STAT303 Sec 507-510
Spring 2012
Exam #2
Form A
Instructor: Julie Hagen Carroll
May 2, 2012

1. **Do not open this exam until you are told to do so.**

2. There are 20 multiple-choice questions on this exam, each worth the same amount. Please mark your answers **clearly** on a GRAY Scantron sheet. Multiple marks will be counted wrong.

3. You **must mark** your Scantron form with
   
   (a) Your NAME and UIN.
   
   (b) Your correct SECTION (Thursday 11:10 is 507, 12:45 is 508, 2:20 is 509, and 3:55 is 510).
   
   (c) This test FORM (A, B, C, or D).
   
   (d) Your Form letter which is above.

4. You will have only 50 minutes to finish this exam.

5. You may use the following:
   
   (a) One $8\frac{1}{2} \times 11$ formula sheet (both sides) of your own making.
   
   (b) A copy of the CI and HT handout.
   
   (c) A copy of the $Z$ tables.
   
   (d) A stand-alone calculator, i.e., one that cannot communicate with the internet or anything outside itself.

6. You must put all possessions besides, the materials listed and your scantron, pencil(s) and eraser, along the walls or at the front of the room out of everyone else’s way. This includes cell phones, which must be turned off.

7. If you have questions, please write out what you are thinking on this test so that we can discuss it after your results are returned to you.

8. 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.

9. When you are finished please make sure you have marked your Section and Form and have an answer for every question, then turn in your scantron and show your ID.

10. Good luck!
1. You should use a significance level $\alpha = 0.01$ when
   A. you want more evidence in order to reject.
   B. you want your test to be more powerful.
   C. you want your confidence interval to be narrower but have a high confidence level.
   D. All of the above are reasons to use $\alpha = 0.01$.
   E. Only two of the above are correct reasons.

2. We suppose we test $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ and get a $p$-value = 0.072, then
   A. we would reject at the 1 and 5% levels and claim the true means are different.
   B. we would fail to reject at the 1 and 5% levels and claim the true means are the same.
   C. we would reject at the 10% level and claim the true means are different.
   D. we would reject at the 10% level and claim the sample means are different.
   E. Two of the above are correct.

3. According to the survey taken at the beginning of the semester, a 95% confidence interval for the expected average starting salary for STAT303 students is ($15,550, \$79,250$). From this we can say
   A. no one in this survey expects to make over $80,000.
   B. the true average starting salary, according to the survey, is $47,400.
   C. we are confident that the true average starting salary, according to the survey, is $47,400.
   D. Two of the above are true.
   E. None of the above are true.

4. I want to find out if, due to the sky-rocketing price of gold, Aggie rings are now lighter than they were in the past. Which set of hypotheses should I test?
   A. $H_0 : \mu_{old} = \mu_{now}$ vs. $H_A : \mu_{old} \neq \mu_{now}$
   B. $H_0 : \mu_{old} \leq \mu_{now}$ vs. $H_A : \mu_{old} > \mu_{now}$
   C. $H_0 : \mu_{old} \geq \mu_{now}$ vs. $H_A : \mu_{old} < \mu_{now}$
   D. $H_0 : \mu_{old} < \mu_{now}$ vs. $H_A : \mu_{old} \geq \mu_{now}$
   E. Did you know that the pre-1900 rings had lead in them?

5. If I run a test of hypotheses, $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$, and reject 45 out of 50 times (sampling from the same populations), which of the following is true?
   A. If $H_A$ is true, I should have rejected all 50 times, so it must be false.
   B. I made a mistake in 5 of the tests since I should have gotten the same answer for each test.
   C. My estimated power is 90%.
   D. All of the above are true.
   E. None of the above are true.

6. Which of the following is the best interpretation of the $p$-value = 0.001 when testing $H_0 : \mu = 50$ vs. $H_A : \mu < 50$? Note: $\bar{x} = 48$.
   A. 0.1% of the time we will reject the null hypothesis.
   B. 0.1% of the time we will get a sample mean, $\bar{x} = 48$ or more, when the true mean, $\mu = 50$.
   C. 0.1% of the time we will get a sample mean, $\bar{x} = 50$ or more, when the true mean, $\mu = 48$.
   D. 0.1% of the time we will get a sample mean, $\bar{x} = 48$ or less, when the true mean, $\mu = 50$.
   E. 0.1% of the time we will get a sample mean, $\bar{x} = 50$ or less, when the true mean, $\mu = 48$.

7. Suppose the sample size used in the problem above was, $n = 25$. If we took a different random sample of size $n = 40$ and still got $\bar{x} = 48$, what would be the new $p$-value after running the same set of hypotheses?
   A. 0.001 since it’s the same $\bar{x}$, hence the same distance between $\bar{x} = 48$ and $\mu_0 = 50$.
   B. 0 because it’s impossible to get the exact same sample mean, $\bar{x} = 48$.
   C. something less than 0.001 since there would be less standard deviations between $\bar{x} = 48$ and $\mu_0 = 50$.
   D. something more than 0.001 since there would be more standard deviations between $\bar{x} = 48$ and $\mu_0 = 50$.
   E. something less than 0.001 since there would be more standard deviations between $\bar{x} = 48$ and $\mu_0 = 50$.

8. Which of the following is/are true?
   A. A matched pairs test should always be used since it reduces the variance of the data.
   B. Two-sided tests have more power since you divide $\alpha$ for each of the tails.
   C. Confidence intervals are more useful than two-sided tests since they also give us an idea of the size of the effect, i.e., how different the sample and hypothesized means are.
   D. All of the above are true.
   E. Only two of the above are true.

9. Suppose we have a hypothesis $H_0 : \mu = 20$, an $\alpha = 0.05$, and a two-sided alternative. Unfortunately, the data was lost but a 95% confidence interval for $\mu$ was recovered. The interval was (0.5, 19.9). What decision rule should you make using this information?
   A. We cannot make a decision without a $p$-value.
   B. We should reject the null hypothesis at the 5% level since 20 is not within the interval.
   C. The results are inconclusive as the interval may or may not contain $\mu$.
   D. We should fail to reject the null hypothesis at the 5% level since 19.9 is very close to 20.
   E. None of the above.
10. Using the definition of a Type I error, which of the following would be a Type I error when we test: $H_0: \mu \geq 2$ vs. $H_A: \mu < 2$ at the $\alpha = 0.05$ level?

A. The $p$-value for our sample is 0.03 and the true mean, $\mu = 1$.
B. The $p$-value for our sample is 0.06 and the true mean, $\mu = 1$.
C. The $p$-value for our sample is 0.03 and the true mean, $\mu = 3$.
D. The $p$-value for our sample is 0.06 and the true mean, $\mu = 3$.
E. You can’t determine whether you made a Type I error or not because you don’t know if the null is true or not.

11. Suppose we are testing $H_0: \mu_1 = \mu_2$ vs $H_A: \mu_1 < \mu_2$. Sample 1 has 16 and sample 2 has 12, and we find a test statistic of -2.3. Assuming that both samples came from normal populations, what is the range $p$-value for the test?

A. between 0.025 and 0.02
B. between 0.05 and 0.04
C. between 0.02 and 0.01
D. between 0.04 and 0.02
E. The $t$ table cannot be used for negative numbers.

12. Suppose we ran a hypothesis test at the 5% significance level. Which of the following is true?

A. If we repeatedly sampled the data and ran the same test, we would reject about 5% of the time.
B. If we repeatedly sampled the data and ran the same test, we would fail to reject about 95% of the time.
C. If we repeatedly sampled the data and ran the same test, we would make a Type II error about 95% of the time.
D. If we repeatedly sampled the data and ran the same test, we would make a Type I error about 5% of the time.
E. Exactly two of the above are true.

13. A 95% CI for the mean $\mu$ of a population is computed from a random sample and found to be $12 \pm 4$. We may conclude:

A. there is a 95% probability that $\mu$ is between 8 and 14.
B. there is a 95% probability that the true mean is 12 and there is a 95% chance that the true margin of error is 4 divided by the sample size.
C. if we took many, many additional random samples and from each computed a 95% CI for $\mu$, approximately 95% of these intervals would contain $\mu$.
D. All of the above.
E. None of the above.

14. A new teacher at a university read an article that discussed a study of the amount of time (in hours) college freshman study each week. The study reported that the mean study time is 7.06 hours. The teacher feels that freshman at her university study more than 7.06 hours per week on average. The null and alternative hypotheses are $H_0: \mu = 7.06$ vs $H_A: \mu > 7.06$. The teacher selected a simple random sample of 15 freshmen at her university and found the observed sample mean study time to be 8.43 hours and the observed variance to mean standard study time to be 8.43 hours and the observed variance to be 18.66. Assume that study time for freshman at her university follow a normal distribution. The observed test statistic and degrees of freedom are

A. 0.317 with df = 14
B. 0.019 with df = 28
C. 1.23 with df = 14
D. 7.56 with df 28
E. None of the above are correct.

15. Suppose that the study time for freshman did not follow a normal distribution. What advice would you give the teacher?

A. Tell the teacher to use the $t$-test since we have a small sample size and the population standard deviation is unknown.
B. Tell the teacher to use the $z$-test since you are testing $\mu > 7.06$.
C. Tell the teacher to at least double the size of her sample and still use the $t$-test.
D. Tell the teacher to use a non-parametric test (which is a test for non-normal data).
E. Two of the above would work.

Hypothesis test results:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>DF</th>
<th>T-Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>seconds</td>
<td>31.49</td>
<td>0.75</td>
<td>464</td>
<td>2.6530</td>
<td>0.0082</td>
</tr>
</tbody>
</table>

16. It’s plausible that the number of seconds in the length of any song is a random number between 0 and 59. The output above is from the survey. What conclusions can we make?

A. The test is wrong because it should have tested 30 not 29.5 seconds.
B. We would reject the null and claim that the true mean is not 29.5 seconds.
C. It’s plausible that this is a Type I error since this should have been a random sample.
D. Two of the above are true.
E. None of the above are true.
17. The level of calcium carbonate ($\text{CaCO}_3$) in the wells of a certain area is of interest. The average $\text{CaCO}_3$ levels for 2001 and 2011 were to be compared. Ten randomly selected wells were measured for $\text{CaCO}_3$ level in 2001. The same ten wells are measured in 2011. In order to test whether the average $\text{CaCO}_3$ level in 2001 differs from the average $\text{CaCO}_3$ level in 2011, one should use

A. a histogram
B. an independent samples $t$-test
C. a paired $t$-test
D. a test that requires knowing the population standard deviation of differences between well readings.
E. the SAT test.

18. Suppose we wanted to test $H_0 : \mu = 16$ vs. $H_A : \mu \neq 16$. Our sample of size 24 had a mean, $\bar{x} = 20$ and $s_x = 5.5$, so the $p$-value for the test was 0.002. If we still got the same sample mean, $\bar{x} = 20$, but from a smaller sample of size, $n = 10$, how would the conclusion change from the one above?

A. The mean would not be normal anymore, so we couldn’t do the test.
B. The null hypothesis would be more likely to be true.
C. The null hypothesis would be less likely to be true.
D. We would decrease our chance of making a Type II error.
E. We would increase our chance of making a Type I error.

19. A government testing agency studies aspirin capsules to see if customers are getting cheated with capsules that contain lesser amounts of medication than advertised. Suppose the agency concludes the capsules contain an average amount below the advertised level when in fact the advertised level is the true mean. This is an example of

A. the agency misusing statistics.
B. a correct decision in hypothesis testing.
C. a hypothesis test with weak (low) power.
D. a Type I error.
E. a Type II error.

90% (6.1404199, 6.8595801)
95% (6.0689941, 6.9310059)
99% (5.9252143, 7.0747857)

20. Using the information above, what is the correct range of the $p$-value if I wanted to test $H_0 : \mu = 7$ vs. $H_A : \mu \neq 7$?

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.