

# STAT303: Sections 508 – 510

Spring 2004

Exam #3

**Form A**

1. Don't EVEN open this until you are told to do so.
2. Be sure to mark your section number and your test form (A, B or C) on the scantron and sign your name where indicated on your scantron.
3. There are 20 multiple-choice questions, each worth 5 points. There is partial credit – you should only include the answers you are sure of to maximize your points. Be sure to mark your scantron CLEARLY. Multiple answers will be wrong!
4. You will have 60 MINUTES to complete this exam.
5. This exam is worth 100 points and constitutes 20% of your course grade.
6. If you are caught cheating or helping someone else cheat, you will both be given a ZERO for this exam. This includes discussing this exam with another student before next Thursday, April 29<sup>th</sup>!!!
7. GOOD LUCK! If you have any questions or comments, please include them on this exam. Turn in your scantron, help sheets and this exam when you are finished.
8. You may have the Review guide, flowchart and 3 other sheets of notes.

1. Are avid readers more likely to wear glasses than those who read less frequently? Three hundred men in the Korean army were selected at random and characterized as to whether they wore glasses and whether the amount of reading they did was above average, average, or below average. What type of test should we run to answer the question?

- A. a  $\chi^2$  test for multiple proportions
- B. an ANOVA  $F$  test for multiple means
- C. a 2-sample  $t$  test for the mean of those wearing glasses vs. the mean of those who don't
- D. a 2-sample  $z$  test for the proportion of those wearing glasses vs. the proportion of those who don't
- E. A non-parametric test since we don't know if the data is normal or not

Or

1. In the computer science department of a large university, many students change their major after the first year. A detailed study of the 256 students enrolled as first-year computer science majors in one year was undertaken to help understand this phenomenon. Students were classified on the basis of their status at the beginning of their second year, and several variables measured at the same time of their entrance to the university were obtained. If they decided that the students' high school math grades were the best variable to use, what type of test should they use? Assume they divided the students into three second-year majors: computer science, other engineering or science, other, and all the necessary assumptions are met.

- A. A  $\chi^2$  test of independence if they used A, B, C, D or F as the grade categories.
- B. A One-way ANOVA using numeric grades.
- C. A paired  $t$  test since it has the most power.
- D. Any of the above would be ok.
- E. Only two of the above would work.

2. What is the *best* conclusion given the output below?

	Unstandardized Coefficients		t	Sig
	B	Std. Error		
(Constant)	1928.206	888.680	2.170	.031
Beginning Salary	1.909	.047	40.28	.000

a. Dependent Variable: Current Salary

A. Since the  $p$ -value is 0.031, we can conclude at the 5% level that `Beginning Salary` is a good predictor for `Current Salary`.

B. Since the  $p$ -value is 0.031, we can conclude at the 5% level that `Beginning Salary` is NOT a good predictor for `Current Salary`.

C. Since the  $p$ -value is 0.000, we can conclude at the 5% level that `Beginning Salary` is a good predictor for `Current Salary`.

D. Since the  $p$ -value is 0.000, we can conclude at the 5% level that `Beginning Salary` is NOT a good predictor for `Current Salary`.

E. Since the  $p$ -value is 0.031, we can conclude at the 1% level that `Beginning Salary` is a good predictor for `Current Salary`.

3. Suppose I need to know whether the true test score is under 70, so I want to test  $H_0: \mu = 70$  vs.  $H_A: \mu < 70$ . If I sample the population 20 times and reject (conclude the true mean is under 70) twice (2 out of 20 times), what does this tell me?

A. The true mean really is under 70 since I rejected twice.

B. The true mean is probably not under 70. The 2 out of 20 rejections, 10%, is just my sample estimate of  $\alpha$ , the chance of making a Type I error.

C. The true mean is probably not under 70. The 2 out of 20 rejections, 10%, is just my sample estimate of  $\beta$ , the chance of making a Type II error.

D. The true mean is under 70. The 2 out of 20 rejections, 10%, is just my sample estimate of  $\alpha$ , the chance of making a Type I error.

E. The true mean is under 70. The 2 out of 20 rejections, 10%, is just my sample estimate of  $\beta$ , the chance of making a Type II error.

4. Which of the following is NOT an assumption of One-way ANOVA?

- A. Each sample must be random.
- B. The samples must be independent.
- C. The samples must be normal.
- D. The samples must have equal means.
- E. All of the above are valid assumptions.

5. Which of the following is an example of a matched pairs design?

A. A teacher compares the pre-test and post-test scores of students.

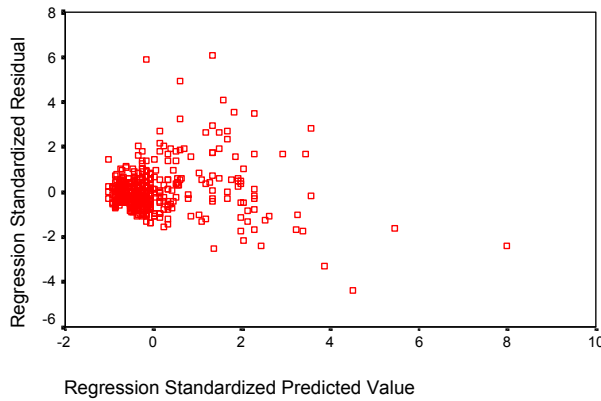
B. A teacher compares the scores of students using a computer-based method of instruction with the scores of other students using a traditional method.

C. A teacher compares the scores of students in her class on a standardized test with the national average score.

D. A teacher compares her class' average score on a standardized test with the national average score.

E. A teacher calculates the average of scores of students on a pair of tests and compares the two averages.

6. What does this residual plot tell us?



- A. The data is not linear.
- B. The data is not normal.
- C. The residuals do not have constant variance.
- D. The  $x$ 's do not have constant variance.
- E. The mean of the residuals is not zero.

7. Which would be a Type II error in the test:

$H_0: \mu_1 = \mu_2$  vs.  $H_A: \mu_1 > \mu_2$ ?

- A. claiming that  $\mu_1 < \mu_2$
- B. failing to prove  $\mu_1 > \mu_2$  when it's really true
- C. claiming that  $\mu_1 < \mu_2$  when it's really greater
- D. claiming that  $\mu_1 > \mu_2$  when it's not true
- E. failing to prove  $\mu_1 < \mu_2$  when it's really true

8. Suppose you tested  $H_0: \mu_1 = \mu_2$  vs.  $H_A: \mu_1 > \mu_2$  and got a  $p$ -value of 0.04. Which of the following statements would be correct?

- A. 0 would be in the 95% confidence interval for the difference  $\mu_1 - \mu_2$ .
- B. Conclude at the 5% level that  $\mu_1$  differs from  $\mu_2$ .
- C. 4% of the time we would see at least this much difference in sample means.
- D. Two of the above are correct.
- E. None of the above statements are correct.

9. What conclusion can be made about the *strain reading* for the different *machines* based on the output below?

Strain Reading					
	SS	df	MS	F	Sig.
Between	45.08	4	11.27	.914	.460
Within	924.9	75	12.33		
Total	970.0	79			

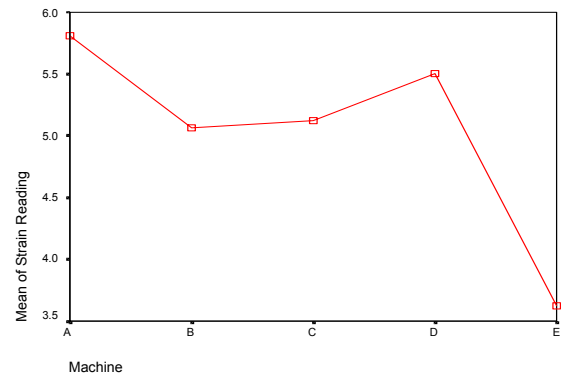
- A. There's a 46 % chance that the machines have an effect on the strain readings.
- B. Since the  $p$ -value is large, we have failed to prove that the machines have an effect on the strain readings.
- C. Since the  $p$ -value is large, we have proved that the machines have an effect on the strain readings.
- D. Since the  $p$ -value is large, we have failed to prove that the strain readings have an effect on the machines.
- E. Since the  $p$ -value is large, we have proved that the strain readings have an effect on the machines.

10. Which of the following statements is the correct interpretation of the  $p$ -value in the previous ANOVA  $F$  test?

- A. How often we would see a 46% difference in the sample means
- B. How often we would see at least this large of a difference in sample means when the true means are actually equal
- C. How often we would see at least this large of a difference in true means when the sample means are actually equal
- D. How often we would see at least a 46% difference in the sample means.
- E. How often we would get a 46% effect on the means.

Or

10. This is the plot of the means for the machines. Which of the following statements is/are true?



- A. The ANOVA  $F$ -test must be wrong because the means are obviously different.
- B. Since the means are not all different, we couldn't conclude that there was an effect.
- C. The variance of the data was so large that the difference in the sample means was not statistically significant.
- D. The plot of the means does not tell us whether the means are statistically different or not.
- E. None of the above statements are true for this plot.

11. What is the purpose of residual plots?
- A. to determine if  $x$  and  $y$  are linearly related
  - B. to determine if the assumptions for One-Way ANOVA test of equal means are valid
  - C. to determine if the true slope of the line,  $\beta_1$ , is 0 or not
  - D. Both A. and C. are correct.
  - E. None of the above are correct.

Or

11. The US Golf Association regularly tests golf equipment to ensure that it conforms to USGA standards. Suppose it wishes to compare the mean distance traveled by four different brands of golf balls when struck by a driver. How should they gather their data?
- A. get the data from the 4 different companies and run an ANOVA test to compare the means
  - B. poll golfers at a randomly selected golf course and record their brand and average drive (distance the ball is driven)
  - C. take a random sample of each brand and have a golfer hit all of them, in random order
  - D. take a random sample of golf courses and record which brands are used
  - E. take a random sample of each brand and have four golfers hit all of them, in random order
12. Rejecting  $H_0: \beta_1 = 0$  in the regression inference procedure says
- A. the  $x$ 's are not useful in predicting the  $y$ 's.
  - B. the slope of the regression line is large (*i.e.*, a big number).
  - C. the probability of making a Type II error is NOT 0.
  - D. Exactly two of the above are correct.
  - E. None of the above are correct.

13. The desired percentage of silicon dioxide in a certain type of cement is 5.0. A random sample of  $n=36$  cement specimens gave had an average percentage of 5.21 with a standard deviation of 0.38. If we wanted to test whether the cement specimens came from a population that had a true mean equal to the desired silicon dioxide percentage, which type test should we do?
- A. Case 6: the 1-sample test of a proportion using the desired percentage as the null value.
  - B. Case 1: the 1-sample test of a mean using the desired percentage as the true mean null value and 0.38 as the true standard deviation.
  - C. Case 3: the 1-sample test of a mean using the desired percentage as the true mean null value and 0.38 as the sample standard deviation.

- D. Case 11: the 2-sample test for 2 proportions using the desired percentage as one and the sample proportion as the other.
- E. Case 9: the 2-sample test for 2 means using the desired percentage as the true mean and the sample mean as the other.

14. Which of the following would lead us to believe that the  $t$  procedures were NOT safe to use for small samples?
- A. The sample means and medians for the two groups were slightly different.
  - B. The distributions of each group were moderately skewed.
  - C. The sample standard deviations were markedly different.
  - D. One of the samples had a few large outliers.
  - E. Any of the above would indicate we should use a non-parametric procedure.

15. An SRS of 100 flights of a large airline (call this airline 1) showed that 64 were on time. An SRS of 100 flights of another large airline (call this airline 2) showed that 80 were on time. Let  $\pi_1$  and  $\pi_2$  be the proportion of all flights that are on time for these two airlines. 90, 95 and 99% confidence intervals for the difference  $\pi_1 - \pi_2$  are:  $(-0.259, -0.055)$ ,  $(-0.279, -0.035)$  and  $(-0.317, 0.003)$  respectively. What is the range of the  $p$ -value for testing  $H_0: \pi_1 = \pi_2$  vs.  $H_A: \pi_1 \neq \pi_2$ ?
- 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 hypothesized value to know whether it's in the confidence interval or not. We don't have one, so we can't determine the  $p$ -value.

And

16. Which of the following are necessary assumptions for a  $\chi^2$  test of independence?
- A. We must have at least 30 in each cell.
  - B. If we don't have 30, then the data must be normal.
  - C. The standard deviations (variances) must be equal.
  - D. All of the above are necessary.
  - E. None of the above are necessary.

Or

15. A Simple Random Sample of 25 male faculty members at a large university found that 10 felt that the university was supportive of female and minority faculty. An independent SRS of 20 female faculty found that only 5 felt that the university was supportive of female and minority faculty. A 95%

confidence interval for the difference in the true proportions of male,  $\pi_M$ , and female,  $\pi_F$ , faculty at the university who felt that the university was supportive of female and minority faculty at the time of the survey is:  $0.15 \pm 0.270$ . If we want to test  $H_0: \pi_M = \pi_F$  vs.  $H_A: \pi_M \neq \pi_F$ , what conclusion should we make at the 5% significance level?

- A. We cannot decide without running a test of hypotheses.
- B. Since both sample proportions (0.4 for males and 0.25 for females) are not less than 0.05, we fail to reject  $H_0$  and conclude we could not prove that the true proportions are different.
- C. Since the sample proportions (0.4 for males and 0.25 for females) are not equal, we reject  $H_0$  and conclude that the true proportions are not equal.
- D. Since the confidence includes 0, we fail to reject  $H_0$  and conclude that we could not prove the true proportions are different.
- E. Since the confidence does not include 0, we reject  $H_0$  and conclude that the true proportions are different.

And

16. Which of the following is/are necessary assumptions for the data and test procedure above?

- A. We must have at least 30 in each sample.
- B. Since the sample sizes are less than 30, the data must be normal.
- C. The standard deviations (variances) must be equal.
- D. All of the above are necessary.
- E. None of the above are necessary.

17. Suppose we have two SRS's from two distinct populations and the samples are independent. We measure the same variable for both samples. Suppose both populations of the values of these variables are normally distributed but the means and standard deviations are unknown. For the purposes of comparing the two means we can use the pooled  $t$ -test procedure if

- A. the means for the two populations are actually equal.
- B. the standard deviations for the two populations are actually equal.
- C. the sample means for the two populations are equal.
- D. the sample sizes are the same.
- E. Two of the above are correct.

18. A SRS of size 100 is taken from a population having proportion 0.8 of successes. An independent SRS of size 400 is taken from a population having

proportion 0.5 of successes. The sampling distribution for the difference in the sample proportions,  $p_1 - p_2$ , has mean

- A. 0.3
- B. 0.65
- C. 0.5
- D. 0.8
- E. 0.13

19. What does it mean when we say a hypothesis test is 'statistically significant'?

- A. It means that we rejected the null hypothesis.
- B. In One-way ANOVA, it means that the effect on the means is significant.
- C. In regression, it means that the line is a significant prediction equation for  $y$ .
- D. All of the above are true.
- E. None of the above are true.

20. What conclusion can be made from the output below?

Minority Classification \* Gender Crosstabulation

		Gender		Total
		Female	Male	
No	Count	176	194	370
Yes	Count	40	64	104
Count		216	258	474

χ

	Value	df	Asymp. Sig. (2-sided)
Chi-Square	2.714	1	.099

- A. There are less minority females than males at the 10% level of significance.
- B. There is not a linear relationship between minority classification and gender in the data above.
- C. Gender has no effect on minority classification in the data above.
- D. There is not a statistically significant relationship between minority classification and gender in the data above at the 5% level of significance.
- E. There is a statistically significant relationship between minority classification and gender in the data above at the 5% level of significance.

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