

STAT303 Secs 508–510

Fall 2001

Exam #3

**Form A**

Instructor: Julie Hagen Carroll

1. **Don't even open this until you are told to do so.**
2. Be sure to write your instructor's name in the space provided and your name beneath.
3. 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.
4. You will have 60 minutes to finish this exam.
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. This exam is worth 100 points.
7. Good luck!

1. 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.
  
2. The sample size affects many things. Which of the following does it NOT affect?
  - A. the width of all of the confidence intervals covered in class
  - B. the chance of making a Type II error
  - C. the  $\alpha$ -level
  - D. the p-value
  - E. All of the above are affected.
  
3. 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
  
4. 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.
  
5. 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.
  
6. 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 course 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

7. What age do babies learn to crawl? Does it take longer to learn in the winter, when babies are often bundled in clothes that restrict their movement? Data on what month the child was born and the age (in weeks) at which the child was first able to creep or crawl a distance of four feet within one minute was reported by the parents. Grouped by birth month, the results are:

Month	Average Age	Std Dev.	n
January	29.84	7.08	32
May	28.58	8.07	27
September	33.93	6.93	38

The null hypothesis for the ANOVA  $F$ -test is that the population mean crawling ages are equal for all three birth months. The alternative hypothesis is

- A. that the population mean crawling age is larger for September than for the other two months.
  - B. that the population mean crawling age is smaller for May than for the other two months.
  - C. that the population mean crawling age for January is larger than for May but not for September.
  - D. at least the population mean crawling age for September is larger than for May.
  - E. None of the above are the correct statement of the alternative hypothesis.
8. Which of the following is true for the last problem?
- A. This is a randomized, designed experiment.
  - B. The data show no strong evidence of violating the ANOVA assumption of equal variances.
  - C. ANOVA is NOT the correct method to test this data since the sample sizes are not equal.
  - D. We don't need to run an ANOVA  $F$ -test since it is obvious that the means are not all equal.
  - E. Two of the above are true statements.

9. 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.
10. 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.

11. A sportswriter wished to see if a football filled with helium travels farther, on average, than a football filled with air. He used 18 adult males, randomly divided into 2 groups: one kicked the helium filled footballs, the other kicked air filled footballs. What are the correct hypotheses to test?
- A.  $H_0 : \mu = X$  vs.  $H_A : \mu \neq X$ , where  $X$  is the true average distance an air filled football travels when kicked
- B.  $H_0 : \mu_{air} = \mu_{helium}$  vs.  $H_A : \mu_{air} > \mu_{helium}$
- C.  $H_0 : \mu_{air} = \mu_{helium}$  vs.  $H_A : \mu_{air} \neq \mu_{helium}$
- D.  $H_0 : \mu_{air} = \mu_{helium}$  vs.  $H_A : \mu_{air} < \mu_{helium}$
- E.  $H_0 : \pi_{air} = \pi_{helium}$  vs.  $H_A : \pi_{air} > \pi_{helium}$ , where  $\pi$  is the proportion of field goals kicked
12. If we can assume that the two groups are independent and the data is approximately normal, what else would help us determine if helium filled footballs travel farther?
- A. if the kickers don't know which type of football they kicked, air or helium
- B. if the kickers don't know who the sportswriter is (or what paper he's with)
- C. if the kickers don't know where the uprights (goal) are
- D. if the kickers don't know how far others kick
- E. if the kickers don't know if they made the team or not
13. 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.

14. A researcher is studying treatments for agoraphobia with panic disorder. The treatments are to be the drug Imipramine at doses of 1.5mg per kg of body weight and 2.5mg per kg of body weight. There will also be a control group given a placebo. Thirty patients were randomly divided into 3 groups of 10 each. After 24 weeks of treatment, each of the subject's symptoms were evaluated through a series of tests, where a high score indicates a lessening of symptoms. Assume the data for the 3 groups are independent and approximately normal.

Group	Mean Test Score	Std Dev.
control	75.7	12.61
1.5mg	84.1	18.44
2.5mg	102.4	20.82

The p-value for this test is 0.007. Which of the following is correct?

- A. The standard deviation for the 2.5mg dose is not equal to the other two.
- B. The true variances for the three groups are not all equal.
- C. The true means for the three groups are not all equal.
- D. The 2.5mg dose works the best since it has the largest mean test score.
- E. The 2.5mg dose works 7% better than the other two.
15. The experiment above used a control group. Why is that a good idea?
- A. We need a control to determine the size of the effect.
- B. Having a control enables us to control the effect of confounding variables.
- C. Having a control makes the differences more statistically significant.
- D. Having a control makes the assumption of independence and normality more reliable.
- E. Having a control enables us to control the effect of the drug.

16. A political analyst was curious if younger adults were becoming more conservative. He decided to see if the mean age of registered Republicans was lower than that of registered Democrats. He selected a SRS of 128 Republicans from a list of those registered and got a mean  $\bar{x}_R = 39$  years with a sd  $s_R = 8$ . He also took a SRS of 200 registered Democrats and found the mean to be  $\bar{x}_D = 40$  years with a sd  $s_D = 10$ . The resulting p-value for testing whether the average age of Republicans is less than that of Democrats is 0.171. Which of the following is the best interpretation of this p-value?
- A. 17.1% of the time we would see a sample mean age of Republicans at least this much smaller than that of Democrats even though the true mean ages are the same.
  - B. 17.1% of the Republicans are younger than the average Democrat.
  - C. 17.1% of the time we would see a sample mean age of Republicans at least this different from that of Democrats even though the true mean ages are the same.
  - D. The Republicans are 17.1% younger than the Democrats.
  - E. 17.1% of the time we would find Republicans younger than Democrats even though they're the same age on average.
17. If the sample mean age of Republicans was 36 instead of 39, which of the following would be true (given everything else stayed the same)?
- A. More than 17.1% of the Republicans would be younger than the average Democrat.
  - B. The test would no longer be valid since the sample means are so different.
  - C. The p-value would be smaller.
  - D. The results would be more significant.
  - E. Two of the above are correct.
18. Twelve runners are asked to run a 10 kilometer race on each of two consecutive weeks. In one of the races, picked at random, the runners wear one brand of shoe and another brand in the second race. The runners are asked to run their best in both races and their times are recorded. Which method should we use to test whether the brand of shoe has an effect on run time?
- A. ANOVA  $F$ -test, assuming the data is normal.
  - B. Case 2: the 1-sample test of the mean, assuming that the data is normal.
  - C. Case 10: the paired test, pairing the brands by runner.
  - D. Non-parametric case for paired data since we don't know if the data is normal.
  - E. Case 9: the 2-sample test, assuming that the data is normal.
19. The pooled  $t$ -test and the 2-sample test for proportions pool some of the information. Why do we 'pool'?
- A. We don't know the true value, so we have to guess.
  - B. We can't assume that they're equal (either the sd's or the  $p$ 's), so we need a better estimate.
  - C. We get better estimates.
  - D. It makes the data more random (it comes from a 'pool' of choices).
  - E. None of the above are correct.
20. 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
- 1A,2C,3B,4D,5E,6C,7E,8B,9C,10D,11D  
12A,13B,14C,15B,16A,17E,18D,19C,20A