AP Statistics – Unit 07
2-Sample CI & ST Review

Name Key
Period

Multiple Choice

1. Which of the following examples involves paired data?
   a) A study compared the average number of courses taken by a random sample of 100 freshmen at a university with the average number of courses taken by a separate random sample of 100 freshmen at a community college.
   b) A group of 100 students were randomly assigned to receive vitamin C (50 students) or a placebo (50 students). The groups were followed for 2 weeks and the proportions with colds were compared.
   c) A group of 50 students had their blood pressures measured before and after watching a movie containing violence. The mean blood pressure before the movie was compared with the mean pressure after the movie. same group
   d) A group of 40 students were blocked by gender and then randomly assigned different levels of caffeine to drink (no caffeine or 3 cups of black coffee). The blood pressures of each block were compared at the end of the day.
   e) None of the above.

C

2. In a large Midwestern university (the class of entering freshmen being on the order of 6000 or more students), an SRS of 100 entering freshmen in 1993 found that 20 finished in the bottom third of their high school class. Admission standards at the university were tightened in 1995. In 1997 an SRS of 100 entering freshmen found that 10 finished in the bottom third of their high school class. Let p_1 and p_2 be the proportion of all entering freshmen in 1993 and 1997, respectively, who graduated in the bottom third of their high school class. $\hat{p}_1 = 0.2$ $\hat{p}_2 = 0.1$

   Is there evidence that the proportion of freshman who graduated in the bottom third of their high school class in 1997 has been reduced, as a result of the tougher admission standards adopted in 1995, compared to the proportion in 1993? To determine this, you test the hypotheses: H_0: p1 = p2, H_A: $p_1 > p_2$.

   The P-value of your test is:
   a) Above 0.10
   b) Between 0.10 and 0.05
   c) Between 0.05 and 0.01
   d) Between 0.01 and 0.001
   e) Below 0.001

   2-Prop Z Test
   x_1: 20
   n_1: 100
   x_2: 10
   n_2: 100
   $p_1 > p_2$

   test statistic: 1.98
   p-value: 0.0238

C

3. To use the two-sample t procedure to perform a significance test on the difference between two means, we assume
   a) ~~The populations' standard deviations are known.~~
   b) The samples from each population are independent.
   c) ~~The distributions are exactly normal in each population.~~
   d) ~~The sample sizes are very large.~~
   e) ~~The populations' variances are equal.~~

   we check all of the others!

4. We wish to test if a new feed increases the mean weight gain compared to an old feed. At the conclusion of the experiment it was found that the new feed gave a 10 kg bigger gain than the old feed. A two-sample t-test with the proper one-sided alternative was done and the resulting P-value was 0.082. This means:
   a) There is an 8.2% chance the null hypothesis is true.
   b) There was only a 8.2% chance of observing an increase greater than 10 kg (assuming the null hypothesis was true).
   c) There was only an 8.2% chance of observing an increase greater than 10 kg (assuming the null hypothesis was false).
   d) There is an 8.2% chance the alternate hypothesis is true.
   e) There is only an 8.2% chance of getting a 10 kg increase.

B

5. In a comparison of the life expectancies of two models of washing machines, the average years before breakdown in an SRS of 10 machines of one model, which is compared with that of 15 machines of a second model, The 95 percent confidence interval estimate of the difference is (6,12). Which of the following is the most reasonable conclusion?
   a) The mean life expectancy of one model is twice that of the other.
   b) The mean life expectancy of one model is 6 years, while the mean life expectancy of the other is 12 years.
   c) The probability the life expectancies are different is 0.95.
   d) The probability the difference in life expectancies is greater than 6 years is 0.95
   e) We should be 95% confident that the difference in life expectancies is between 6 years & 12 years.

E

Free Response

1. A study aims to compare the size of butterfly wings between two related species. Eight randomly selected butterflies from each species yield the following wing sizes as measured in centimeters:

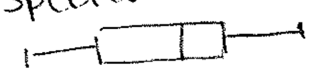
Species 1:	4	5	5	3	6	4	5	7
Species 2:	4	3	3	4	3	2	5	4

Calculate and interpret a 90 percent confidence interval estimate for the difference between the mean wing sizes of the two species.

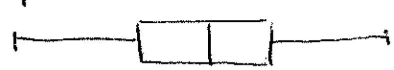
State: μ_1 = the true mean wing size of butterfly species #1 (in cm)
μ_2 = the true mean wing size of butterfly species #2 (in cm)
We want to find the true mean difference $\mu_1 - \mu_2$ with 90% confidence. $\bar{x}_1 = 4.875$cm $\bar{x}_2 = 3.5$cm

Plan: random: the data come from 2 independent random samples ✓
10% condition: $n_1 = 8$ 80 < all butterflies of species #1
$n_2 = 8$ 80 < all butterflies of species #2
Normal/Large: $n_1 = 8 < 30$ $n_2 = 8 < 30$
but a graph of each data set shows no strong skewness and no outliers:

species #1



species #2



Because our conditions are met, we will construct a 2-sample t-interval for the difference between 2 means $\mu_1 - \mu_2$.

Do: 2-SampTInt

$\bar{x}_1$: 4.875 $\bar{x}_2$: 3.5 C-level: 0.9
Sx1: 1.246 Sx2: 0.926 pooled: no = (0.40241, 2.3476)
n_1: 8 n_2: 8 df: 12.92 t^*: 1.772

Conclude: We are 90% confident that the interval from 0.40241cm and 2.3476cm captures the true difference in mean wing size between butterfly species #1 and #2 $(\mu_1 - \mu_2)$.

2. An SRS of 100 students at schools using an innovative math program scored and average of 357 with a standard deviation of 54 on a state test: an SRS of 150 students at schools using a traditional approach scored an average of 343 with a standard deviation of 62 on the same state test.

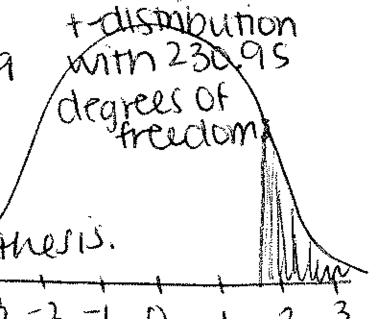
   a) Is there evidence that students using the innovative approach have a higher average score than students using the traditional approach? Give statistical justification for your answer by performing a significance test.

State: μ_1 = the true mean score of students in the innovative math program on the state test
μ_2 = the true mean score of students in the traditional math program on the state test.
$\bar{x}_1 = 357$ points $\bar{x}_2 = 343$ points $\alpha = 0.05$

$H_0: \mu_1 = \mu_2$ $H_A: \mu_1 > \mu_2$

Plan: random: the data come from 2 independent SRSs. ✓
10% condition: $n_1 = 100$ 1000 < all students in the innovative math program ✓
$n_2 = 150$ 1500 < all students in the traditional math program ✓
Normal/Large: $n_1 = 100 \geq 30$ ✓ $n_2 = 150 \geq 30$ ✓
Because our conditions are met, we will perform a 2-sample t-test for the difference of 2 means $\mu_1 - \mu_2$.

DO: 2-SampTTest
$\bar{x}_1$: 357 $\bar{x}_2$: 343 $\mu_1 > \mu_2$ p-value: 0.0299
Sx_1: 54 Sx_2: 62 pooled: no df: 230.95
n_1: 100 n_2: 150 test statistic: 1.891



Conclude: Because our p-value = 0.0299 is less than our significance level $\alpha = 0.05$ we reject the null hypothesis. There is convincing evidence that the true mean score on the state test is higher for students in the innovative math program than the traditional math program.

   b) Suppose a study using this design resulting in a P-value less than 0.01. Would it be reasonable for all school boards to push for adoption of the innovative approach? Explain.

No! We did not control any potentially confounding variables that may have impacted the scores earned by these students. Also, we cannot generalize our results to another (potentially very different) population.

3. The drug AZT was the first drug that seemed effective in delaying the onset of AIDS. Evidence for AZT's effectiveness came from a large randomized comparative experiment. The subjects were 870 volunteers who were infected with HIV, the virus that causes AIDS, but did not yet have AIDS. The study assigned 435 of the subjects at random to take 500 milligrams of AZT each day and another 435 to take a placebo. At the end of the study, 38 of the placebo subjects and 17 of the AZT subjects had developed AIDS.

   a. Do the data provide convincing evidence at the $\alpha = 0.05$ level that taking AZT lowers the proportion of infected people who will develop AIDS in a given period of time.

State: p_1 = the true proportion of AZT patients who develop AIDS.
p_2 = the true proportion of non-AZT patients who develop AIDS.

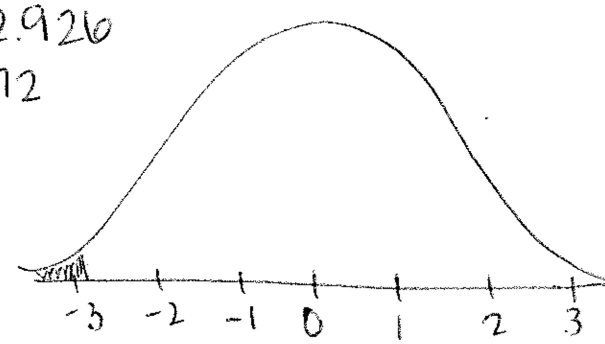
$H_0: p_1 = p_2$ $\hat{p}_1 = 17/435 = 0.0391$
$H_A: p_1 < p_2$ $\hat{p}_2 = 38/435 = 0.0874$ $\alpha = 0.05$

Plan: random: 870 volunteers randomly assigned to 2 independent treatment groups ✓
10% condition: $n_1 = 435$ 4350 < all HIV patients ✓
$n_2 = 435$ 4350 < all HIV patients ✓
Large counts: $n_1\hat{p}_1 = 17 \geq 10$ ✓ $n_2\hat{p}_2 = 38 \geq 10$ ✓
$n_1\hat{q}_1 = 418 \geq 10$ ✓ $n_2\hat{q}_2 = 397 \geq 10$ ✓
Because our conditions are met, we will perform a 2-sample z-test for the difference of 2 proportions $p_1 - p_2$.

Do: 2-PropZTest
x_1: 17
n_1: 435
x_2: 38
n_2: 435
$p_1 < p_2$

test statistic: −2.926
p-value: 0.00172

N(0,1)



Conclude: Because our p-value = 0.00172 is less than our significance level $\alpha = 0.05$, we reject the null hypothesis. There is convincing evidence that the true proportion of AZT patients who develop AIDS is less than for non-AZT patients, and that AZT may lower the proportion of infected people who will develop AIDS in a given period of time.

b. Describe a Type I Error and Type II Error in this setting and give a consequence of each error. Based on your conclusion in part (a), which error could have been made in this study?

Type I Error: AZT does not slow the onset of AIDS but we think it does. People take AZT (spend money) but still develop and die of AIDS as a result. We could have made this type of error.

Type II Error: AZT slows the onset of AIDS but we don't think it does. We don't use it (people die) but we continue research for a better drug/solution to AIDS.

c. Using the data given in the problem, create and interpret a 90% confidence interval for the true difference between the proportion of participants who developed AIDS while taking the placebo and the proportion of participants who developed AIDS while taking AZT.

State: We want to estimate the true difference in proportion $p_1 - p_2$ with 90% confidence.

Plan: Because our conditions are met, we will construct a 2-sample z-interval for the difference of 2 proportions $p_1 - p_2$.

DO: 2-PropZInt

x_1: 17

n_1: 435 = (-0.0753, -0.0213)

x_2: 38

n_2: 435

C-level: 0.90

Conclude: We are 90% confident that the interval from -0.0753 to -0.0213 captures the true difference in proportion $p_1 - p_2$ of HIV patients who developed AIDS while taking AZT and taking the placebo.