1. Do not open this exam until you are told to do so.

2. There are 20 multiple-choice questions on this exam, each worth the same amount. Please mark your answers clearly on a GRAY Scantron sheet. Multiple marks will be counted wrong.

3. You must mark your Scantron form with
   (a) Your NAME and UIN.
   (b) Your Form letter which is above.

4. You will have only 60 minutes to finish this exam.

5. You may use the following:
   (a) One $8\frac{1}{2} \times 11$ formula sheet (both sides) of your own making.
   (b) A copy of the Percent Curve handout.
   (c) A copy of the $Z$ tables.
   (d) A copy of the recap of Week 1 & 2.
   (e) A stand-alone calculator, i.e., one that cannot communicate with the internet or anything outside itself.

6. You must put all possessions besides, the materials listed and your scantron, pencil(s) and eraser, along the walls or at the front of the room out of everyone else’s way. This includes cell phones, which must be turned off.

7. If you have questions, please write out what you are thinking on this test so that we can discuss it after your results are returned to you.

8. If you are caught cheating or helping someone to cheat on this exam, you both will receive a grade of zero on the exam. You must work alone.

9. When you are finished please make sure you have marked your Section and Form and have an answer for every question, then turn in your scantron and show your ID.

10. Good luck!
1. If $X \sim N(30, 6^2)$, how likely are you to get an observation above 32?

A. 0.3333  
B. 0.6293  
C. -0.43  
D. 0.3707  
E. 0.43

2. If a dataset has outliers

A. the mean is a better measure of center than the median.
B. the IQR is a better measure of spread than the standard deviation.
C. $z$-scores won’t have a mean of 0 and a standard deviation of 1.
D. All of the above are correct statements.
E. Only two of the above are correct statements.

3. Which of the following is true based on the graph above which is from our class survey? The bar on the left in each category is females; the one on the right is males.

A. Females like animals (pets) more than males.
B. Males like dogs better than cats.
C. Females are more likely to have a pet than males.
D. All of the above are true.
E. None of the above are true.

4. If $Min = 1, Q_1 = 9, M = \bar{x} = 15, Q_3 = 23, Max = 31$, which of the following best describes the shape of this distribution?

A. normal  
B. uniform  
C. skewed left  
D. skewed right  
E. skewed left with an outlier

5. If we multiplied the data in the previous problem by 10, what would the new summaries be?

A. $Min = 10, Q_1 = 90, M = \bar{x} = 150, Q_3 = 230, Max = 310$  
B. The mean would be 10 times larger.  
C. The standard deviation would be 10 times larger.  
D. All of the above are true.  
E. Only A and B are true.

6. Suppose you get a $z$-score = 1.2 on this exam. Question #10 is really hard, so I decide to give everyone credit, i.e., I give 5 points to everyone who missed it. What SHOULD you think about this if you got the problem right originally?

A. It doesn’t matter since your $z$-score would stay the same.
B. You get gypped (lose) since your $z$-score would decrease.
C. You benefit since your $z$-score would increase.
D. Your $z$-score would increase, but so would everyone else’s.
E. Your $z$-score would decrease, but so would everyone else’s.

7. If the mean for a set of data is 25 and the standard deviation is 8, then

A. The $z$-score for 20 is 0.625.  
B. About 95% of the observations will between 9 and 41.  
C. 57 would be an outlier.  
D. All of the above are true.  
E. None of the above are true.

8. The middle 90% of the $Z \sim N(0, 1^2)$ distribution is between

A. ±1.645  
B. ±1.28  
C. ±0.8159  
D. ±0.13  
E. ±0.5199

9. Which of the following is true about categorical data?

A. Because it is not numbers, you can’t calculate any statistics from it.  
B. You can’t describe the shape of the distribution.  
C. You can always use either a pie chart or a bar chart to plot the data, but never a boxplot or histogram.  
D. All of the above are true.  
E. None of the above are true.

10. When can we correctly apply the Empirical Rule?

A. When the normal quantile plot is approximately a straight line.  
B. When the data is bell-shaped.  
C. When the data is symmetric.  
D. All of the above are true.  
E. Exactly two of the above are true.

11. Suppose I randomly pick 20 students from a statistics class and ask them how many roommates each of them have. Which of the following must be true?

A. The mean number, $\bar{x}$, must be an integer since the data is all integers (discrete).  
B. The median number, $\tilde{x}$, must be an integer since the data is all integers.  
C. The mode must be an integer since the data is all integers.  
D. The standard deviation cannot be an integer.  
E. Two of the above are true.
<table>
<thead>
<tr>
<th>taking</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>0.0087</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.0130</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.0043</td>
</tr>
<tr>
<td>12</td>
<td>33</td>
<td>0.1429</td>
</tr>
<tr>
<td>13</td>
<td>28</td>
<td>0.1212</td>
</tr>
<tr>
<td>14</td>
<td>18</td>
<td>0.0779</td>
</tr>
<tr>
<td>15</td>
<td>77</td>
<td>0.3333</td>
</tr>
<tr>
<td>16</td>
<td>50</td>
<td>0.2165</td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>0.0433</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>0.0173</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>0.0173</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>0.0043</td>
</tr>
</tbody>
</table>

12. Which of the following is the correct *5Number Summary* for the data in the table above?

A. $Min = 3, Q_1 = 11, M = 14.5, Q_3 = 17.5, Max = 20$
B. $Min = 3, Q_1 = 7.25, M = 11.5, Q_3 = 15.75, Max = 20$
C. $Min = 3, Q_1 = 13, M = 14, Q_3 = 15, Max = 20$
D. $Min = 3, Q_1 = 13, M = 15, Q_3 = 16, Max = 20$
E. It cannot be determined from this table.

13. Jimmy made a 75 on his test and Janie made an 80 on hers, but the mean for Jimmy’s was a 72 with a standard deviation of 6 and Janie’s had a mean of 78 with a standard deviation of 4. Who did the best relative to their class?

A. Janie, since 80 is higher than 75.
B. Jimmy, since he’s 3 points higher than the average and Janie is only 2 points higher.
C. Janie since her standard deviation is smaller.
D. They did relatively the same.
E. We can’t compare since we don’t know how many students took the test.

14. Which of the following is true about the data in the graph above?

A. The minimum height for females is about 55 inches.
B. Three-fourths (75%) of the males are taller than three-fourths of the females.
C. The average (mean) height of females is about 65 inches.
D. All of the above are true.
E. Only two of the above are true.

15. A nine-year-old genius appeared on *The David Letterman Show*. Dave asked him what his IQ was, and he replied that it was “off the bell curve”. For his IQ to be an outlier about how large must it be? NOTE: Assume IQ’s are normally distributed with mean = 100 and standard deviation = 15.

A. over 130
B. over 115
C. over 100
D. over 150
E. normal distributions don’t have outliers

16. The sample proportion of red M&M’s, in bags with a total of 50, is $p_{red} \sim N(0.2, (0.057^2))$. How likely are you to get only 2 reds (so $p_{red} = 2/50 = 0.04$)? What is $P(p_{red} \leq 0.04)$?

A. −2.81
B. 0.0025
C. 0.025
D. 0.0218
E. −2.017
17. The graph above is for the head circumference from the class survey. Which of the following is true?

A. More than half of the class has a head circumference between 20 and 25 inches.
B. Some students entered their measurements wrong.
C. If we deleted the observations below 10, the average and standard deviation of the data would increase.
D. All of the above are true.
E. Only two of the above are true.

18. Suppose $X \sim N(30,12^2)$. Which of the following are true?

A. The IQR is 1.35.
B. $Q_1$ is about 8.
C. The minimum value shouldn’t be less than $-18$.
D. All of the above are true.
E. Two of the above are true.

19. Suppose $X \sim N(5,6^2)$. What is $P(4.1 < X < 7.1)$?

A. 0.1964
B. 0.2000
C. 0.8036
D. 0.5596
E. 0.6368

20. Outliers affect our statistics so

A. we should always plot the data first.
B. we should use the median instead of the mean when they are present.
C. we should delete them since they don’t fit the rest of the data.
D. All of the above are correct.
E. Only two of the above are correct.

1D,2B,3C,4B,5D,6B,7E,8A,9B,10E,11C,