1. The box plot of a given set of data is shown.

Which of the following statements is true?

(A) The data are strongly skewed to the right with two outliers.
(B) The data are uniformly distributed and the outliers do not affect the median.
(C) The data are normally distributed with two outliers.
(D) Only one of the outliers is less than \{Quartile1 - 1.5\times InterQuartile Range\}.
(E) The data are strongly skewed to the left with two outliers.

2. A radio station is interested in predicting the proportion of registered voters who support an increase in the state sales tax. Listeners were asked to go to the station's web site and indicate whether they favored an increase to support additional regional parks. 1744 listeners logged on and 922 (53%) were against the increase. The population of interest is

(A) the 1744 listeners who logged on.
(B) all registered voters.
(C) all people who listen to the station.
(D) only that portion of the 1744 listeners who voted that are actually registered to vote.
(E) the 922 listeners who were against the proposed increase.

3. A recent newspaper article cited a poll that reported 4% of the adults in Wyoming believed that overcrowding was a serious problem while 36% of respondents in California had the same opinion. The reporter said that an average of 20% of the population of the two states felt that overcrowding was therefore a problem. A statistician criticized the article stating the conclusion was inaccurate. Which of the following reasons support the statistician's reasoning?

(A) The number of people in the two states are very different.
(B) The polls did not represent an SRS from each state. Wyoming has less overcrowding than California, so the results are biased on this issue.
(C) The correct value should have been 40%.
(D) The sample size from each state was not large enough to support a 95% confidence interval.
(E) The standard deviations of the two states are different.
4. One of the topics studied by The Marriage Project at Rutgers University was "I am confident that I can find a marriage partner that will meet my needs in a marriage." Preliminary research showed that 88% of respondents agreed with the statement. How many people must be sampled to estimate the true proportion within ±0.04 with 95% confidence?

\[ 0.04 = \sqrt{\frac{0.88 \cdot 0.12}{n}} \cdot 1.96 \]

(A) 178  
(B) 254  
(C) 391  
(D) 601  
(E) 1201

5. Sixty-eight people from a random sample of 128 residents of Uppsala, Sweden had blue eyes. 45 people from a random sample of 110 people from Preston, England had blue eyes. Let \( p_1 \) represent the proportion of people in Uppsala with blue eyes and \( p_2 \) represent the proportion of people in Preston with blue eyes. What is the P-value of the hypothesis test: \( H_0: p_1 = p_2 \) vs. \( H_a: p_1 \neq p_2 \)?

\[
\begin{align*}
\chi^2 &= \frac{68 - 45}{128 + 110} = 1.88 \\
\chi^2 &= \sqrt{\frac{113 \cdot 17.5}{238} \left( \frac{1}{128} + \frac{1}{110} \right)} \\
P(\chi^2 > 1.88) &= 0.0299 \\
\chi^2 &= 0.599
\end{align*}
\]
6. A random sample of 100 homes counted the number of television sets in each home. The results are shown in the histogram below.

![Histogram showing the distribution of television sets per home]

Based on the histogram, which of the following statements is TRUE.

I. The data are approximately normal in shape.
II. The 3 homes without television sets are outliers.
III. The 95% confidence interval is (2, 7).

(A) I only
(B) II only
(C) III only
(D) I and II only
(E) I and III only
An efficiency expert wanted to see if there is a relationship between the number people invited to a meeting and the number of minutes late that the meeting started. The table shows the results with the accompanying scatter plot in Figure 1.

<table>
<thead>
<tr>
<th>Number of people invited</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of minutes late the meeting started</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>14</td>
<td>29</td>
</tr>
</tbody>
</table>

Figure 2 represents the scatter plot with the point (10, 29) removed from the data. Which of the following is NOT true about the point (10, 29).

(A) It is an influential observation.
(B) It is a small residual.
(C) It is an outlier.
(D) The P-value of a Linear Regression t-Test changes little when the point is removed.
(E) None of the above.

8. A sports physiologist wants to compare the effects of two different exercise machines on the flexibility of gymnasts. There are 20 gymnasts available. She decides to group the gymnasts in matched pairs. Which of the following would give the least biased result?

(A) Use a random number table to divide the gymnasts into 10 separate pairs. Then assign the first person to the first machine and the second person to the second machine.
(B) Use a random number table to divide the gymnasts into 10 separate pairs. Then flip a coin to assign the each member of a pair gymnasts to the two machines.
(C) Use a random number table to divide the gymnasts into two groups. Flip a coin to assign each group to the machines.
(D) Pair the gymnasts subjectively (keeping both members of the pair as similar as possible). Use a random number table to assign each member of the pair to a machine.
(E) Use a random number table to divide the gymnasts into two groups. Have both groups use both machines, using a coin flip to decide which one machine is used first. Then do a pre-post test for each gymnast.
9. The number of goals scored per game in a full season of soccer games for a professional league is strongly skewed to the right with a mean of 2.3. An SRS of sample size \( n = 15 \) is taken from the population and the sample mean is computed. This is repeated for a total of 175 trials. Which of the following best describes the shape of the sampling distribution?

(A) Skewed to the right with a mean of 2.3 goals.

(B) Skewed to the right with a mean of \( \frac{2.3}{\sqrt{15}} \) goals.

(C) Binomially distributed with a mean of 2.3 goals.

(D) Approximately normally distributed with a mean of \( \frac{2.3}{\sqrt{15}} \) goals.

(E) Approximately normally distributed with a mean of 2.3 goals.

10. Students in a statistics class drew circles of varying diameters and counted how many Cheerios\textsuperscript{®} could be placed in the circle. The scatter plot shows the results.

![Scatter Plot](image)

The students wanted to determine an appropriate equation for the relationship between diameter and the number of Cheerios\textsuperscript{®}. The students decided to transform the data to make it appear more linear before computing a least squares regression line. Which of the following transformations would be reasonable for them to try.

I. Take the square root of the number of Cheerios\textsuperscript{®}.
II. Cube the number of Cheerios\textsuperscript{®}.
III. Take the log of number of Cheerios\textsuperscript{®}.
IV. Take the log of the diameter.

(A) I and II
(B) I and III
(C) II and III
(D) II and IV
(E) III and IV
11. To determine the viscosity of a liquid, scientists measure the length of time it takes for a fixed amount of the liquid to run through a tube. Two different types of motor oil were tested with thirty trials each. The first type of oil had a mean time of 10 seconds with a standard deviation of 2 seconds. The second type of oil had a mean of 14 seconds with a standard deviation of 3 seconds. The difference between the mean time of the two types of oil is 4 seconds with a standard deviation of $\sqrt{\frac{2^2 + 3^2}{14 + 9}}$.

(A) $\frac{1}{4}$ seconds
(B) $\frac{5}{\sqrt{20}}$ seconds
(C) $\sqrt{13}$ seconds
(D) $\sqrt{\frac{13}{30}}$ seconds
(E) $\frac{5}{\sqrt{40}}$ seconds

12. According to data from Major League Baseball, the mean salary for professional major league baseball players has been increasing for the past 15 years. A sportswriter looks at the data for all players for the 2002 season and decides to construct a 95% confidence interval for the mean salary for all players. Which of the following is the reason the writer's reasoning is incorrect?

(A) The sportswriter did not take a Simple Random Sample.
(B) A confidence interval cannot be constructed when population parameter is known.
(C) Since a very few ball players make over $20,000,000 a year, the data is highly skewed, which violates the rules of thumb to construct a confidence interval.
(D) Since the number of players is not listed, the degrees of freedom is unknown.
(E) Using the data from only one year is not sufficient to verify the trend from previous years.
13. A random variable $X$ has a probability distribution as shown in the graph. What is the probability that $X$ will be greater than or equal to 1.75?

\[ P(X) \]

\[ 0 \quad 1 \quad 2 \quad 3 \]

\[ 0.5 \]

\[ 3 - 1.75 = 1.25 \]

(A) 0.250
(B) 0.375
(C) 0.625
(D) 0.750
(E) 1.750

14. Two different brands of blood pressure cuffs were tested on two groups of patients. The results are shown in the table

<table>
<thead>
<tr>
<th></th>
<th>Brand A</th>
<th>Brand B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean</td>
<td>127 mg</td>
<td>134 mg</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>23 mg</td>
<td>20 mg</td>
</tr>
<tr>
<td>Sample size</td>
<td>20</td>
<td>17</td>
</tr>
</tbody>
</table>

What is the 95% confidence interval for the difference between the two means?

\[ df = 16 \]

(A) \(-7 \pm 3.23\)
(B) \(-7 \pm 11.63\)
(C) \(-7 \pm 12.34\)
(D) \(-7 \pm 13.85\)
(E) \(-7 \pm 14.98\)
15. At a fund raiser for charity, 15 college basketball players were randomly matched with 15 people who were not basketball players who gave a donation to the charity. Each person then shot 10 free throws and the difference between the player and the partner was recorded (College player - Partner). A plot of the data showed the differences were approximately normally distributed. The results are summarized in the computer output below.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>MEAN</th>
<th>STDEV</th>
<th>SE MEAN</th>
<th>95 PERCENT C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIFF</td>
<td>15</td>
<td>2.600</td>
<td>3.290</td>
<td>0.849</td>
<td>(0.77768, 4.42232)</td>
</tr>
</tbody>
</table>

Based on the information, which of the following statements is true?

(A) There is a 95% probability that the true mean difference between college players and the randomly selected partners is between 0.78 and 4.42.

(B) Since the confidence interval did not contain 0, it is likely that the mean number of shots made by college players is greater than the mean number of shots made by randomly selected people.

(C) A sample size of 15 is too small to draw any conclusion.

(D) Since there is no relationship between the two groups, a two sample t-test is more appropriate than a one sample t-test on the difference.

(E) All college players made more shots than their partner.

16. Suppose that 40% of the cars in a certain town are white. A person stands at an intersection waiting for a white car. Let X = the number of cars that must drive by until a white car drives by. The probability that a white car drives by before X = 5 is

\[
P(X < 5) = \sum_{x=1}^{4} \binom{4}{x} (0.4)^x (0.6)^{4-x}
\]

(A) 0.0518

(B) 0.296

(C) 0.2592

(D) 0.8704

(E) 0.9482

17. According to the Central Florida Blood Bank, 13% of first-time donors return to make a second donation within 3 months. In the fall of 2001 they tracked 6000 first-time donors and found that 891 donated a second time within 90 days. What is the probability that a random sample of 6000 first-time donors has 13% or more who return within 90 days to donate again?

(A) Less than 0.001

(B) 0.007

(C) 0.014

(D) 0.492

(E) 0.983

\[
Z = \frac{891 - 1485}{\sqrt{\frac{1485 \cdot 891}{6000}}}
\]

\[
Z = \frac{-13}{\sqrt{\frac{1485 \cdot 8515}{6000}}} = -25.813
\]
18. The mean blood pressure for 47 year old males in the United States is 139 mg with a standard deviation of 26 mg. A doctor tells his 47 year old male patient he is at the lowest 10% of all people in this population. Which of the following is nearest to the patient’s blood pressure?

(A) 96
(B) 106
(C) 108
(D) 125
(E) 127

\[ \frac{x - 139}{26} = -1.282 \]

19. A pharmacist wants to study the effect of temperature (0°F, 30°F, 60°F, and 90°F) on the potency of a headache pain reliever when it is stored for extended periods of time. 30 pills were randomly assigned to each of the temperatures for the specified storage time. The time it took for a person to gain relief from a headache was measured for each pill. Which of the following is the correct description of the treatment, experimental unit, and response?

(A) Specific temperature, relief time, pill.
(B) Specific temperature, the pharmacist, relief time.
(C) Pill, relief time, specific temperature.
(D) Random assignment, pill, relief time.
(E) Specific temperature, pill, relief time.

20. A statistics class took a random sample of the students at the school to find the proportion of those who claimed to be vegetarians. Another statistics class in another school took a similar random sample of the students at its school. The results from the first school were 12 out of 75 claimed to be vegetarians while 9 out of 45 students from the second school were vegetarians. What is the 90% confidence interval for the difference between the population proportions of the two schools that claim to be vegetarians?

(\[ \frac{12}{75} - \frac{9}{45} \pm 1.645 \sqrt{\frac{12 \times \frac{33}{75} + 9 \times \frac{36}{45}}{75 \times 45}} \]

16.9 \%-20.8%)

21. A company with 16 employees gives everyone a $2000 bonus. What will be the change in the standard deviation of the employees' income after the bonus is awarded?

(A) It will increase by $2000.
(B) It will stay the same.
(C) It will increase by $500.
(D) It will increase by $\sqrt{2000}$.
(E) It will be multiplied by $2000$. 

39
22. A local dealer has two video stores in a town, one on Foothill Drive and the other on Grand Avenue. The Foothill Drive store does 70% of the dealer’s business in the town and the Grand Avenue store does the rest. In the Foothill Drive store 40% of all rentals are DVD’s. At the Grand Ave. store 30% of all rentals are DVD’s. If a customer is selected at random, what is the probability that she will rent a DVD?

(A) 0.175
(B) 0.33
(C) 0.35
(D) 0.37
(E) 0.70

23. A community kept track of the number of people who used the local jogging track. They recorded the frequency of use for people each week.

<table>
<thead>
<tr>
<th>Number of days walked per week</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Season</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring</td>
<td>52</td>
<td>34</td>
<td>26</td>
<td>20</td>
<td>17</td>
<td>13</td>
<td>10</td>
<td>172</td>
</tr>
<tr>
<td>Summer</td>
<td>104</td>
<td>69</td>
<td>52</td>
<td>41</td>
<td>35</td>
<td>26</td>
<td>21</td>
<td>348</td>
</tr>
<tr>
<td>Fall</td>
<td>69</td>
<td>51</td>
<td>39</td>
<td>31</td>
<td>26</td>
<td>19</td>
<td>15</td>
<td>250</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>225</td>
<td>154</td>
<td>117</td>
<td>92</td>
<td>78</td>
<td>58</td>
<td>46</td>
<td>770</td>
</tr>
</tbody>
</table>

How much does the cell corresponding to 3 times per week during the summer contribute to the total chi-square statistic?

\[
\text{Expected} = \frac{348 \times 117}{770} = 52.88 \\
(\frac{52-52.88}{52.88})^2 = 0.145
\]
24. The following scatter plot shows the mean skid length for a car weighing 1500 pounds traveling at 40 mph on a new set of tires when plotted against the outside temperature (in °C). Which of the following statements is true?

![Scatter Plot]

(A) The data show a weak, positive correlation.
(B) The data show a strong, positive correlation.
(C) The data show a weak, negative correlation.
(D) The data show a strong, negative correlation.
(E) No linear relationship exists.

25. A sales person ranks in the top 5% of all sales people in a large company. If the sales in the company are normally distributed and the annual mean sales amount is $750,000 with a standard deviation of $150,000, how much does the person sell each year?

\[ \frac{X - 750,000}{150,000} = 1.645 \]

(B) $996,750

(C) $1,044,000

(D) $1,697,100

(E) $2,650,500
26. The following information is a summary about a data set from pulse rates for 3 year old children.

\[
\begin{align*}
\text{1-Var Stats} \\
f &= 97.00 \\
\Sigma x &= 2910 \\
\Sigma x^2 &= 290518 \\
S_x &= 11.93931 \\
\sigma_x &= 11.73864 \\
n &= 30 \\
\text{MinX} &= 74 \\
Q1 &= 88 \\
\text{Med} &= 98.5 \\
Q3 &= 107 \\
\text{MaxX} &= 118
\end{align*}
\]

What is the smallest integer value for an outlier on the upper end?

(A) 120  
(B) 125  
(C) 127  
(D) 136  
(E) 173

27. 25 hamsters were trained to run a maze. The times for each hamster to complete the maze were normally distributed with a mean of 48 seconds and a standard deviation of 10 seconds. What is the range for the middle 95% of all times to run the maze?

(A) 18 to 78 seconds  
(B) 28 to 68 seconds  
(C) 38 to 58 seconds  
(D) 44 to 52 seconds  
(E) 46 to 50 seconds
28. A random sample of cell phone customers for a company was taken from each of two large cities to measure cell phone usage. The number of calls in a month for each customer was recorded. The results are summarized below.

<table>
<thead>
<tr>
<th>Level</th>
<th>Minimum</th>
<th>10%</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
<th>90%</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>City A</td>
<td>1.9</td>
<td>1.9</td>
<td>4.3</td>
<td>20.2</td>
<td>58.1</td>
<td>97.4</td>
<td>100.0</td>
</tr>
<tr>
<td>City B</td>
<td>1.9</td>
<td>2.2</td>
<td>9.5</td>
<td>20.3</td>
<td>61.5</td>
<td>103.4</td>
<td>141.0</td>
</tr>
</tbody>
</table>

Means and standard deviations

<table>
<thead>
<tr>
<th>Level</th>
<th>n</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Std Err Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>City A</td>
<td>12</td>
<td>39.7</td>
<td>30.4</td>
<td>8.8</td>
</tr>
<tr>
<td>City B</td>
<td>12</td>
<td>44.6</td>
<td>31.1</td>
<td>9.0</td>
</tr>
</tbody>
</table>

Which of the following statement are correct?

I. Customers, on average, in City A make fewer cell phone calls than those in City B.
II. The data from City A is more strongly skewed than the data from City B.
III. The maximum from City B is an outlier.

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

29. A company has 500 sales people. A survey of 10 representatives found the mean years with the company was 8.7 years with $s = 3.2$ years. If the survey is an SRS of the company's sales force, what is the 95% confidence interval for the mean years of service of all sales people?

(A) $8.7 \pm 1.98$
(B) $8.7 \pm 2.09$
(C) $8.7 \pm 2.29$
(D) $8.7 \pm 2.63$
(E) $8.7 \pm 2.79$
30. A researcher wishes to test the yield of a new type of fertilizer. He divides a small field 100' plots as shown in the diagram. Each row gets a different concentration of fertilizer. The design of the experiment requires that 10 different plots be sampled.

<table>
<thead>
<tr>
<th></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>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>92</td>
<td>93</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Because there is a strong wind that blows from the east, the researcher decides to randomly select one plot in each row (1-10, 11-20, 21-30, etc). This type of sampling is known as

(A) multi-stage sample
(B) simple random sample
(C) survey sample
(D) convenience sample
(E) stratified sample
31. The resting pulse rate of 11 people was measured before and after a television commercial was shown them. The researcher wanted to check the mean difference in pulse rate for each subject. The results are shown in the following table.

<table>
<thead>
<tr>
<th>Subject</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>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse rate before</td>
<td>85</td>
<td>79</td>
<td>73</td>
<td>76</td>
<td>75</td>
<td>79</td>
<td>74</td>
<td>67</td>
<td>80</td>
<td>69</td>
<td>71</td>
</tr>
<tr>
<td>Pulse rate after</td>
<td>83</td>
<td>78</td>
<td>85</td>
<td>82</td>
<td>81</td>
<td>90</td>
<td>75</td>
<td>69</td>
<td>73</td>
<td>82</td>
<td>70</td>
</tr>
</tbody>
</table>

What is the degrees of freedom for the appropriate t-test for testing if the mean increase is greater than zero?

(A) 10
(B) 11
(C) 20
(D) 21
(E) 22

32. A researcher wanted to measure the sensitivity of a hypothesis test to detect a difference between the mean of a sample and the hypothesized mean of the item. When the difference was great, she could use a small sample size. When the difference between the means was small, she had to increase the sample size. This is an illustration of:

(A) two-sample t-test
(B) power
(C) correlation coefficient
(D) type II error
(E) convenience sampling
A strength coach wants to determine if there is a relationship between the weight of an individual and how much weight they can lift. The data was collected and analyzed using a statistics software program. The output is shown below.

<table>
<thead>
<tr>
<th>TotalWeightLifted</th>
<th>Under 150</th>
<th>Over 150</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 200</td>
<td>44</td>
<td>43</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>40.0</td>
<td>47.0</td>
<td>87.0</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>0.34</td>
<td>0.00</td>
</tr>
<tr>
<td>Over 200</td>
<td>25</td>
<td>38</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>28.0</td>
<td>34.0</td>
<td>63.0</td>
</tr>
<tr>
<td></td>
<td>0.55</td>
<td>0.47</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>81</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>69.0</td>
<td>81.0</td>
<td>150.0</td>
</tr>
</tbody>
</table>

The number of rows with at least one missing value is 0.

Chi-Square Statistics Section
Chi-square: 1.745200
Degrees of Freedom: 1.0000
Probability level: 0.18841
Phi: 0.107864
Cramer's V: 0.107864
Kappa's t value: 1.32106
McNemar's Probability Level: 0.029049

What is the expected value for the number of students who weight more than 150 pounds and could lift more than 200 pounds?

(A) 0.47
(B) 34
(C) 38
(D) 63
(E) 81
34. A scatter plot and Least Squares Regression line are shown in the figure below. If the highlighted point (20, 21) is removed from the data set, which of the following statements is true?

(A) The slope will decrease and the y-intercept will decrease.
(B) The slope will increase and the y-intercept will decrease.
(C) The slope will decrease and the y-intercept will increase.
(D) The slope will increase and the y-intercept will increase.
(E) No conclusion can be drawn since the coordinates of the other data points is unknown.

35. A large bakery has a many different products for sale. Suppose that 70% of all customers of the bakery order donuts, 50% order cinnamon rolls, and 40% order both. If a customer is randomly selected, what is the probability that he ordered either donuts or cinnamon rolls but not both?

(A) 20%
(B) 48%
(C) 40%
(D) 60%
(E) 24%
36. A building engineer is trying to determine the wait times for a bank of elevators. He takes a random sample of waiting times for riders. The summary results and box plot of the data are shown below.

\[ \bar{x} = 1.31 \text{ minutes.} \]
\[ s = 0.85 \text{ minutes.} \]
\[ n = 21 \]

Wait Times

Which of the following conditions necessary for an inference test has not been met?
I. The sample size must be at least 30.
II. The \( \alpha \) level must be stated.
III. The distribution of the sample data must be approximately normal.

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

Assuming sample size cannot change.
37. Two lists of numbers were created. List 1 was random integers from 0 to 25. List 2 was random integers from 5 to 30. The difference of List 1 - List 2 was sampled 100 times and the differences plotted. Which of the following graphs represents the data?

(A) Sample of Graph A

(B) Sample of Graph B

(C) Sample of Graph C

Additional graphs are on next page
38. An advertiser compared the number of television commercials for a product viewed in one month with viewers’ rating of that product. The results are summarized in the table.

<table>
<thead>
<tr>
<th>Number of exposures</th>
<th>0-5</th>
<th>6-10</th>
<th>11-15</th>
<th>15-20</th>
<th>21+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>23</td>
<td>29</td>
<td>42</td>
<td>48</td>
<td>51</td>
</tr>
<tr>
<td>Good</td>
<td>25</td>
<td>33</td>
<td>44</td>
<td>53</td>
<td>49</td>
</tr>
<tr>
<td>Fair</td>
<td>31</td>
<td>29</td>
<td>25</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>Poor</td>
<td>38</td>
<td>32</td>
<td>25</td>
<td>18</td>
<td>12</td>
</tr>
</tbody>
</table>

The advertiser decided to use a \( \chi^2 \)-test to see if there was a relationship. The degrees of freedom for this test is

\[
\frac{4 \cdot 3}{12} = 12
\]

(A) 3
(B) 4
(C) 12
(D) 18
(E) 19

39. A marine biologist wants to determine the best treatment to eliminate a water-borne parasite from salmon which causes a weight loss. There are 40 Pacific salmon and 60 Atlantic salmon. The biologist wants to check on three different treatments against a control group. A block design is chosen with 10 Pacific salmon and 15 Atlantic salmon in each treatment. Which of the following is a valid reason to use a block design for this experiment?

(A) The conclusions can apply equally to both species.
(B) Atlantic salmon are smaller than Pacific salmon.
(C) it is a requirement of a double blind experiment.
(D) Each species may respond differently to the treatments.
(E) There might be different parasites in each ocean.

40. 23 of 65 people randomly selected at a baseball game report owning a cell phone. 18 of 52 people randomly selected at a rock concert occurring at the same time across town reported owning a cell phone. A researcher wants to test the claim that the proportion of cell phone ownership is not the same between the two groups. Assuming the two groups are independent, what is the P-value of this test?

\[
\begin{align*}
 z &= \frac{23 - 18}{\sqrt{\frac{23 \times 18}{65 + 52}}} = \frac{41}{117} \\
 &= 0.2633 \\
\end{align*}
\]

\[
\begin{align*}
 p(Z > 0.089) &= 0.931 \\
\end{align*}
\]

(A) 0.045
(B) 0.090
(C) 0.465
(D) 0.910
(E) 0.931