Name ______________________________________ Midterm Exam, Econ 371
Instructor: Ron Michener Fall 2004

This test is closed book and pledged, but you may use a calculator and the formula sheet that will be supplied to you. For credit you must show your work. If necessary, use the backs of the pages. *Please circle your answers.*

1) The weights of a particular kind of pill produced by a pharmaceutical company are known to have a normal distribution. A random sample of 5 pills are selected, and their weights (in grams) are discovered to be:

11.8 11.5 11.7 12.4 11.6

a) Compute the sample mean and sample standard deviation of the weights.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{59}{5} = 11.8$$

$$s = \sqrt{\frac{(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{.5}{4}} = .3536$$

b) Suppose that it is very important that the pills all contain the same amount of medicine, and as a consequence the manufacturer must be sure that the population variance of pill weights does not exceed 0.1. Use the data given above to test the following null hypothesis using $\alpha = .05$. Be sure to state the observed and critical values of the test statistic, and whether you accept or reject the null hypothesis.

$$H_0 : \sigma^2 \leq .10$$

$$H_A : \sigma^2 > .10$$

$$\chi^2_{.05} = 9.488$$

$$\chi^2_{obs} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{4(.125)}{.1} = 5$$

$$\chi^2_{obs} < \chi^2_{.05}$$

Accept H_0

2) A random sample of 800 supermarket shoppers will be taken, and asked whether they prefer generic brand items if the price is lower. Let p be the population proportion of shoppers who prefer generics when the price is lower. The data will be used to test the null hypothesis $H_0 : p \ge .50$ against the alternative $H_A : p < .50$ using $\alpha = .10$.

   a) If it is actually the case that 45% of all shoppers prefer generic brand items when the price is lower, what is the **power** of the proposed test?

$$-z_{.10} = -1.282 = \frac{\bar{p}^* - .50}{\sqrt{\frac{(.50)(.50)}{800}}}$$

$$\bar{p}^* = .4773$$

$$\textit{Reject if } \bar{p} < .4773$$

$$z = \frac{.4773 - .45}{\sqrt{\frac{(.45)(.55)}{800}}} = 1.55$$

$$power = P\left(\bar{p} < .4773 \middle| p = .45\right) = P\left(z < 1.55\right) = .9394$$

   b) When the sample is drawn, 378 of 800 shoppers say they prefer generic brands when they are cheaper. **Compute the p-value** associated with this outcome. Say whether you would accept or reject the null, using $\alpha = .10$, based on this p-value.

$$\bar{p} = 378/800 = .4725$$

$$z = \frac{.4725 - .500}{\sqrt{\frac{(.5)(.5)}{800}}} = -1.56$$

$$p - value = P\left(z < -1.56\right) = .0594$$

$$\textit{Reject } H_0;\ p - value = .0594 < .10 = \alpha$$

c) Indicate whether the following are **true** or **false.**

1) The significance level of a test is the probability that the null hypothesis is true. FALSE

2) A Type I error occurs when a true null hypothesis is rejected. TRUE

3) A null hypothesis is rejected at the .025 level, but is accepted at the .01 level. This means the p-value is between .01 and .025. TRUE

4) The power of a test is the probability of accepting a null hypothesis that is true. FALSE

5) If a null hypothesis is rejected against an alternative at the 5% level, then using the same data, it must always be rejected against that alternative at the 1% level. FALSE

6) If a null hypothesis is rejected against an alternative at the 1% level, then using the same data, it must always be rejected against that alternative at the 5% level. TRUE

7) The p-value of a test is the probability that the null hypothesis is true. FALSE

3) In October 1981 the *Chicago Sun Times* ran an expose on suicides (?) of prisoners in Chicago jails. In a 22-month period there had been 17 prisoner suicides in Chicago jails. During the same period there were only 16 prisoner suicides in Los Angeles and New York cities combined. The implication of the story was that the Chicago police were probably responsible for some of the so-called suicides. The obvious question not explicitly addressed by the *Sun Times* is: How likely is it that such a disparity in suicide rates could occur by chance? As a crude answer to this question, I looked up the population of these cities from the 1980 census. The population of Chicago was 3,000,000, while the population of New York and Los Angeles (combined) was 10,100,000.

a) If p_1 represents the proportion of Chicago residents who died in jailhouse suicides, and p_2 represents the proportion of New York and Los Angeles residents who died in jailhouse suicides, use this data to test the following null hypothesis using $\alpha = .002$ *(Note the unusual alpha!)* Do you accept or reject the null hypothesis?

$$H_0 : p_1 = p_2$$
$$H_A : p_1 \neq p_2$$

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{17+16}{13{,}100{,}000} = .000002519$$

$$\sigma_{\bar{p}_1-\bar{p}_2} = \sqrt{(.000002519)(.999997481)\left[\left(\frac{1}{3000000}+\frac{1}{10100000}\right)\right]}$$

$$\sigma_{\bar{p}_1-\bar{p}_2} = .000001044$$

$$z = \frac{\frac{17}{3000000} - \frac{16}{10100000}}{.000001044} = 3.91$$

$$z_{.001} = 3.09, \quad 3.91 > 3.09, \text{ so reject}$$

b) The disparity in suicide rates is striking and therefore newsworthy. Unfortunately, pre-test bias may undermine the validity of the formal hypothesis test. Explain.

The newspaper that broke this story probably had data on several different cities, and perhaps several different years' data as well; they almost certainly picked the comparison cities and (perhaps) the year to get the largest and most newsworthy disparity. Once you have sifted data, even informally, to focus on the most "interesting" anomalies, you can not use the anomalous data you find as the basis for a hypothesis test and expect alpha to be meaningful.

Please Write and Sign the Pledge!