1. Don’t even open this until you are told to do so.

2. There are 20 multiple-choice questions on this exam, each worth 5 points. There is partial credit. Please mark your answers clearly. Multiple marks will be counted wrong.

3. You will have 50 minutes to finish this exam.

4. If you have questions, please write out what you are thinking on the back of the page so that we can discuss it after I return it to you.

5. 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.

6. When you are finished please make sure you have marked your CORRECT section (Tuesday 11:10 is 507, 12:45 is 508, 2:20 is 509, and 3:55 is 510) and FORM (A, B, C or D) and 20 answers, then turn in JUST your scantron.

7. Good luck!
1. Outliers
   A. affect z-scores.
   B. affect all measures of spread.
   C. don’t affect measures of location.
   D. All of the above are true.
   E. Only two of the above are true.

2. Which of the following is true for the histogram above?
   Each bin contains only the value to the left and there are no half sizes.
   A. The median 8 since it’s the tallest bin.
   B. The mean is larger than the median since it’s skewed to the left.
   C. The 5 Number Summary is 4, 7, 9, 11, 15.
   D. The 5 Number Summary is 4, 7, 8, 10, 15.
   E. More than one of the above are true.

3. Which of the following is a categorical variable?
   A. Social Security Number
   B. Classification: senior, junior, etc.
   C. days of the week
   D. All of the above are categorical variables.
   E. Exactly two of A, B, and C are categorical variables.

4. Sample size, \( n \), matters because
   A. it affects the shape of the distribution of the sample data.
   B. it affects the center of the distribution of the sample data.
   C. it affects the spread of the distribution of the sample data.
   D. Two of the above.
   E. None of the above.

5. This is the data from our survey showing how many hours of sleep you usually get per night vs. gender. Does it appear that there is a relationship between the two?
   A. Yes, there are more females in every category.
   B. Yes, the proportions are about the same for males and females.
   C. No, the proportions are about the same for males and females.
   D. No, the numbers are different for males and females.
   E. There is not enough evidence to tell.

6. Referring to the previous table, how likely is a female to sleep between 8 and 10 hours per night?
   A. \( \frac{104}{170} \)
   B. \( \frac{51}{170} \)
   C. \( \frac{32}{170} \)
   D. \( \frac{32}{104} \)
   E. \( \frac{32}{51} \)

7. Which of the following is the best description of the probability \( \frac{35}{90} \)?
   A. how likely a male is to sleep 5 to 7 hours per night
   B. how likely you get a male if they sleep 5 to 7 hours per night
   C. how likely you sleep 5 to 7 hours per night if you are male
   D. of the males, how likely you get someone who sleeps 5 to 7 hours per night
   E. how likely you get someone who sleeps 5 to 7 hours per night and they are male

8. Still referring to the table, are being a female and sleeping 5 to 7 hours per night independent events?
   A. Yes, \( \frac{55}{170} = (\frac{104}{170}) \ast (\frac{90}{170}) \) (within 0.000)
   B. Yes, \( \frac{55}{104} = (\frac{90}{170}) \) (within 0.000)
   C. Yes, \( \frac{55}{90} = (\frac{104}{170}) \) (within 0.000)
   D. All of the above prove they are independent.
   E. Since all of the above are true (except D), they are dependent.
9. Jack is 6’2” (74 in.) and Jill is 5’8” (68 in.) tall. Using the summaries above, who is relatively taller.

A. Jack
B. Jill
C. We can’t compare males to females since males on average are taller.
D. Jack, of course, 74 > 68
E. They are relatively the same height.

10. Corn variety #1 yielded 140 bushels per acre last year at a research farm. This year, corn variety #2, planted in the same location, yielded only 110 bushels per acre. Unfortunately, we don’t know whether the difference is due to the superiority of variety #1 or to the effect of this year’s drought. This is an example of

A. bias due to voluntary response.
B. random sampling error.
C. confounding.
D. the placebo effect.
E. nonsampling error.

11. Which of the following is most likely the approximate mean of this boxplot?

A. 12
B. 13
C. 16
D. 17
E. 18

12. BLANK A in a sampling method means that the sample results will systematically misrepresent the population in the same way when we take repeated samples. For example, if we contact only people listed in telephone directories, the sample suffers from BLANK B. If some people chosen for the sample refuse to participate, the sample suffers from BLANK C. Both BLANK B and BLANK C are common sources of BLANK A. BLANK A should read

A. bias
B. random sampling error
C. high variability
D. undercoverage
E. nonresponse

13. BLANK B in the previous problem should read

A. bias
B. random sampling error
C. high variability
D. undercoverage
E. nonresponse

14. If we wanted to show the relationship between drinking (yes or no) and income (in dollars), which of the following should we use?

A. a two-way table with drinking as the column variable and income as the row variable
B. side-by-side boxplots of income for drinkers and non
C. a scatterplot of income vs. drinks per week
D. two pie charts since drinking is a categorical variable
E. Any of the above would work.

15. In a tire-treadwear study one brand of tires was tested at different temperatures. What is the explanatory variable??

A. brand of tire
B. temperature
C. both brand and temperature
D. tire-treadwear
E. both tire-treadwear and temperature

16. What is the shape of the distribution above?

A. skewed left
B. skewed left with outliers
C. skewed right
D. bell shaped
E. flat (uniform)
17. Who spends more? Based on our survey, the boxplots above compare how much males vs. females spend on haircuts and lunch. (I deleted the $150 haircut and the guys who spent more than $100 on lunch. Wouldn’t you like them to take you to lunch??) Which of the following is/are true statements about these comparisons?

A. Seventy five percent of the females spend more on their hair than on lunch.
B. Males would rather spend their money on food than haircuts.
C. Males are much more consistent (less variable) in what they spend on haircuts than females.
D. All of the above are true.
E. Only two of the above are true.

18. If I eliminated everyone who spent more than $50 on either a haircut or lunch, what would the new boxplots look like?

A. Male haircuts would not change.
B. Female lunch would still have outlier(s).
C. Male lunch would be more symmetric.
D. All of the above are true.
E. Only two of the above are true.

19. If I converted dollars to euros, what would happen?

A. The y-axis would change, but the boxplots would look the same.
B. The 5 Number summaries would change, but the range and IQR’s would stay the same.
C. All the statistics (locations and spreads) would change.
D. Two of the above are true.
E. None of the above are totally correct.

20. Finally, what measures of center and spread should be used to describe the distributions represented by the boxplots?

A. the median and IQR
B. the median and IQR for lunch and the mean and standard deviation for haircut since they are more symmetric
C. the mean and IQR for all