

STAT303 Secs 509–511
Spring 2003
Exam #3
Form A

Instructor: Julie Hagen Carroll

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** on the exam. Multiple marks will be counted wrong.
3. You will have 60 minutes to finish this exam.
4. 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.
5. This exam is worth 100 points, and will constitute 20% of your final grade.
6. Good luck!

1. Suppose you tested $H_0 : \mu_1 = \mu_2$ vs. $H_a : \mu_1 \neq \mu_2$. Your data consisted of two samples with $\bar{x}_1 = 10$ and $\bar{x}_2 = 12$ and the resulting p-value was 0.806. Which of the following is the best interpretation of the p-value for this test?
 - A. There is an 80.6% chance that the two true means are equal.
 - B. There is an 80.6% chance of seeing at least this big of a difference in sample means when the true means are equal.
 - C. If we took many samples from these same populations, 80.6% of the time we would fail to reject H_0 .
 - D. If we took many samples from these same populations, 80.6% of the time we would see at least this big of a difference in true means when the sample means are equal.
 - E. If we took many samples from these same populations, 80.6% of the time we would see at least this big of a difference in sample means when the true means are equal.
2. Which of the following best describes a Type II error in the χ^2 test for independence?
 - A. We claim the row and column variables are independent when they actually have some relationship.
 - B. We claim the row and column variables have some relationship when they actually are independent.
 - C. We don't prove that there is a relationship between the row and column variables even though one exists.
 - D. We don't prove that the row and column variables are independent even though they are independent.
 - E. We claim we are financially independent, but mom and dad still pay the bills.
3. Many people say that dentists make them nervous. One way to determine whether someone is stressed, *e.g.*, from having to go to the dentist, is to measure their blood pressure. So, we want to know if visiting the dentist causes the mean blood pressure to increase. Which method would be the best to use?
 - A. Take a SRS of people who aren't nervous when they go to the dentist and another SRS of people who are nervous and compare the mean blood pressure for the two groups using a 2-sample t -test.
 - B. Take a SRS of people who aren't nervous when they go to the dentist and another SRS of people who are nervous and compare the mean blood pressure for the two groups using a pooled t -test since it's likely that the groups would have quite similar standard deviations.
 - C. Take a SRS of people, measure their blood pressure on a day that they aren't going to the dentist and again the day they do go, compare the results using a paired t -test.
 - D. Take a SRS of people who aren't nervous when they go to the dentist and another SRS of people who are nervous, measure their blood pressure before and after visiting the dentist and compare the results using a χ^2 test for proportions since nervous or not is a categorical variable.
 - E. none of the above
4. Does cooking in iron pots increase the iron content of food? A study was done using an Ethiopian beef dish. Four samples each were cooked in iron, clay and aluminum pots. What are the appropriate degrees of freedom for the F -test?
 - A. numerator df = 2, denominator df = 9
 - B. numerator df = 2, denominator df = 11
 - C. numerator df = 3, denominator df = 12
 - D. numerator df = 3, denominator df = 4
 - E. numerator df = 2, denominator df = 3

5. The composition of the earth's atmosphere may have changed over time. One attempt to discover the nature of the atmosphere long ago studies the gas trapped in bubbles inside ancient amber. The gas in bubbles within amber should be a sample of the atmosphere at the time the amber was formed. Measurements on specimens of amber from the late Cretaceous era give these percents of nitrogen:
- | | | | | |
|------|------|------|------|------|
| 63.4 | 65.0 | 64.4 | 63.3 | 54.8 |
| 64.5 | 60.8 | 49.1 | 51.0 | |
- If we also know that the present nitrogen level in our atmosphere is 78.1%. Which set of hypotheses should we test to determine if the nitrogen level has changed over time?
- A. $H_0 : \pi_{old} = \pi_{now}$ vs. $H_A : \pi_{old} \neq \pi_{now}$
 B. $H_0 : \pi_{old} = 0.781$ vs. $H_A : \pi_{old} \neq 0.781$
 C. $H_0 : \mu_{old} = \mu_{now}$ vs. $H_A : \mu_{old} \neq \mu_{now}$
 D. $H_0 : \mu_{old} = 78.1\%$ vs. $H_A : \mu_{old} \neq 78.1\%$
 E. $H_0 : \mu_{now} = 78.1\%$ vs. $H_A : \mu_{now} \neq 78.1\%$
6. What assumptions are necessary for us to be able to perform the test in the last problem?
- A. The sample of amber (actually the nitrogen level in the bubbles in the amber) must be random.
 B. The nitrogen levels in ancient amber must follow a normal distribution or we have to use a non-parametric procedure.
 C. The sample size must be large enough for $n\pi_{old}$ and $n\pi_{now}$ to be greater than 10.
 D. All of the above are necessary.
 E. Only two of the above are necessary.
7. When testing a statistical hypothesis, we fail to reject H_0 at the $\alpha\%$ level if
- A. H_0 is true.
 B. the hypothesized value under H_0 is less than the sample value.
 C. α is less than the p-value of the test statistic.
 D. All of the above are true.
 E. None of the above are true.
8. The *power* of a test is
- A. the probability that H_0 is true.
 B. the probability that H_0 is false.
 C. how often we correctly reject H_0 .
 D. the probability of rejecting a true H_0 .
 E. Exactly two of the above are correct.
9. A study was done to determine whether taking painkillers before surgery (the experimental group) was more effective than the regular practice of waiting until after to start the pain therapy (the control group). After $9\frac{1}{2}$ weeks, 12 of the 60 men in the experimental group were still feeling pain, and 18 of the 30 in the control group were still feeling pain. Which test should be used?
- A. the paired t-test for mean differences, pairing the men with similar pain symptoms.
 B. the pooled t-test for the difference of means since it's likely the the variability of the two groups is the same.
 C. the 2-sample *t*-test since both samples had more than 30 men.
 D. the 2-sample *z*-test for proportions
 E. the 2-sample *z*-test for means
10. When is a confidence interval more 'useful' than a hypothesis test? Consider only means here, and not proportions.
- A. when we want to run a two-sided test
 B. when we want an idea of the size of effect (mean or difference)
 C. when we don't know the true standard deviation
 D. when we don't know the true mean
 E. when we don't know if the data is normal
11. Suppose we tested $H_0 : \pi_1 = \pi_2$ vs. $H_A : \pi_1 < \pi_2$, found our test statistic to be -2.45 . What is our *p*-value?
- A. $P(Z > -2.45) = 0.9929$
 B. $P(Z < -2.45) = 0.0071$
 C. $2 * P(Z < -2.45) = 0.0142$
 D. $P(Z < 2.45) = 0.9929$
 E. $P(Z > 2.45) = 0.0071$
12. What does it mean when we say the ANOVA *F*-test is statistically significant?
- A. It means the effect being tested is significant.
 B. It means the means being tested are not all equal.
 C. It means the *p*-value is smaller than the α -level.
 D. All of the above are correct.
 E. Only 2 of the above are correct.

CI for Mean of normal, sigma known:
 $n = 16$, $\sigma^2 = 4$, $\bar{x} = 11$
 Lower Limit = 10.177573
 Upper Limit = 11.822427 90%

 Lower Limit = 10.020018
 Upper Limit = 11.979982 95%

 Lower Limit = 9.7120853
 Upper Limit = 12.287915 99%

13. Three confidence intervals based on the same data are given above. What is the correct range of the p -value for testing $H_0 : \mu = 10$ vs. $H_A : \mu \neq 10$?
- A. $p\text{-value} > 0.10$
 - B. $0.10 > p\text{-value} > 0.05$
 - C. $0.05 > p\text{-value} > 0.01$
 - D. $0.01 > p\text{-value}$
 - E. The sample mean is 11, not 11.8, so we can't use this information.

	ANOVA				
SCORE	SSqs	df	MSq	F	Sig.
Between	66.22	3	22.07	1.109	0.348
Within	2208.56	111	19.90		
Total	2274.78	114			

14. I decided to test if there was a difference in the four versions of the bonus quiz after the last exam. I ran the One-way ANOVA above and normal probability plots and found that the points fell along a line. What conclusion can be made?
- A. Since the normal probability plots weren't bell-shaped, the data's normal and the procedure is invalid for this data.
 - B. The data's fairly normal and as long as the standard deviations are relatively the same size, the F -test tells me that there was no significant difference in the four quizzes.
 - C. The data's fairly normal and as long as the standard deviations are relatively the same size, the F -test tells me that there's a 34.8% variation in the quiz scores.
 - D. The data's fairly normal and as long as the standard deviations are relatively the same size, the F -test tells me that the average of the four quizzes is 34.8%.
 - E. Since the F -test statistic is 1.109, the average on all the quizzes was just over 1, *i.e.*, on average everyone got one correct on the quiz.

15. When is the ONLY time you should run a paired t -test?
- A. when you have some link between the two samples
 - B. when you want less variability
 - C. when you want more power
 - D. when you can't tell if the variances are equal
 - E. when you don't know the true standard deviation

	next yr lose	next yr win	
this yr lose	83	37	120
Expected	59.0	61.0	120.0
this yr win	35	85	120
Expected	59.0	61.0	120.0
Total	118	122	240

Chi Sq = 38.411 $p\text{-value} = 0.000$

16. This table is data about stock performance for two successive years as to whether they were winners (their rate of return was higher than the median of all funds) or losers (their rate of return was less). If the performance of a stock fund is due to the skill of the manager, then a 'winner' fund should continue to do well in the following year. This is called persistence of fund performance. Is there evidence in favor of persistence of fund performance in this table?
- A. Yes, the p -value is 0, so we would conclude that there is a relationship between this year's performance and next year's.
 - B. No, the p -value is 0, so there is no persistence of fund performance.
 - C. No, the p -value is 0, so we couldn't prove there was a relationship between this year's performance and next year's.
 - D. Yes, the p -value is 0, so we claim performance from year to year is independent.
 - E. No, the war has changed people's perception of the value of stocks.

17. What assumption must be met in order for the χ^2 -test in the last problem to be valid?
- A. The data must be normal or the sample size (count) greater than 30.
 - B. The row and column variables must be independent.
 - C. The counts must be at least 5.
 - D. All of the above are necessary.
 - E. Only two of the above are necessary.
18. Suppose we had one sample of size 10 with sample mean and standard deviation, $\bar{x}_1 = 62$, $s_1 = 4.2$ and another of size 20 with $\bar{x}_2 = 67$, $s_2 = 5.1$. We test $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ and get a test statistic = -2.673 . What is the range of our p -value? Note: you must also figure out the correct test which gives you the degrees of freedom. Assume the data is normal.
- A. $df = 9$, $0.02 > p\text{-value} > 0.01$
 - B. $df = 9$, $0.04 > p\text{-value} > 0.02$
 - C. $df = 28$, $0.01 > p\text{-value} > 0.005$
 - D. $df = 28$, $0.02 > p\text{-value} > 0.01$
 - E. $df = 14$, $0.01 > p\text{-value} > 0.005$
19. Suppose you p -value for the last problem was 0.068. Which of the following is the best interpretation of this value?
- A. 6.8% of the time we will conclude the true means are different.
 - B. 6.8% of the time we will sample means of 62 and 67 when the true means are equal.
 - C. 6.8% of the time we will see a difference in the sample means of at least 5 when the true means are actually the same.
 - D. 6.8% of the time we will see a difference in the true means of at least 5 when the sample means are actually the same.
 - E. 6.8% of the time we will get a difference in the sample means even though the true means are actually the same.
20. I will not be here Thursday. If you want to know how you did on this exam, you may come to the lab at 11:30!!!! and Kristen will let you in, but she has class at 12:45, so come at 11:30. If you can't come then, the file will still be on the S: drive under instructors, jhc, exam3.xls and you can access it during Open Lab Hours. Please read Chapter 10 for next week. So how do you think you did? (mark what you think, but they all get 5 points)
- A. A
 - B. B
 - C. C
 - D. D
 - E. F
- 1E,2C,3C,4A,5D,6E,7C,8C,9D,10B,11B
12D,13C,14B,15A,16A,17C,18D,19C