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

2. Be sure to mark your section number (508, 509 or 510) and your test form (A, B, C or D) on the scantron!

3. Sign your name where indicated on your scantron and write your Tuesday section number and computer number beside it. Also, you must place your scantron in the correct section stack (for next Tuesday).

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

5. You will have 60 minutes to finish this exam.

6. 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.

7. This exam is worth 100 points, and will constitute 25% of your final grade.

8. Good luck!

Grades for this exam will be available by 5PM, THURSDAY. The Statistics office WON’T have them, however, so I would like my test score and final grade posted by the last four digits of my Social Security Number. This means it will appear on the Statistics Lab bulletin board no later than Thursday at 5PM, and I won’t call the office asking about it!

Printed Name:

Signature:

SSN:
1. A statistic is an unbiased estimator if
   A. the mean of all possible values of the statistic from all possible samples of size \( n \) from the population is \( \mu \).
   B. it has the smallest variance of all the estimators.
   C. the sampling distribution of the statistic is normal.
   D. the mean of the sampling distribution of the statistic is the parameter of interest.
   E. the original population from which the sample is drawn is normal.

2. The phrase \((1 - \alpha)100\%\) confidence means
   A. the probability of the confidence interval NOT capturing the value of the population parameter is \( \alpha \).
   B. for a large enough sample size \( n \), we can be \((1 - \alpha)100\%\) confident of capturing the true value of \( \mu \).
   C. the null hypothesis will be true if the hypothesized mean, \( \mu_0 \), is in the confidence interval.
   D. the estimator \( \bar{X} \) is unbiased for \((1 - \alpha)100\%\) of the possible values of \( \bar{X} \) based on all possible sample of size \( n \), provided \( n \) is large enough.
   E. if we could take all possible samples of size \( n \) and calculate the \((1 - \alpha)100\%\) confidence interval for each sample, then exactly \((1 - \alpha)100\%\) of these intervals would contain the value of the population parameter.

3. In testing hypotheses, the difference between a Type I error and a Type II error is
   A. the Type I is always more significant.
   B. you can only make a Type I error if the null hypothesis is true, but you can always make a Type II error.
   C. they are complements, i.e., \( \beta = 1 - \alpha \).
   D. All of the above are true.
   E. None of the above are true.

4. The article "What Kinds of People Do Not Use Seat Belts?" (Amer. J. of Public Health(1977)) reported on a survey with the objective of studying characteristics of drivers and seat belt usage. Let \( \pi \) denote the proportion of all drivers of cars with seat belts who use them. Suppose that at the outset of the study, the investigators wished to estimate \( \pi \) to within 2\% with 90\% confidence. What is the required sample size?
   A. 83
   B. 9604
   C. 1692
   D. 2401
   E. 6765

5. What is the power of the test for the test in the graph above?
   A. 0.05
   B. 0.19
   C. 0.81
   D. 0.95
   E. 25

6. What elements of a one-sample statistical test may be affected by changing the sign of the alternative hypothesis test? (Assume the data, sample size, and \( \alpha \) are the same)
   A. the value of the test statistic
   B. the p-value
   C. both A. and B.
   D. the sign of the alternative is always =, and therefore, cannot change
   E. None of the above are correct.

7. Suppose the true mean and standard deviation of Duracell Alkaline AA battery lifetime are 4.1 hr and 1.8 hr, respectively. Those of Eveready batteries are 4.5 hr and 2.0 hr, respectively. If \( \bar{X}_D \) is the sample mean of 100 Duracells and \( \bar{X}_E \) is the sample mean of 100 Evereadys, what is the approximate sampling distribution of \( \bar{X}_D - \bar{X}_E \)?
   A. \( N(-0.4, 0.269^2) \)
   B. \( N(-0.4, 0.2^2) \)
   C. \( N(-0.4, -0.2^2) \)
   D. \( N(-0.4, 0.0724^2) \)
   E. \( N(-0.4, 0.02^2) \)
8. A fast food franchiser is considering building a restaurant in a certain location. Based on financial analyses, a site is acceptable only if the number of pedestrians passing the location averages more than 100 per hour. This number is normally distributed, and we can assume that the population standard deviation is known to be 12. The number of pedestrians observed for each of forty hours was recorded. Which type of statistical test should we employ to decide if the site is acceptable?

A. Case 1 since the population standard deviation is known and we are testing one mean.
B. Case 3 since the sample size is large (> 30) and we are testing one mean.
C. Case 8 since we can assume that the variances are both $12^2$.
D. Case 9 since we are not told that the variances are equal.
E. Case 10 since this test reduces the variability.

9. What conclusion can be made from the graph above?

A. We fail to reject at the 5% level and conclude that the true mean is not 0.
B. We fail to reject at the 5% level and conclude that there is not enough evidence to say the true mean is not 0.
C. We fail to reject at the 5% level and conclude that the two means are not equal.
D. We fail to reject at the 5% level and conclude that there is not enough evidence to the two means are not equal.
E. None of the above are correct.

10. In the previous problem, we assumed the variances were equal. The graph above tests this assumption for the same data. Which of the following is/are true?

A. There is no way the the two variances are equal since one standard deviation is 5.7 and the other is only 4.
B. The p-value is less than 5%, therefore, we the assumption of equal variances is valid.
C. The p-value is less than 5%, therefore, we the assumption of equal variances is invalid.
D. The p-value is less than 5%, therefore, the difference of the variances is only 1.
E. The p-value is less than 5%, therefore, the difference of the variances is not 1.

11. Again referring to problem 9 above about testing the means, which of the following statements would be true?

A. 0 would be in a 90% confidence interval for the differences of the two means, but 0 would not be in a 95 nor 99%.
B. 0 would not be in a 90% confidence interval for the differences of the two means, but 0 would be in a 95 and a 99%.
C. $\mu_1$ and $\mu_2$ would be in a 90% confidence interval for the differences of the two means, but they would not be in a 95 nor 99%.
D. $\mu_1$ and $\mu_2$ would not be in a 90% confidence interval for the differences of the two means, but they would be in a 95 and 99%.
E. Exactly two of the above are correct.
12. Which of the following is NOT true?

A. The standard test procedures allow the user to control \( \alpha \), but they provide no direct control over \( \beta \).
B. Choosing a small value for \( \alpha \) implies that the user wants to employ a procedure for which the risk of a Type I error is quite small.
C. By collecting more data, both \( \alpha \) and \( \beta \) can be reduced simultaneously.
D. All of the above are true.
E. None of the above are true.

13. Suppose you are testing the hypothesis \( H_0 : \mu = 3 \) vs. \( H_A : \mu \neq 3 \), at \( \alpha = 0.10 \) and the only information available from a sample of size \( n = 10 \) from a normally distributed population is the 90% confidence interval (2.73, 3.17). Then

A. since 3 is in the confidence interval, and we are 90% certain the \( \mu \) is in the confidence interval, \( \mu \) could not be 3, so we reject \( H_0 \).
B. since 3 is in the confidence interval, and we are 90% certain the \( \mu \) is in the confidence interval, \( \mu \) could not be 3, so we fail to reject \( H_0 \).
C. since 3 is in the confidence interval, and we are 90% certain the \( \mu \) is in the confidence interval, \( \mu \) could be 3, so we reject \( H_0 \).
D. since 3 is in the confidence interval, and we are 90% certain the \( \mu \) is in the confidence interval, \( \mu \) could be 3, so we fail to reject \( H_0 \).
E. we could not determine the correct conclusion since the sample size is too small.

14. What would be the consequence of a Type II error when testing the hypotheses above?

A. You claim the true mean is 3 when it really is 3.
B. You claim the true mean is 3 when it really is NOT 3.
C. You claim the true mean is NOT 3 when it really is 3.
D. You claim the true mean is NOT 3 when it really is NOT 3.
E. You claim you hate math so you can’t do statistics.

15. Why do we use the \( t \) distribution rather than a \( z \) when \( \sigma \) is unknown?

A. It is easier to get a rejection with the \( t \) (i.e., the confidence intervals are narrower).
B. The degrees of freedom for the \( t \) make it more accurate than the \( z \).
C. The \( t \) distribution allows for the variability of \( s \) as an estimate for \( \sigma \).
D. Exactly two of the above are correct.
E. None of the above are correct.

16. Which of the following statements is the correct interpretation of the power of the test?

A. The power of the test is when you reject a false null hypothesis.
B. If the null hypothesis is false, the power of the test is the proportion of all test statistics from all possible samples of size \( n \) that will result in a test statistic that rejects the null.
C. If the null hypothesis is false, the power of the test is all test statistics from all possible samples of size \( n \) that will result in a test statistic that rejects the null.
D. If the null hypothesis is true, the power of the test is the proportion of all test statistics from all possible samples of size \( n \) that will result in a test statistic that fails to reject the null.
E. If the null hypothesis is true, the power of the test is all test statistics from all possible samples of size \( n \) that will result in a test statistic that fails to reject the null.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>38</td>
<td>25</td>
<td>.8111071</td>
<td>23.63159 26.36841</td>
</tr>
<tr>
<td>95%</td>
<td>38</td>
<td>25</td>
<td>.8111071</td>
<td>23.35654 26.64346</td>
</tr>
<tr>
<td>99%</td>
<td>38</td>
<td>25</td>
<td>.8111071</td>
<td>22.79751 27.20249</td>
</tr>
</tbody>
</table>

17. What is the appropriate range for the p-value for testing \( H_0 : \mu = 27 \) vs. \( H_A : \mu \neq 27 \)?

A. p-value > 0.10
B. 0.10 > p-value > 0.05
C. 0.05 > p-value > 0.01
D. 0.01 > p-value
E. The test statistic is needed in order to determine the p-value.
18. Suppose that a random sample of 50 bottles of a particular brand of cough medicine is selected and the alcohol content of each bottle is determined. Let \( \mu \) denote the true average alcohol content for the population of all bottles of the brand under study. If the sample of 50 results in a 95% confidence interval for \( \mu \) of \((7.8, 9.4)\), which of the following statements is correct?

A. There is a 95% chance that the true mean alcohol content, \( \bar{x} \), of this brand falls between 7.8 and 9.4.
B. The estimate, \( \bar{x} \), is within 95% of the true mean, \( \mu \).
C. If the process of selecting a sample of size 50 and then computing the corresponding 95% confidence interval is repeated many times, approximately 95% of the resulting intervals will include \( \mu \).
D. If the process of selecting a sample of size 50 and then computing the corresponding 95% confidence interval is repeated many times, approximately 95% of the resulting intervals will fall within \((7.8, 9.4)\).
E. Both B. and C. are correct.

19. The inspection division is interested in whether the actual amount of soft drink that is placed in a 2-liter bottle at the local bottling plant of a large nationally known soft drink company is really 2 liters. A random sample of 100 bottles produced an average fill of only 1.99 liters. The p-value of the test was less than 5%, so the inspection division concluded that the bottles actually contained less than 2 liters. Which of the following is true?

A. The inspection division made a Type I error since they rejected a true \( H_0 \).
B. The inspection division made a Type II error since the bottling company failed the test.
C. The large sample size caused a difference of 10 ml (0.01 l) to be significant when it really isn’t practically significant to the consumer, i.e., we couldn’t tell the difference even by looking at the bottles.
D. Without knowing the exact p-value, we cannot determine if an error was made.
E. Coke would never be dishonest, so this must have been Pepsi.

20. The level of confidence is compromised when the population from which a sample is taken is not normally distributed. This implies

A. although we state a level of confidence equal to \((1 - \alpha)100\%\), the true level of confidence may not be \((1 - \alpha)100\%\), but something quite different.
B. although we say we are calculating a 95% confidence interval, we’re really only calculating a 68% confidence interval.
C. if the sample size is large enough, the distribution of the confidence intervals will be symmetric.
D. unless the sample size is large enough, we will get confidence intervals that are too wide.
E. the confidence intervals will be biased.

21. The mayors of Bryan and College Station have a bet going. Whoever’s city has the higher average family income wins $10. Since the mayor of College Station is doing the test, what significance level would he prefer to use? Remember \( H_0 \) is always that the two are equal.

A. 0.10
B. 0.05
C. 0.01
D. It depends on which city is sample 1 and which is sample 2.
E. It doesn’t matter since a hypothesis can’t conclude one is higher than the other.