to observe and categorize any side-effects. Because of maturity and body mass, it is decided to test the adults separately from the children.

The grouping of the adults separate from the children is an example of what?

A) Blocking
B) Matching
C) Reduction of confounding factors
D) Controlling
E) Stratifying

A pharmaceutical company will be testing a new "one-dosage fits all" medication for treating depression. The pharmaceutical company has 100 willing test subjects that will undergo a 30-day, double-blind experiment. The experiment uses the following diagram as a guide for administering the medication at various doses, from 5% to 125%.

What is the most obvious factor missing from this proposed experiment?

A) A 100% dosage
B) A control group
C) Blocking must be used.
D) Stratifying must be used.
E) Nothing is missing. Basic variations have been accounted for.
Describe the design of the experiment (completely randomized or blocked).

45) A researcher wants to investigate the effectiveness of a cholesterol-lowering medication. The 450 participants are first divided into two groups - those suffering from diabetes and those not suffering from diabetes. Within each group the participants are randomly assigned to one of three groups. The first group will receive a low dose of the medication, the second group will receive a high dose of the medication, and the third group will receive a placebo. At the end of a six-month period, each person’s cholesterol will be tested and the change in their cholesterol level will be recorded.

A) Completely randomized over two factors (medication and diabetes), blocked by cholesterol level
B) Completely randomized over one factor (diabetes), blocked by medication
C) Completely randomized over one factor (medication), blocked by diabetes and cholesterol level
D) Completely randomized over one factor (medication), blocked by diabetes
E) Completely randomized over two factors (medication and diabetes)

46) An education researcher was interested in examining the effect of the teaching method and the teacher on students’ reading levels. 257 students participated in an experiment. There were two different teachers (Juliana and Felix) and three different teaching methods (A, B, and C). Students were randomly assigned to a teaching method and teacher. At the end of the semester the students took a reading test on which they received a score out of 20. Students who studied with Felix using method B achieved the highest scores.

A) Completely randomized over one factor (teaching method)
B) Completely randomized over one factor (teacher), blocked by teaching method
C) Completely randomized over two factors (teaching method and teacher)
D) Completely randomized over one factor (teaching method), blocked by teacher
E) Blocked by teacher, blocked by teaching method

A designed experiment is described. Identify the specified element.

47) An education researcher was interested in examining the effect of the teaching method and the teacher on students’ reading levels. 257 students participated in an experiment. There were two different teachers (Juliana and Felix) and three different teaching methods (A, B, and C). Students were randomly assigned to a teaching method and teacher. At the end of the semester the students took a reading test on which they received a score out of 20. Students who studied with Felix using method B achieved the highest scores. Identify the factor(s) in the experiment and the number of levels for each.

A) Teaching method (3 levels) and teacher (2 levels)
B) Juliana and method A, Juliana and method B, Juliana and method C, Felix and method A, Felix and method B, Felix and method C (6 levels)
C) Teaching method (3 levels), teacher (2 levels), score on reading test (20 levels)
D) The teacher and teaching method combination (6 levels)
E) Score on reading test (20 levels)
In a clinical trial, 780 participants suffering from high blood pressure were randomly assigned to one of three groups. Over a one-month period, the first group received a low dosage of an experimental drug, the second group received a high dosage of the drug, and the third group received a placebo. The diastolic blood pressure of each participant was measured at the beginning and at the end of the period and the change in blood pressure was recorded. The biggest decrease in blood pressure was for those who received the low dosage of the drug. Identify the response variable measured.

A) The dosage of the drug
B) Change in diastolic blood pressure
C) The treatment received (placebo, low dosage, high dosage)
D) The participants in the experiment
E) The one-month period

An observational study is described. Identify the specified element.

Among a group of Canadian men who were tracked for ten years, those who had scored over 130 on intelligence tests were more likely to suffer severe depression than those who had scored below 130 on intelligence tests. Identify the parameter of interest.

A) Number of years since last episode of severe depression
B) Score on intelligence test
C) Gender
D) Nationality
E) Risk of suffering from severe depression

Among a group of married women who were tracked for ten years, those who worked full time were more likely to divorce than those who did not work full time. Identify the subjects studied.

A) Divorced women
B) Married women working full time or not
C) Women who work full time
D) Divorced women who work full time
E) Married women who work full time
52) In a group of 500 women, those who smoked moderately did worse on tests of reaction time than those who did not smoke. Determine the nature and scope of the conclusion the study can reach.

A) Since there is no random assignment, there is no way to know that smoking causes subjects to do worse on tests of reaction time; there may have been lurking variables.
B) Smoking moderately causes women to do worse on tests of reaction time.
C) Since there is no random assignment, there is no way to know that smoking causes subjects to do worse on tests of reaction time; there may have been confounding variables.
D) Since the study involved only women, there is no way to know that smoking causes people to do worse on tests of reaction time.
E) Smoking moderately causes people to do worse on tests of reaction time.

Provide an appropriate response.

53) Which statement about bias is true?

I. Bias results from random variation and will always be present.
II. Bias results from a sampling method likely to produce samples that do not represent the population.
III. Bias is usually reduced when sample size is larger.

A) III only
B) I only
C) I and III only
D) II and III only
E) II only

54) Suppose the state decides to randomly test high school wrestlers for steroid use. There are 16 teams in the league, and each team has 20 wrestlers. State investigators plan to test 32 of these athletes by randomly choosing two wrestlers from each team. Is this a simple random sample?

A) Yes, because stratified samples are a type of simple random sample.
B) Yes, because each wrestler is equally likely to be chosen.
C) No, because not all possible groups of 32 wrestlers could have been the sample.
D) Yes, because the wrestlers were chosen at random.
E) No, because a random sample of teams was not first chosen.

55) Suppose your local school district decides to randomly test high school students for attention deficit disorder (ADD). There are three high schools in the district, each with grades 9-12. The school board pools all of the students together and randomly samples 250 students. Is this a simple random sample?

A) Yes, because they could have chosen any 250 students from throughout the district.
B) No, because we can’t guarantee that there are students from each school in the sample.
C) Yes, because each student is equally likely to be chosen.
D) Yes, because the students were chosen at random.
E) No, because we can’t guarantee that there are students from each grade in the sample.
Answer the question.

56) The human resources department of a large, well-known telecommunications firm would like to know the job satisfaction of the employees working at its company. This HR department hires an outside, impartial consultant to sample its management and its non-management employees separately. For this scenario, what is the best sampling method to use?

A) Cluster sampling
B) Stratified sampling
C) Convenience sampling
D) Attempted census
E) Voluntary response sampling

Identify the bias.

57) Management at a particular post office wants to know about the quality of its face-to-face service at its customer service windows. For one week, the post office will ask the customers directly. After a customer is helped, a different employee asks the customer a short question about his or her customer service experience. The customer’s comments are then transcribed by a secretary. This survey runs smoothly and a lot of customer comments are collected. What, if any, is the most noticeable bias for this survey regarding customer service at this post office for that week?

A) Undercoverage of the population
B) Voluntary response bias
C) Response bias
D) Nonresponse bias
E) There does not seem to be any bias.

58) A newspaper reporter decided to sample residents of all major metropolitan areas in the entire country. This sampling asked if the residents liked living in their city, or are they considering moving, and if so, where? Unfortunately, this survey went over budget and was stopped, and only the east coast of the country was adequately interviewed. With regards to having this survey labeled as representing the entire country, what, if any, will be the most noticeable bias for this survey as it is now?

A) Undercoverage of the population
B) Voluntary response bias
C) Nonresponse bias
D) Response bias
E) There does not seem to be any bias.
59) A cereal manufacturer randomly samples households across the country by telephone. The survey asks the consumer to explain honestly in fifty words or less why he or she likes or dislikes its brand of cereals. It is also stated that for those who participate in this telephone survey, a random drawing will be held for an all-expense paid, two-week cruise. What, if any, will be the most noticeable bias for this survey?

A) Response bias  
B) Nonresponse bias  
C) Undercoverage of the population  
D) Voluntary response bias  
E) There does not seem to be any bias.

Provide an appropriate response.

60) Each person in a simple random sample of 2,000 received a survey, and 317 people returned their survey. How could nonresponse cause the results of the survey to be biased?

A) Those who did not respond reduced the sample size, and small samples have more bias than large samples.  
B) Those who did not respond caused a violation of the assumption of independence.  
C) Those who did not respond represent a stratum, changing the simple random sample into a stratified random sample.  
D) Those who did not respond were indistinguishable from those who did not receive the survey.  
E) Those who did respond may differ in some important way from those who did not respond.

61) At the local college a survey was being done on whether or not the students liked the cafeteria food. The survey was located in the college newspaper and was to be filled out and sent to the editing office. Identify the type of sampling used in this example.

A) Stratified sampling  
B) Voluntary response sampling  
C) Cluster sampling  
D) Attempted census  
E) Systematic sampling

62) An education researcher randomly selects 38 schools from one school district and interviews all the teachers at each of the 38 schools. Identify the type of sampling used in this example.

A) Stratified sampling  
B) Systematic sampling  
C) Cluster sampling  
D) Voluntary response sampling  
E) Attempted census
63) Jennifer is a quality control inspector for a well-known computer modem manufacturer. Jennifer oversees five assembly lines, each assembly line produces the same number of modems per day. Jennifer randomly selects one assembly line each morning and performs further sampling and quality control procedures on that assembly line's modems for the rest of the day. This morning, Jennifer randomly selected the modems of assembly line #5 for quality control inspection. Furthermore, Jennifer will then perform systematic sampling on that assembly line's modems. What has been excluded from being in the sampling frame today?

A) Some of the modems of assembly line #5
B) A sampling method for assembly line #5
C) The modems of assembly lines #1, 2, 3, and 4
D) An unbiased sample for assembly line #5
E) Nothing has been excluded from the sampling frame.

64) A regression analysis of students' college grade point averages (GPAs) and their high school GPAs found $R^2 = 0.311$. Which of these is true?

I. High school GPA accounts for 31.1% of college GPA.
II. 31.1% of college GPAs can be correctly predicted with this model.
III. 31.1% of the variance in college GPA can be accounted for by the model

A) None  B) III only  C) II only  D) I and II  E) I only

65) The relationship between the number of games won by a minor league baseball team (x) and the average attendance at their home games (y) is analyzed. The mean number of games won was 66 with a standard deviation of 17. The mean attendance was 6177 with a standard deviation of 1480. The correlation between the games won and attendance was 0.46.

\[ \hat{y} = 431 + 87.1 \text{ wins} \]
\[ \hat{y} = 6180 + 0.00528 \text{ wins} \]
\[ \hat{y} = 1880 + 65.0 \text{ wins} \]
\[ \hat{y} = 3070 + 140 \text{ wins} \]
\[ \hat{y} = 3530 + 40.0 \text{ wins} \]
Answer the question appropriately.

66) A golf ball is dropped from 15 different heights (in inches) and the height of the bounce is recorded (in inches.) The regression analysis gives the model \( \text{bounce} = -0.3 + 0.65 \text{drop} \). Interpret the meaning of the \( y \)-intercept.

A) According to the model, a ball dropped from 0.65 inches high will bounce - 0.3 inches. (This may not actually happen.)
B) According to the model, a ball dropped from - 0.3 inches high will bounce 0 inches. (This may not actually happen.)
C) According to the model, a ball dropped from 0 inches high will bounce - 0.3 inches. (This may not actually happen.)
D) According to the model, a ball dropped from 0 inches high will bounce 0.65 inches. (This may not actually happen.)
E) According to the model, a ball dropped from 0.65 inches high will bounce 0 inches. (This may not actually happen.)

Provide an appropriate response.

67) The correlation coefficient between the hours that a person is awake during a 24-hour period and the hours that same person is asleep during a 24-hour period is most likely to be

A) near - 0.8
B) exactly +1.0
C) near 0
D) near +0.8
E) exactly - 1.0

68) All but one of the statements below contain a mistake. Which one could be true?

A) The correlation between gender and age is - 0.171.
B) The correlation between weight and length of foot is 0.488.
C) If the correlation between blood alcohol level and reaction time is 0.73, then the correlation between reaction time and blood alcohol level is - 0.73.
D) The correlation between the breed of a dog and its weight is 0.435.
E) The correlation between height and weight is 0.568 inches per pound.

Find the lurking variable.

69) A study shows that the amount of chocolate consumed in Canada and the number of automobile accidents is positively related. Find the lurking variable, if there is one.

A) Vacation
B) Speed
C) Population growth
D) Children
E) No lurking variable
Solve the problem.

70) A Science instructor assigns a group of students to investigate the relationship between their grade point average and their consistency in studying. The correlation between GPA and studying consistency is 0.849. Assume that the relation is linear. Describe the association.

A) Strong linear relation in a positive direction
B) Weak linear relation in a positive direction
C) Strong curved relation in a positive direction
D) No evidence of relation
E) Strong linear relation in a negative direction

71) Here are some scores from a recent Mathematics exam: 95.5, 65.9, 93.2, 80.6, 56.8, 50, 86.4, 54.5, 40.9, 77.3, 79.5, 10, 65.9, 70.5, 15, 77.3, 81.8, 12, 50, 79.5, 60.2. Which is a better summary of the scores, the mean or the median? Explain.

A) Mean, the data is so skewed to the right
B) Mean, the data is so skewed to the left
C) Median, the data is so skewed to the left
D) Median, the data is so skewed to the right
E) Either, the data is symmetric

72) The annual snowfall in a town can be represented by the Normal model N(44, 7.3). What amount of snowfall would you expect to be unusually low for this town?

A) 65.9 inches
B) 22.1 inches
C) 40.35 inches
D) 51.3 inches
E) 36.7 inches

Solve the problem. Round to the nearest tenth.

73) Based on the Normal model for car speeds on an old town highway N(77, 9.1), what is the cutoff value for the highest 15% of the speeds?

A) about 11.6 mph
B) about 63.1 mph
C) about 86.5 mph
D) about 67.5 mph
E) about 65.5 mph

Solve the problem.

74) A bank’s loan officer rates applicants for credit. The ratings can be described by a Normal model with a mean of 200 and a standard deviation of 50. If an applicant is randomly selected, what percentage can be expected to be between 200 and 275?

A) 43.32%   B) 6.68%   C) 5.00%   D) 42.37%   E) 93.32%
Provide an appropriate response.

75) Which of the following summaries are changed by adding a constant to each data value?
   I. the mean
   II. the median
   III. the standard deviation
   A) I and III
   B) I and II
   C) I only
   D) I, II, and III
   E) III only

Solve the problem.

76) Office workers were asked how long it took them to travel to work one morning. Here is the stem-and-leaf display.

   2 | 0 0 0 2 3 4 4 5 7 8
   3 | 0 2 5 7
   4 | 1 2 7 8 9
   5 | 0 2 8
   6 | 0 5

Without actually finding the mean and the median, would you expect the mean to be higher or lower than the median?
   A) Higher, because the data are skewed to the left.
   B) Lower, because the data are skewed to the right.
   C) Lower, because the data are skewed to the left.
   D) Higher, because the data are skewed to the right.
   E) Neither, because the mean would be equal to the median.

Provide an appropriate response.

77) In an experiment the primary purpose of blinding is to reduce...
   A) variation.
   B) randomness.
   C) confounding.
   D) undercoverage.
   E) bias.
78) Does donating blood lower cholesterol levels? 50 volunteers have a cholesterol test, then donate blood, and then have another cholesterol test. Which aspect of experimental design is present?

A) a control group
B) randomization
C) a placebo
D) blinding
E) none of these

79) Twenty dogs and 20 cats were subjects in an experiment to test the effectiveness of a new flea control chemical. Ten of the dogs were randomly assigned to an experimental group that wore a collar containing the chemical, while the others wore a similar collar without the chemical. The same was done with the cats. After 30 days veterinarians were asked to inspect the animals for fleas and evidence of flea bites. This experiment is...

A) completely randomized with two factors.
B) randomized block, blocked by type of collar.
C) completely randomized with one factor: the type of collar.
D) randomized block, blocked by species.
E) completely randomized with one factor: the species of animal.

Solve the problem.

80) A survey revealed that 34% of people are entertained by reading books, 47% are entertained by watching TV, and 19% are entertained by both books and TV. What is the probability that a person will be entertained by either books or TV?

A) 1 B) 0.6 C) 0.81 D) 0.62 E) 0.19

Find the indicated probability.

81) A group of volunteers for a clinical trial consists of 81 women and 77 men. 21 of the women and 19 of the men have high blood pressure. If one of the volunteers is selected at random find the probability that the person is a man given that they have high blood pressure.

A) 0.253 B) 0.487 C) 0.247 D) 0.475 E) 0.120

82) The table shows the political affiliation of voters in one city and their positions on stronger gun control laws.

<table>
<thead>
<tr>
<th>Stronger Gun Control</th>
<th>Favor</th>
<th>Oppose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Republican</td>
<td>0.110</td>
<td>0.31</td>
</tr>
<tr>
<td>Democrat</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td>Other</td>
<td>0.10</td>
<td>0.08</td>
</tr>
</tbody>
</table>

What is the probability that a voter who favors stronger gun control laws is a Republican?

A) 0.256 B) 0.430 C) 0.420 D) 0.110 E) 0.262
83) The sale bin in a clothing store contains an assortment of t-shirts in different sizes. There are 7 small, 8 medium, and 4 large shirts. Alan is looking for a large shirt. He starts grabbing shirts one at a time and checking the size. After he checks each shirt, he leaves it outside the bin. What is the probability that at least one of the first four shirts he checks is a large?

A) 0.612  
B) 0.117  
C) 0.388  
D) 0.352  
E) 0.648

Determine whether the events are disjoint and give a reason.

84) Melissa is looking for the perfect man. She claims that at her college 27% of men are smart, 22% are funny, and 14% are both smart and funny. If Melissa is right, are being smart and being funny disjoint events?

A) No, 14% are both smart and funny  
B) No, the probability that a man is funny is 0.22, but the probability that a man is funny given that he is smart is 0.52.  
C) Yes, no man is both smart and funny  
D) Yes, because P(S or F) = P(S) + P(F)  
E) Yes, the probability that a man is smart is the same as the probability that a man is smart given that he is funny.

Use a tree diagram to find the indicated probability.

85) Shameel has a flight to catch on Monday morning. His father will give him a ride to the airport. If it rains, the traffic will be bad and the probability that he will miss his flight is 0.04. If it doesn't rain, the probability that he will miss his flight is 0.02. The probability that it will rain on Monday is 0.20. What is the probability that Shameel misses his flight?

A) 0.0080  
B) 0.030  
C) 0.04  
D) 0.0160  
E) 0.0240

Find the expected value of the random variable. Round to three decimal places.

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

<table>
<thead>
<tr>
<th>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>0.26</td>
<td>0.36</td>
<td>0.14</td>
<td>0.10</td>
</tr>
</tbody>
</table>

A) 9  
B) 9.12  
C) 5.73  
D) 7.98  
E) 8.4
Provide an appropriate response.

87) 3. A survey of some AP Stats students recorded gender and whether the student was left or right-handed. Results were summarized in a table like the one shown. If it turned out that handedness was independent of gender, how many of the AP Stat students were lefty girls?

<table>
<thead>
<tr>
<th></th>
<th>Lefty</th>
<th>Righty</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy</td>
<td>?</td>
<td>66</td>
<td>20</td>
</tr>
<tr>
<td>Girl</td>
<td>54</td>
<td>?</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>120</td>
<td>100</td>
<td>220</td>
</tr>
</tbody>
</table>

A) 7  
B) 4  
C) 9  
D) It cannot be determined.  
E) 10

Use a tree diagram to find the indicated probability.

88) 3.3% of a population are infected with a certain disease. There is a test for the disease, however the test is not completely accurate. 94% of those who have the disease will test positive. However 4.4% of those who do not have the disease will also test positive (false positives). What is the probability that any given person will test positive? Round your answer to three decimal places if necessary.

A) 0.492  
B) 0.043  
C) 0.074  
D) 0.031  
E) 0.94

Find the expected value of the random variable. Round to three decimal places.

89) A contractor is considering a sale that promises a profit of $29,000 with a probability of 0.7 or a loss (due to bad weather, strikes, and such) of $3000 with a probability of 0.3. What is the expected profit?

A) $20,300  
B) $19,400  
C) $22,400  
D) $26,000  
E) $21,200

90) You pick a card from a deck. If you get a club, you win $90. If not, you get to draw again (after replacing the first card). If you get a club the second time, you win $30. If not, you lose. Find the expected amount you will win.

A) $36.56  
B) $28.13  
C) $32.34  
D) $30.00  
E) $45.00

Find the standard deviation of the random variable. Round to two decimal places if necessary.

91) 

<table>
<thead>
<tr>
<th>x</th>
<th>P(X = x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

A) 0.72  
B) 0.66  
C) 0.68  
D) 0.69  
E) 0.44

35
Find the indicated probability.

92) In the 4 ×100 relay event, each of four runners runs 100 meters. A college team is preparing for a competition. The means and standard deviations of the times (in seconds) of their four runners are shown in the table:

<table>
<thead>
<tr>
<th>Runner</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.01</td>
<td>0.12</td>
</tr>
<tr>
<td>2</td>
<td>12.40</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>11.95</td>
<td>0.10</td>
</tr>
<tr>
<td>4</td>
<td>11.81</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Find the probability that in the competition they will beat their previous best time of 47.668 sec. Assume that the runners' performances are independent and that each runner's time follows a Normal model.

A) 0.989   B) 0.011   C) 0.036   D) 0.018   E) 0.982

93) Miguel buys a large bottle and a small bottle of juice. The amount of juice that the manufacturer puts in the large bottle is a random variable with a mean of 1016 ml and a standard deviation of 8 ml. The amount of juice that the manufacturer puts in the small bottle is a random variable with a mean of 510 ml and a standard deviation of 5 ml. If the total amount of juice in the two bottles can be described by a normal model, what's the probability that the total amount of juice in the two bottles is more than 1540.2 ml?

A) 0.055   B) 0.919   C) 0.933   D) 0.067   E) 0.081

Provide an appropriate response.

94) Some marathons allow two runners to "split" the marathon by each running a half marathon. Alice and Sharon plan to split a marathon. Alice's half-marathon times average 92 minutes with a standard deviation of 4 minutes, and Sharon's half-marathon times average 96 minutes with a standard deviation of 2 minutes. Assume that the women's half-marathon times are independent. The expected time for Alice and Sharon to complete a full marathon is 92 + 96 = 180 minutes. What is the standard deviation of their total time?

A) 4.5 minutes   B) It cannot be determined   C) 20 minutes   D) 2 minutes   E) 6 minutes

Find the indicated probability.

95) Suppose a computer chip manufacturer rejects 3% of the chips produced because they fail presale testing. What is the probability that the seventh chip you test is the first bad one you find?

A) 0.21   B) 0.0007   C) 0.0219   D) 0.0250   E) 0.03

96) An archer is able to hit the bull's-eye 49% of the time. If she shoots 10 arrows, what is the probability that she gets exactly 4 bull's-eyes? Assume each shot is independent of the others.

A) 0.7870   B) 0.2130   C) 0.1267   D) 0.0010   E) 0.0576
Find the probability of the outcome described.

97) A test consists of 10 true/false questions. If a student guesses on each question, what is the probability that the student will answer at least 9 questions correctly.

A) 0.011  B) 0.010  C) 0.999  D) 0.9  E) 0.001

98) A car insurance company has determined that 5% of all drivers were involved in a car accident last year. If 15 drivers are randomly selected, what is the probability that at least 3 were involved in a car accident last year?

A) 0.0307  B) 0.9638  C) 0.4995  D) 0.9693  E) 0.0362

Provide an appropriate response.

99) A researcher wants to compare the effect of a new type of shampoo on hair condition. The researcher believes that men and women may react to the shampoo differently. Additionally, the researcher believes that the shampoo will react differently on hair that is dyed. The subjects are split into four groups: men who dye their hair, men who do not dye their hair, women who dye their hair, women who do not dye their hair. Subjects in each group are randomly assigned to the new shampoo and the old shampoo. This experiment...

A) is completely randomized.
B) has two factors (shampoo type and whether hair is dyed) blocked by gender.
C) has three factors (shampoo type, gender, whether hair is dyed).
D) has two factors (gender and whether hair is dyed) blocked by shampoo type.
E) has one factor (shampoo type), blocked by gender and whether hair is dyed.