

STAT303 Sec 508-510

Spring 2011

Exam #2

Form A

Instructor: Julie Hagen Carroll

Name: \_\_\_\_\_

1. **Don't even open this until you are told to do so.**
2. All graphs are on the last page which you may remove.
3. 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.
4. You will have **60 minutes** to finish this exam.
5. 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.
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 the 15% of your course grade.
8. 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 to the correct pile for your section.
9. Good luck!

1. Comparing the  $p$ -value to  $\alpha$  is the same as comparing
  - A. the critical (table) value to the test statistic.
  - B. the sample mean to the hypothesized mean.
  - C. the hypothesized mean to a confidence interval.
  - D. Two of the above are correct.
  - E. None of the above are correct.
2. Suppose we have a 95% confidence interval for the true test average,  $\mu$ , of (73.6, 85.8). Then we can say
  - A. 95% of the test takers made a B (in the 80's) or C (in the 70's).
  - B. 95% of all exams of this type will have an average in this same range.
  - C. 95% of the possible samples from this same population of test grades will have an average in this range.
  - D. 95% of the possible samples from this same population of test grades will produce intervals containing this range.
  - E. None of the above are correct.
3. The test below compares the women's labor force in 1972 and 1968.

Hypothesis test results:

mu1 : mean of 1972, mu2 : mean of 1968

mu1 - mu2 : mean difference

H0 : mu1 - mu2 = 0 vs. HA : mu1 - mu2 > 0

(with pooled variances)

Difference	Sample Mean	Std. Err.	DF	T-Stat	P-value
mu1 - mu2	0.0337	0.0225	36	1.496	0.0717

Which of the following is the correct conclusion?

- A. At the 10% level of significance, we can say that the women's labor force changed from 1968 to 1972.
  - B. At the 1 and 5% levels, we can say that there is no change in the women's labor force from 1968 to 1972.
  - C. At the 1 and 5% levels, we can say that there is a decrease in the women's labor force from 1968 to 1972.
  - D. At the 10% level of significance, we can say that there is an increase in the women's labor force from 1968 to 1972.
  - E. Two of the above are correct.
4. Using the three confidence intervals below, what is the correct range of the  $p$ -value if I wanted to test  $H_0 : \mu = 8.5$  vs.  $H_A : \mu \neq 8.5$ ?
 

90%	(5.456, 8.004)
95%	(5.211, 8.248)
99%	(4.735, 8.725)

    - 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. You need a test statistic value to determine the  $p$ -value.

5. Suppose I tested  $H_0 : \mu = 5$  vs.  $H_A : \mu \neq 5$  with a sample mean,  $\bar{x} = 7$ , sample size,  $n = 20$  and standard deviation,  $s = 2$  from a normal population. In which of the following situations would I be more likely to reject? Holding all other variables constant.
  - A. a larger sample size
  - B. a larger sample mean
  - C. a larger sample  $\mu_0$
  - D. All of the above would have a smaller  $p$ -value, so I would be more likely to reject.
  - E. Only two of the above would have a smaller  $p$ -value, so I would be more likely to reject.
6. What are the missing values in the ANOVA table below? Remember that the median of an  $F$  distribution is 1.

Location	n	Mean	Std. Error
Aurora	200	14.34	0.24875973
Denver	203	14.51	0.2827668
Parker	224	14.125	0.26259512

ANOVA table

Source	df	SS	MS	F-Stat	P-value
Treatments		16.10	8.052		
Error					
Total		9202.20			

- A. num df = 3, den df = 623,  $F = 14.72$ ,  $p$ -v is very small
  - B. num df = 2, den df = 624,  $F = 14.72$ ,  $p$ -v is very small
  - C. num df = 2, den df = 624,  $F = 0.55$ ,  $p$ -v > 0.50
  - D. num df = 624, den df = 626,  $F = 0.55$ ,  $p$ -v < 0.50
  - E. Too many values are missing to fill them all in correctly.
7. Say we created 100 confidence intervals from 100 samples of the same population. If we decrease the confidence level (99 to 95 to 90%),
    - A. less of the intervals will contain the true mean.
    - B. we will have more conservative intervals.
    - C. we would have to take larger samples (increase  $n$ ) to keep the widths the same.
    - D. All of the above are correct.
    - E. Two of the above are correct.
  8. Which of the following is/are true statements about the assumptions for hypothesis testing?
    - A. The more assumptions we can make (valid assumptions), the greater power our test has.
    - B. Assumptions give us the correct distribution for our data, so the  $p$ -value is correct.
    - C. We always need to assume that the distribution of the sample mean,  $\bar{x}$  is normal.
    - D. All of the above are correct.
    - E. Only two of the above are correct.

9. Rejecting at the 5% significance level means
- I will reject 5% of the time.
  - I will also reject at the 1% significance level.
  - I will make a Type II error 95% of the time since the probability of making a Type II error,  $\beta = 1 - \alpha$ .
  - I don't require 'as much evidence' to reject as I would at the 10% significance level.
  - None of the above are correct.
10. I want to test  $H_0 : \mu_1 = \mu_2$  vs.  $H_A : \mu_1 \neq \mu_2$ , but all I have is a 95% confidence interval for the true difference in the population means.
- If 0 is the center of the interval, then  $\mu_1 = \mu_2$ .
  - If 0 is in the interval, then  $\mu_1 = \mu_2$ .
  - If 0 is in the interval, then  $\bar{x}_1 = \bar{x}_2$ .
  - Two of the above are true.
  - None of the above are true.
11. This data shows the Sales Per Week for 3 Ice Cream Stores by Flavor of Ice Cream.
- | Column     | n   | Mean      | Std. Error |
|------------|-----|-----------|------------|
| Chocolate  | 628 | 14.85008  | 0.14720057 |
| Strawberry | 628 | 20.01595  | 0.25493953 |
| Vanilla    | 628 | 14.318979 | 0.15287355 |
- 
- | ANOVA table |      |         |         |        |         |
|-------------|------|---------|---------|--------|---------|
| Source      | df   | SS      | MS      | F-Stat | P-value |
| Treatments  | 2    | 12439.3 | 6219.68 | 270.03 | <0.0001 |
| Error       | 1881 | 43326.0 | 23.03   |        |         |
| Total       | 1883 | 55765.3 |         |        |         |
- What is the correct conclusion for this test?
- Since we are not given the standard deviations, we can't say that the assumption of equal variances is valid, so we can't make any conclusions.
  - Since the  $p$ -value is basically 0, there is no difference (0 difference) in the sample means.
  - Since the  $p$ -value is basically 0, there is no difference (0 difference) in the true (population) means.
  - Since the  $p$ -value is basically 0, we can say that *Strawberry* sells the most.
  - Since the  $p$ -value is basically 0, we can say that flavor does effect the average sale of these three choices.
12. What is the correct interpretation of the previous  $p$ -value?
- Never should we see this big of a difference in sample means if the true means are really the same.
  - Never should we see no difference in sample means if the true means are the same.
  - Never should we see 0 difference in the true means if the sample means aren't the same.
  - Never should we see 0 difference in the sample means if the true means aren't the same.
  - Never should we see an effect on average sales due to the flavor of ice cream.
13. What would be the consequence of a Type II error in the previous example from the store owner's point of view?
- buying more strawberry ice cream when each flavor sells about the same
  - buying the same amount of the three flavors when they really don't sell at the same rate
  - buying the same amount of the three flavors when they have equal sales
  - buying more chocolate or vanilla when you should have bought more strawberry
  - Seriously? strawberry's the favorite?
14. The larger the sample size,  $n$ , the
- less chance there is of making a Type I error.
  - less chance there is of making a Type II error.
  - less chance there is of having a biased sample.
  - All of the above are true.
  - Only two of the above are true.
15. Which of the following would give me the most power in a 1-sample test of the mean?
- $n = 100, \alpha = 0.10$
  - $n = 100, \alpha = 0.01$
  - $n = 50, \alpha = 0.01$
  - $n = 50, \alpha = 0.05$
  - $n = 100, \alpha = 0.05$
16. I want to test whether the lottery drawing is actually random, meaning every value from 000 to 999 is just as likely to be drawn. There are many tests that need to be done, but we'll start with testing the mean. Which set of hypotheses should we use?
- $H_0 : \mu = 0$  vs.  $H_A : \mu \neq 0$
  - $H_0 : \mu \leq 0$  vs.  $H_A : \mu > 0$
  - $H_0 : \mu \geq 0$  vs.  $H_A : \mu < 0$
  - $H_0 : \mu = 499.5$  vs.  $H_A : \mu \neq 499.5$
  - I would need A, B and C to be sure.
17. Which of the following is true for the previous question if you are the lottery commission?
- A Type I error would be claiming the lottery is random when it's not, so I should use  $\alpha = 0.10$ .
  - A Type II error would be failing to prove the lottery is not random when it's actually not random, so I should use  $\alpha = 0.10$ .
  - A Type I error would be claiming the lottery is not random when it is, so I should use  $\alpha = 0.01$ .
  - A Type II error would be claiming the lottery is random when it's not, so I should use  $\alpha = 0.01$ .
  - A Type II error would be failing to prove the lottery is random when it is, so I should use  $\alpha = 0.10$ .

18. What is the  $z$  critical value, the  $z_{\alpha/2}$ , for a 25% confidence interval?
- $\pm 0.675$
  - $\pm 0.32$
  - 0.5987
  - 0.646
  - 0.7734
19. We want to test whether directed reading activities in the classroom help elementary school students improve aspects of their reading ability. A treatment class of 23 third-grade students participated in these activities for eight weeks, and a control class of 23 third-graders followed the same curriculum without the activities. After the eight-week period, students in both classes took a Degree of Reading Power (DRP) test which measures the aspects of reading ability that the treatment is designed to improve. *Reference: Moore, David S., and George P. McCabe (1989). Introduction to the Practice of Statistics. Original source: Schmitt, Maribeth C., The Effects on an Elaborated Directed Reading Activity on the Metacomprehension Skills of Third Graders, Ph.D. dissertation, Purdue University, 1987.* What type test should we use if we can assume that all of the data is normal?  $\mu_1$  is the control group mean and  $\mu_2$  is the treatment group mean
- a 1-sample  $t$ -test testing the treatment mean is greater than the control mean
  - a 2-sample  $t$ -test testing the treatment mean is greater than the control mean
  - a paired  $t$ -test testing if the control mean minus the treatment mean greater than 0
  - a 2-sample  $t$ -test testing the control mean is greater than the treatment mean
  - a 1-sample  $t$ -test testing the control mean is greater than the treatment mean
20. Results of an experiment to test the question above are given below. This is a pooled  $t$ -test, not one of the choices above.
- Summary statistics:
- | Column    | n  | Mean     | Std. Dev. | Std. Err. |
|-----------|----|----------|-----------|-----------|
| Control   | 23 | 41.52174 | 17.148733 | 3.575758  |
| Treatment | 21 | 51.47619 | 11.007357 | 2.402002  |
- Hypothesis test results:
- mu1 : mean of Control, mu2 : mean of Treatment  
mu1 - mu2 : mean difference  
H0 : mu1 - mu2 = 0 vs. HA : mu1 - mu2 not = 0
- | Difference | Sample Mean | Std. Err. | DF | T-Stat | P-value |
|------------|-------------|-----------|----|--------|---------|
| mu1 - mu2  | -9.954      | 4.392     | 42 | -2.266 | 0.0286  |
- Which of the following is/are true?
- At the 5 and 10% levels, we could claim that the treatment (activities) worked (improved reading skills).
  - The true difference of the populations means would be in a 99% confidence interval, but not in the 95 nor 90%.
  - 0 would be in the 90 and 95% confidence intervals but not in the 99%.
  - Two of the above are true.
  - None of the above are true.
21. Had we done a 2-sample  $t$ -test instead of the pooled test above, what would be the correct range of the  $p$ -value given all the data stayed the same.
- $0.01 < p\text{-value} < 0.02$  with 20 df
  - $0.02 < p\text{-value} < 0.04$  with 20 df
  - $0.98 < p\text{-value} < 0.99$  with 20 df
  - $0.02 < p\text{-value} < 0.04$  with 22 df
  - $0.01 < p\text{-value} < 0.02$  with 22 df
- 1A,2E,3D,4C,5E,6C,7A,8E,9E,10E,11E,  
12A,13B,14B,15B,16D,17B,18B,19B,20E,21B