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

2. There are 18 multiple-choice questions on this exam, each worth 5 points, and 5 True/False, each worth 2 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. You think the true proportion of your population is 45%. If you are correct, which of the following would have the most power?
   A. $H_0 : p \leq 0.3$ vs. $H_A : p > 0.3$ with a sample size, $n = 35$ and $\alpha = 0.01$.
   B. $H_0 : p \leq 0.3$ vs. $H_A : p > 0.3$ with a sample size, $n = 35$ and $\alpha = 0.10$.
   C. $H_0 : p \leq 0.4$ vs. $H_A : p > 0.4$ with a sample size, $n = 35$ and $\alpha = 0.10$.
   D. $H_0 : p \leq 0.4$ vs. $H_A : p > 0.4$ with a sample size, $n = 50$ and $\alpha = 0.10$.
   E. $H_0 : p \leq 0.3$ vs. $H_A : p > 0.3$ with a sample size, $n = 50$ and $\alpha = 0.10$.

2. Which of the following would be a Type I error in the previous situation?
   A. claiming your true proportion is greater than 30% when it’s really 45%.
   B. claiming your true proportion is 45% when it’s really 30%.
   C. claiming your true proportion is greater than 40% when it’s really 30%.
   D. claiming your true proportion is greater than 40% when it’s really 45%.
   E. claiming your true proportion is less than 45% when it’s really 30%.

A survey revealed the following information about HOMETOWN size and GENDER. Use the table below to answer the next 5 questions.

<table>
<thead>
<tr>
<th>GENDER</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>male</td>
<td>female</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>HOMETOWN metro</td>
<td>37</td>
</tr>
<tr>
<td>large</td>
<td>12</td>
</tr>
<tr>
<td>city</td>
<td>24</td>
</tr>
<tr>
<td>town or rural</td>
<td>36</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Total</td>
<td>109</td>
</tr>
</tbody>
</table>

Chi-Square Test
Value df Sig.(2-sided)
Chi-Square 0.526 3 0.913
0 cells have expected count less than 5.
The minimum expected count is 13.18.

3. What is the Expected Count for female city dwellers?
   A. 38
   B. 38*62*164/273
   C. 62*164/273
   D. 164/273*62/273
   E. 164*62

4. What is the correct conclusion for the $\chi^2$ test in the previous question?
   A. There is sufficient evidence to conclude that there is a relationship between GENDER and HOMETOWN.
   B. There is insufficient evidence to conclude that there is a relationship between GENDER and HOMETOWN.
   C. There is insufficient evidence to conclude that there is a relationship between males and females by HOMETOWN.
   D. There is insufficient evidence to conclude that GENDER and HOMETOWN are independent.
   E. There is sufficient evidence to conclude that GENDER and HOMETOWN are independent.

5. Which of the following is an example of a Type I error for the previous $\chi^2$ test above?
   A. We claim that GENDER and HOMETOWN are related when they are actually independent.
   B. We claim that males and females are related by their HOMETOWN.
   C. We fail to prove that GENDER and HOMETOWN are independent even though they actually are.
   D. We fail to prove that GENDER and HOMETOWN are related even though they actually are.
   E. We claim that females are more likely to live in a metroplex when actually all of the types of HOMETOWN are about equally likely.

6. How likely is a female townie (living in town or rural area)
   A. $48/164$
   B. $48/273$
   C. $48/84$
   D. $84/273$
   E. Both C and D

7. How likely is a person from a large city?
   A. (12+21)/33
   B. (12+21)/109
   C. 12/109 + 21/164
   D. 33/273
   E. Both C and D
8. Which of the following is TRUE?

A. A 95% confidence interval is ALWAYS narrower than a 99% when using the same sample.
B. Increasing the sample size, $n$, will ALWAYS increase the width of a confidence interval.
C. 2 sample confidence intervals are more accurate than 1 sample intervals since they use more data (have a bigger combined sample size)
D. Two-sample confidence intervals are ALWAYS centered about 0.
E. All of the above are false statements.

9. Which of the following best defines the $p$-value for the $\chi^2$ statistic when assessing a statistically significant relationship?

A. It is the likelihood that the two categories are independent.
B. It is the likelihood that the two categories are dependent.
C. It is the likelihood of seeing this strong of a relationship when there really isn’t any relationship.
D. It is the likelihood of seeing this strong of a relationship when there really IS a relationship.
E. Two of the above are correct.

10. Suppose the US Department of Education is interested in knowing if the proportion of high school seniors planning on attending college is different than 0.6. They conduct a hypothesis test with $\alpha = 0.05$ and get a $p$-value of 0.039. What interpretation should they make?

A. They reject $H_0$, and believe that proportion of high school seniors planning to attend college is less than 0.6.
B. They do not reject $H_0$, and believe that the proportion of high school seniors planning to attend college is 0.60.
C. They reject $H_0$, and believe that the proportion of high school seniors planning to attend college is different than 0.60.
D. They reject $H_0$, and believe that the proportion of high school seniors planning to attend college is greater than 0.60.
E. They do not reject $H_0$, and believe that the proportion of high school seniors planning to attend college is different than 0.60.

11. What is a $p$-value anyway?

A. A $p$-value is probability from the hypothesized curve.
B. A $p$-value is used to help us decide whether the hypothesized value would fall in a particular confidence interval or not.
C. A $p$-value is dependent on the sample data, the sign of the alternative hypothesis and the hypothesized value.
D. All of the above are true statements about a $p$-value.
E. Only two of the above are true statements about a $p$-value.

12. The American Medical Association wants to investigate the proportion of Americans who are allergic to aspirin. They get a 95% confidence interval for the true proportion of (0.087, 0.145). Which of the following are true?

A. At $\alpha = 0.1$, they would reject $H_0 : p = 0.05$.
B. They are 95% confident that the true proportion of Americans with an aspirin allergy is between 0.087 and 0.145.
C. At $\alpha = 0.05$, they would not reject $H_0 : p = 0.09$.
D. B and C are true, we can’t be sure about A.
E. A, B, and C are all true.

13. When deciding whether or not to pick up a television series, network executives often show the pilot episode to test audiences. Suppose ABC wants to know if women report a grade of “Favorable” for the pilot of a certain show at a higher rate than men. What hypothesis should they test?

A. $H_0 : p_W = p_M$ vs. $H_A : p_W \neq p_M$
B. $H_0 : \mu_W \leq \mu_M$ vs. $H_A : \mu_W > \mu_M$
C. $H_0 : p_W \geq p_M$ vs. $H_A : p_W < p_M$
D. $H_0 : p_W \leq p_M$ vs. $H_A : p_W > p_M$
E. $H_0 : \mu_W \geq \mu_M$ vs. $H_A : \mu_W < \mu_M$

14. Suppose my alternative hypothesis is $H_A : p_1 < p_2$ and I get a test statistic of $-1.82$. What can I conclude?

A. I would reject the null at 10% significance but not the 5% significance level.
B. At the 1% significance level, I believe that the group 1’s proportion is less than group 2’s.
C. I would reject the null hypothesis at the 5% significance level but not the 10% significance level.
D. I would not reject for 10%, 5%, or 1% significance levels.
E. I would reject the null for 5% and 10% significance, but not 1%.
15. Which of the following affect the width of a confidence interval for the true proportion, $p$?

A. the sample proportion, $\hat{p}$
B. the sample size, $n$
C. the confidence level, $1 - \alpha$
D. all of the above
E. only two of the above

16. Suppose that for some population the true proportion, $p$, is 0.75. I conduct a hypothesis test where the alternative hypothesis is that $p < 0.8$. I get a $p$-value of 0.071. Which of the following true?

A. At the 10% significance level, I make a correct decision. At 5%, I commit a Type I error.
B. At the 10% significance level, I commit a Type I error. At 5%, I make a correct decision.
C. I commit a Type II error at both 10% and 5% significance.
D. For 10% significance, I make a correct decision. At 5%, I commit a Type II error.
E. There’s not a hypothesized value to determine the $p$-value.

17. If we test whether $H_0 : p = 0.5$ or not and our sample proportion is $p = 0.4$ from a sample of $n = 50$, what is the value of the test statistic and the $p$-value?

A. 1.44 and 0.0749
B. 1.41 and 0.0793
C. −1.41 and 0.1586
D. −1.44 and 0.1498
E. 1.44 and 0.1498

18. Using the three confidence intervals below, what is the correct range of the $p$-value when testing $H_0 : p_1 = p_2$ vs. $H_A : p_1 \neq p_2$?

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>(-0.3334, -0.0066)</td>
</tr>
<tr>
<td>95%</td>
<td>(-0.3647, 0.0247)</td>
</tr>
<tr>
<td>99%</td>
<td>(-0.4259, 0.0859)</td>
</tr>
</tbody>
</table>

A. $p$-value > 0.10
B. 0.10 > $p$-value > 0.05
C. 0.05 > $p$-value > 0.01
D. $p$-value < 0.01
E. There’s not a hypothesized value to determine the $p$-value.

19. If there is no apparent association between the row and column variables in a contingency table, then the observed counts must equal the expected counts.

A. True
B. False

20. A 2×2 contingency table is the equivalent of a 2 sample $z$ test for proportions with the greater than alternative since the $p$-value is always $P(\chi^2 > X)$, where $X$ is the test statistic value.

A. True
B. False

21. We test proportions instead of counts because we can control $\sigma(\hat{p})$ by the sample size.

A. True
B. False

22. The marginal distribution of the column (or row) variable in a contingency table is the distribution of the total row (or column).

A. True
B. False

23. 2 sample tests are more powerful than 1 sample tests.

A. True
B. False

1E,2C,3C,4B,5A,6B,7D,8A,9C,10C,11D,12E,