Exam 3 for Introductory Statistics

Prof. Bernardo Ábrego — Math 140, Spring 2006

Name.______________________________ 05/04/06

Directions.

- You have 1 hour and 50 minutes to complete the test.
- You can use a 3"×5" card with notes, your calculator, and the z-tables provided on the back of this test. No other materials are allowed.
- No scratch paper is permitted. For this purpose or in case you need more space you can use the back of the pages of the test.

1. A recent study of 100 people in Miami found 27 were obese. Find the 99% confidence interval of the population proportion of individuals living in Miami who are obese.

   $n = 100$ is the sample size, $\hat{p} = 27/100 = 0.27$ is the sample proportion. For 99% confidence we use $z^* = 2.576$. So the interval is given by $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ which gives,

   $$(0.15564, 0.38436)$$

2. A university dean of students wishes to estimate with 95% confidence the average number of hours students spend doing homework per week. It has been estimated that the population standard deviation is about 6.2 hours. How large a sample must be selected if he wants to be accurate within 1.5 hours?

   We know $\sigma = 6.2$. We want to find n so that the margin of error is at most 1.5 hours. We know that $E = z^* \frac{\sigma}{\sqrt{n}}$ for the confidence interval of a mean μ when σ is known. In this case $z^* = 1.96$ since the dean wants 95% confidence. Solving for n we get

   $$n = \left(\frac{z^*\sigma}{E}\right)^2 = \left(\frac{(1.96)\,(6.2)}{1.5}\right)^2 \approx 65.63.$$

   So the dean must select a sample of 66 students or more.

   Use the following three pages to conduct the following significance tests.

3. The Medical Rehabilitation Education Foundation reports that the average cost of rehabilitation for stroke victims is \$24,672. To see in the average cost of rehabilitation is different at a particular hospital, a researcher selects a random sample of 35 stroke victims at the hospital and finds that the average cost of their rehabilitation is \$25,226 with a standard deviation of \$3251. At $\alpha = 0.01$, can it be concluded that the average cost of stroke rehabilitation at a particular hospital is different from \$24,672?

4. A nationwide survey of large U.S. cities finds that the average commute time one way is 25.4 minutes. A chamber of commerce executive feels that the commute in his city is less and wants to publicize it. He randomly selects 25 commuters and finds the average is 22.1 minutes, with a standard deviation of 5.3 minutes. At $\alpha = 0.10$, is he correct?

5. In a sample of 80 workers from a factory in city A, it was found that 8 were unable to read, while in a sample of 50 workers in city B, 6 were unable to read. Can it be concluded that there is a difference in the proportions of nonreaders in the two cities? Use $\alpha = 0.05$.

Problem 3

Identify the test (circle one):

1. Significance test for one proportion (1-PropZTest)
2. Significance test for the difference of two proportions (2-PropZTest)
3. Significance test for a mean (T-Test)
4. Significance test for the difference of two means (2-SampTTest)

Data: Significance level $\alpha =$ 0.01

Other: (depending on the test, circle and give the necessary values)

$p_0 =$	$\hat{p} =$	$\overline{x} =$ 25226
$\mu_0 =$ 24672	$\hat{p}_1 =$	$\overline{x}_1 =$
$n =$ 35	$\hat{p}_2 =$	$\overline{x}_2 =$
$n_1 =$	$x_1 =$	$s =$ 3251
$n_2 =$	$x_2 =$	$s_1 =$
	$x =$	$s_2 =$

Check the conditions:

1. Sample is a SRS
2. Money measurements are usually normally distributed $n = 35 > 15$ and close to 40 (15/40) rule.
3. $10n = 350 <$ total population of patients

State the hypotheses:

H_0: $\mu = \mu_0 = 24672$

H_a: $\mu \neq \mu_0 = 24672$

where (circle and describe in words the appropriate symbol(s): $p, p_1, p_2, \mu, \mu_1, \mu_2$)

μ = average cost of rehabilitation for stroke victims at this particular hospital

Compute the test statistic, p-value, and label and complete the sketch:

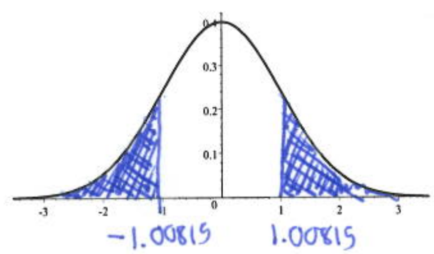
Test statistic: $\boxed{t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}}$ = 1.00815

formula

P-value = 0.3205

Is the sample significant? No

Yes/No



Conclusions:

(Circle one) We reject/don't reject the null hypothesis.

Explain in context.

Since P-value = 0.3205 > 0.01 then we do not reject. If indeed the average cost is \$24,672, then there is a 32.05% chance of getting a sample as extreme or more than \$25,226. Since this is quite likely. There is not enough evidence to reject the fact that the average rehab. expenses are \$24672.

Problem 4

Identify the test (circle one):

1. Significance test for one proportion (1-PropZTest)
2. Significance test for the difference of two proportions (2-PropZTest)
3. Significance test for a mean (T-Test)
4. Significance test for the difference of two means (2-SampTTest)

Data: Significance level $\alpha =$ 0.10

Other: (depending on the test, circle and give the necessary values)

$p_0 =$ ____ $\hat{p} =$ ____ $\overline{x} =$ 22.1

$\mu_0 =$ 25.4 $\hat{p}_1 =$ ____ $\overline{x}_1 =$ ____

$n =$ 25 $\hat{p}_2 =$ ____ $\overline{x}_2 =$ ____

$n_1 =$ ____ $x_1 =$ ____ $s =$ 5.3

$n_2 =$ ____ $x_2 =$ ____ $s_1 =$ ____

$x =$ ____ $s_2 =$ ____

Check the conditions:

1. Sample is a SRS
2. Distance measurements are usually normally distributed. $n=25$ is larger than 15 (15/40 rule). Proceed with caution
3. $10n = 250 <$ total population of commuters

State the hypotheses:

H_0: $\mu = \mu_0 = 25.4$

H_a: $\mu < \mu_0 = 25.4$

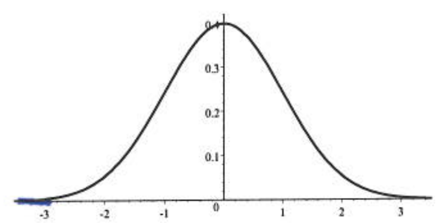
where (circle and describe in words the appropriate symbol(s): $p, p_1, p_2, \mu, \mu_1, \mu_2$)

μ = the average one-way commute time in this particular city

Compute the test statistic, p-value, and label and complete the sketch:

Test statistic: $\boxed{t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}}$ = -3.1132

formula



P-value = 0.002367

Is the sample significant? Yes

Yes/No

Conclusions:

(Circle one) We reject/don't reject the null hypothesis.

Explain in context.

Since the P-value = .0023 < .10 = α then we reject the null hypothesis in favor of the alternate hypothesis. If indeed the average one-way commute time in this city was 25.4 minutes then it would be quite unusual (0.23%) to get a sample less than or equal 22.1. Therefore we have evidence that the standard is no longer valid.

Problem 5

Identify the test (circle one):

1. Significance test for one proportion (1-PropZTest)
2. Significance test for the difference of two proportions (2-PropZTest)
3. Significance test for a mean (T-Test)
4. Significance test for the difference of two means (2-SampTTest)

Data: Significance level $\alpha =$ 0.05

Other: (depending on the test, circle and give the necessary values)

$p_0 =$ ____ $\hat{p} = \frac{14}{130} = 0.10769$ $\overline{x} =$ ____

$\mu_0 =$ ____ $\hat{p}_1 = \frac{8}{80} = 0.1$ $\overline{x}_1 =$ ____

$n =$ 130 $\hat{p}_2 = \frac{6}{50} = 0.12$ $\overline{x}_2 =$ ____

$n_1 =$ 80 $x_1 =$ 8 $s =$ ____

$n_2 =$ 50 $x_2 =$ 6 $s_1 =$ ____

$x =$ ____ $s_2 =$ ____

Check the conditions:

1. We assume the samples are SRS and they are independent
2. $n_1\hat{p}_1 = 8$, $n_1(1-\hat{p}_1) = 72$
   $n_2\hat{p}_2 = 6$, $n_2(1-\hat{p}_2) = 44$
   are all greater than 5
3. $10 n_1 = 800 <$ pop. of city A
   $10 n_2 = 500 <$ pop. of city B

State the hypotheses:

H_0: $p_1 = p_2$ or $p_1 - p_2 = 0$

H_a: $p_1 \neq p_2$ or $p_1 - p_2 \neq 0$

where (circle and describe in words the appropriate symbol(s): $p, p_1, p_2, \mu, \mu_1, \mu_2$)

p_1 = proportion of people in city A that are unable to read

p_2 = proportion of people in city B that are unable to read

Compute the test statistic, p-value, and label and complete the sketch:

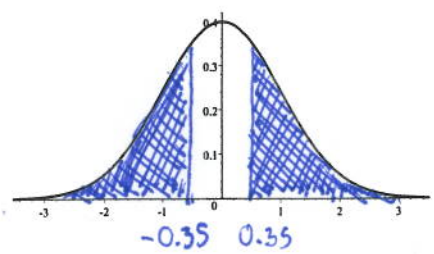
Test statistic: $\boxed{z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}} = -0.357881$

formula

P-value = 0.72043

Is the sample significant? No

Yes/No



Conclusions:

(Circle one) We reject/don't reject the null hypothesis.

Explain in context.

Since the P-value = 0.72043 > 0.05 we don't reject the null hypothesis. If the two cities have the same proportion of people unable to read, then it is quite likely (72.04%) to get a sample as extreme or more than the one we obtained. Thus we have no evidence to suggest that the proportions are different.