STAT303 Sec 508-510
Spring 2010
Exam #3
Form A

Instructor: Julie Hagen Carroll

Name:______________________________

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 60 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 12:45 is 508, 2:20 is 509, and 3:55 is 510) and FORM and 20 answers, then turn in JUST your scantron.
7. Good luck!

1. If we sample from $p_{50} \sim N(0.25, 0.0612^2)$, how likely are we to get a sample proportion of 30% or more?

   A. 0.2061
   B. 0.6179
   C. 0.7939
   D. 0.8170
   E. 0, the z-score is more than 5

```
H0 : pi1 - pi2 = 0
HA : pi1 - pi2 not equal 0
Difference Sample Diff. Std. Err.  Z-Stat P-value
p1 - p2    -0.22807017  0.1322165  -1.72  0.0845
```

2. The test above was to see if men (`pi2`) where *more likely* to help someone carry their bags from the mall than women (`pi1`). Which of the following is/are true?

   A. This should have been a 1-sample test to see if `pi2` was greater than 0.5, so we can't answer the question.
   B. The p-value is 0.0845, so we can conclude at the 10% level that proportion of men who would help is greater than the proportion of women.
   C. The p-value is 0.04225, so we can conclude at the 5% level that proportion of men who would help is greater than the proportion of women.
   D. The p-value is 0.95775, so we cannot conclude that proportion of men who would help is greater than the proportion of women.
   E. At the 5% level we would say the proportions are the same.

3. What is $P(p1_{50} - p2_{50} < 0.20)$ if $p1_{50} - p2_{50} \sim N(0.25, 0.025^2)$?

   A. 1
   B. 0
   C. We can't have a negative proportion, so this is impossible.
   D. 0.0228
   E. 0.9772

4. What does it mean for the χ^2 test statistic to equal 0 in the test for independence of two categorical variables?

   A. It means we would reject the null and conclude the two variables are related.
   B. It means we would reject the null and conclude the two variables are independent.
   C. It means the two variables are completely (exactly) dependent.
   D. It means the two variables are completely (exactly) independent.
   E. It means that the assumptions are not valid since the expected counts are 0.

5. Historically, there are more boys born in the United States than girls, 51.5% of newborns are in fact boys. Does this mean that there is a significantly higher proportion of newborn boys than girls? What should we test?

   A. $H_0 : \pi_b = \pi_g$ vs. $H_A : \pi_b \neq \pi_g$
   B. $H_0 : \pi_b \leq \pi_g$ vs. $H_A : \pi_b > \pi_g$
   C. $H_0 : \pi_b = \pi_g$ vs. $H_A : \pi_b \neq 0.515$
   D. $H_0 : \pi_b \leq 0.515$ vs. $H_A : \pi_b > 0.515$
   E. $H_0 : \pi_g \leq 0.515$ vs. $H_A : \pi_g > 0.515$

6. Suppose we decided to test $H_0 : \pi_b = 0.5$ vs. $H_A : \pi_b \neq 0.5$ (not even one of the choices so don't worry about your last answer). We took a sample of 45 newborns and found 48% of them were boys. What is the correct conclusion?

   A. We can't calculate the test statistic since the standard deviation isn't given.
   B. The true proportion of newborn boys is 39.36% not 50%.
   C. The p-value $= 0.7872$, so we don't have enough evidence to say that the true proportion of newborn boys is 0.5.
   D. The p-value $= 0.7872$, so we don't have enough evidence to say that the true proportion of newborn boys is not 0.48.
   E. The p-value $= 0.7872$, so we don't have enough evidence to say that the true proportion of newborn boys is not 0.5.

7. Let $\mu_{p1_{50}} = 0.5$, $\sigma_{p1_{50}} = 0.0707$ with $n_{p1_{50}} = 50$ and $\mu_{p2_{25}} = 0.75$, $\sigma_{p2_{25}} = 0.0866$ with $n_{p2_{25}} = 25$, what is the distribution of $p1_{50} - p2_{25}$ if the two populations are independent?

   A. Since $p2_{25}$ is not normal, we can only say $\mu_{p1_{50}-p2_{25}} = -0.25$.
   B. Since $p2_{25}$ is not normal, we can only say $\mu_{p1_{50}-p2_{25}} = +0.25$.
   C. $p1_{50} - p2_{25} \sim N(-0.25, 0.0159^2)$
   D. $p1_{50} - p2_{25} \sim N(-0.25, 0.1573^2)$
   E. $p1_{50} - p2_{25} \sim N(-0.25, 0.0125^2)$

8. I have a 95% confidence interval for π, the true proportion of voters in favor of Propostion 2, (0.49, 0.56). Which of the following statements are valid based on this interval?

   A. I could not reject the hypothesis that half of the people are in favor of Propostion 2 at the 10% level.
   B. I could not reject the hypothesis that half of the people are in favor of Propostion 2 at the 5% level.
   C. I could not reject the hypothesis that half of the people are in favor of Propostion 2 at the 1% level.
   D. Two of the above are correct.
   E. None of the above are correct.

	Lived	Died	Total
Crew	212	673	885
	(23.95%)	(76.05%)	(100.00%)
First	202	123	325
	(62.15%)	(37.85%)	(100.00%)
Second	118	167	285
	(41.4%)	(58.6%)	(100.00%)
Third	178	528	706
	(25.21%)	(74.79%)	(100.00%)
Total	710	1491	2201
	(32.26%)	(67.74%)	(100.00%)

Statistic	DF	Value	P-value
Chi-square	3	187.79321	<0.0001

9. The table above lists the survivor count by type of passenger (First Class, etc.) and crew. What statistical conclusion can be made from this data?

   A. The crew stayed behind to help the passengers get out safely.
   B. The rich (First Class passengers) were helped off first, so they were more likely to survive.
   C. The poor (Third Class passengers) were helped off last, so they were less likely to survive.
   D. There is a significant relationship between type of passenger (including crew members) and whether they survived or not.
   E. All of the above are valid conclusions for this data.

10. What is the expected count for a Second Class survivor (rounded to 0 decimals)?

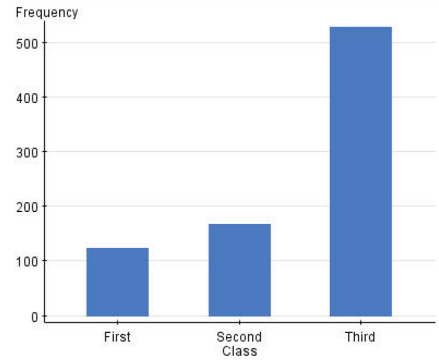
   A. 118
   B. 92
   C. 911
   D. 294
   E. 5

11. What assumption needed to be true for the previous test to be valid?

   A. Each expected cell count needed to be at least 30.
   B. Each expected cell count needed to be at least 10.
   C. Each expected cell count needed to be at least 5.
   D. The average expected cell count needed to be at least 5.
   E. None of the above are right.

12. Which of the following would be a Type I error in the χ^2 test?

   A. We concluded that more crew members died than any type of passenger (First, etc.), but it's not true.
   B. We concluded that the proportion of deaths was different for each class, but it wasn't.
   C. We failed to prove that there was a relationship between type of passenger and surviving.
   D. We concluded that the type of passenger and surviving were not related, but they are.
   E. We concluded that the type of passenger and surviving were related, but they are not.



13. Describe the shape of the distribution above.

   A. It is skewed to the right.
   B. It is skewed to the left.
   C. It is uniform since the bars are equally spaced.
   D. It could be any of the above.
   E. We can't describe the shape since the data is categorical.

14. How can I interpret, (0.12, 0.43), a 95% confidence interval for π, the true proportion of husbands who wished their wives would have at least some plastic surgery? (This is not based on any information. I totally made this up.)

   A. I am confident that the true proportion of husbands who wished their wives would have at least some plastic surgery falls between 12 and 43%.
   B. The true proportion of husbands who wished their wives would have at least some plastic surgery will fall between 12 and 43% 95% of the time in repeated sampling.
   C. 95% of the husbands wished their wives would have at least some plastic surgery between 12 and 43% of the time.
   D. Less than half of the husbands wished their wives would have at least some plastic surgery since the interval does not contain 50%.
   E. Two of the above are correct.

15. Suppose I tested $H_0 : \pi \geq 0.5$ vs. $H_A : \pi < 0.5$, found a sample proportion, $p = 0.34$ and a p-value $= 0.156$. How would I interpret the p-value?

   A. 15.6% of the time I would get a sample proportion of 34% or less even though the true proportion, π is at least 50%.
   B. 15.6% of the time I would get a sample proportion of 34% or more even though the true proportion, π is 50%.
   C. 15.6% of the time I would get a sample proportion of 50% or more even though the true proportion, π is 34%.
   D. 34% of the time I would get a sample proportion of 50% or less even though the true proportion, π is at least 15.6%.
   E. 15.6% of the time I would get a sample proportion of 50% or less even though the true proportion, π is at least 34%.

16. If my probability of success is $\pi = 0.4$ and my sample size is $n = 28$, what is the distribution of the sample proportion?

   A. It is not normal since $n = 28 < 30$.
   B. $p_{28} \sim N(0.4, 0.0926^2)$
   C. $p_{28} \sim N(0.4, 0.0086^2)$
   D. $p_{28} \sim N(0.0143, 0.0086^2)$
   E. $p_{28} \sim N(0.4, 0.0143^2)$

17. The Texas Transportation Institute is funding a study to estimate the number of traffic accidents caused by running red lights. You are hired by the TTI to help them in this study. It's said 10% of drivers run red lights, but you think it's MORE. You observed a sample of 100 cars at the Texas Ave. University Dr. intersection and found that 15% of them ran the red light. Are you right?

   A. The test statistic value is 1.67 with a p-value of 0.0475, so you are right at the 5% level.
   B. The test statistic value is 1.67 with a p-value of 0.095, so you didn't prove you were right at the 5% level.
   C. The test statistic value is 1.40 with a p-value of 0.0808, so you didn't prove you were right at the 5% level.
   D. The test statistic value is 1.40 with a p-value of 0.1616, so you didn't prove you were right at the 5% level.
   E. 15% is obviously more than 10%, so you claim you are right and don't bother to run a hypothesis test.

18. Which of the following statements is/are true?

   A. As the sample size increases, the standard *error* of the sample proportion, p, decreases, but the standard deviation stays the same.
   B. As the sample size increases, the sample proportion, p, will approach (get closer to) the true proportion, π.
   C. As the sample size increases, the sample proportion, p, will become less biased since it will get closer to π.
   D. All of the above are true.
   E. Only two of the above are true.

19. If $p_{20} \sim N(0.5, 0.1118^2)$, what is $P(0.25 < p_{20} < 0.75)$?

   A. 0.5, or 50%, since it's asking for the area between 25% and 75%
   B. 0.5 since $0.75 - 0.25 = 0.5$
   C. Since the sample size is only 20, we can't say the data is normal, so we can't do this calculation.
   D. 1
   E. 0.975

20. I want to know if the true proportion of American citizens who are over 75 is 5% or not, but all I have are these 3 confidence intervals.

```
90% (0.0211, 0.0589)
95% (0.0174, 0.0626)
99% (0.0104, 0.0696)
```

   What is my conclusion?

   A. We can't make hypothesis test conclusions from confidence intervals.
   B. Since 0.05 is in all three intervals, we fail to prove that the true proportion of American citizens over 75 is not 5%.
   C. Since 0.05 is in all three intervals, we fail to prove that the true proportion of American citizens over 75 is 5%.
   D. Since 0.05 is not in the 90% interval, we can conclude at the 10% level that the true proportion of American citizens over 75 is 5%.
   E. Since 0.05 is not in the 90% interval, we can conclude at the 10% level that the true proportion of American citizens over 75 is not 5%.

1A,2C,3D,4D,5B,6E,7A,8D,9D,10B,11D,
12E,13E,14E,15A,16B,17A,18B,19E,20B