Midterm #2

11/4/2013

Instructions: Each sub-question is worth 5 points. The exam is out of 90 points, so you get 10 points for free.

Refer to the provided probability tables as necessary. You may use a calculator, and one sheet of notes. You will never be penalized for showing work, but if what is asked for can be computed directly, points awarded will depend primarily on the correctness of your numerical answer. Good luck!

Problem 1 You wish to estimate μ, the average income of UK economics alumni. To do so, you use a university contact list to randomly sample 400 alumni to determine their income.

a. Explain, in detail, what you know about the sampling distribution of $\overline{x}$, the sample mean.

b. Suppose that income is known to have a standard deviation of $\sigma = 10,000$. What is the probability that your sample mean, $\overline{x}$, differs from μ by more than \$750?

c. After conducting your sample, you determine that $\overline{x} = \$50,000$. Calculate a 95% confidence interval for μ.

Problem 2 Dean Blackwell is concerned that Gatton professors are canceling too many classes. However, it is not feasible for the dean to personally monitor all of the hundreds of courses offered by Gatton each year, so he randomly chooses 25 of them, and checks each day to see if the class is being taught or not. The dean considers an acceptable number of canceled classes to be two or fewer per semester, per course. From past experience, the dean knows that the standard deviation of number of canceled classes is $\sigma = .5$.

a. State the appropriate null and alternative hypotheses for a test of whether or not too many classes are being canceled.

b. For what values of the sample mean $\overline{x}$ do you reject the null hypothesis, with $\alpha = .05$?

c. Suppose in his sample of 25 courses, the dean finds that an average of 2.15 classes are canceled per course. What is the p-value of the test?

d. Explain in words what the p-value you found in part c tells you.

Problem 3 Consider the following regression equation:

$$y = \beta_0 + \beta_1 * X + \epsilon \tag{1}$$

a. Explain in plain English how to interpret the coefficient β_1.

b. Explain in plain English why the ϵ term is necessary.

c. Suppose that x is mileage and y is the selling price of a certain type of used car. You estimate (1) using data provided by a local dealership. Do you expect the regression results to indicate that $\hat{\beta}_1 > 0$ or that $\hat{\beta}_1 < 0$? Why?

d. Suppose that x is the number of McDonald's restaurants in a country and y is the GDP of a country. You estimate (1) using a dataset containing information on 150 countries, and find a large, positive, and statistically significant estimate $\hat{\beta}_1$. Does this mean that McDonald's restaurants cause a high GDP? Explain.

Problem 4 Consider the following regression model:

$$WAGE = \beta_0 + \beta_1 * EXPERIENCE + \epsilon \tag{2}$$

where $WAGE$ is hourly wage and $EXPERIENCE$ is years of full-time work experience. Suppose you estimate the regression parameters using a large Census dataset (a random sample of all U.S. residents). Your results are as follows:

Regression	**Statistics**
Multiple R	.22358
R Square	.24985
Adjusted R Square	.24985
Standard error	5.2198
Observations	189,253

	coefficients	Standard error	t stat	P-value	Lower 95%	Upper 95%
Intercept	7.11	2.00	3.56	.000186	4.0924	14.1276
EXPERIENCE	.96	.31	3.097	.000978	.3524	1.5676

a. Consider the following hypothesis test:

$$H_0 : \beta_1 = 0$$
$$H_A : \beta_1 \neq 0$$

Is there sufficient evidence to support rejecting H_0, with $\alpha = .05$?? How do you know?

b. Interpret the estimate of β_1 precisely.

Differing levels of full-time experience is often suggested as one reason why women earn a lower wage than men. Suppose that, among 45-year old workers, women have an average of 17.7 years of full-time work experience, while men have an average of 22.6 years of full time work experience.

c. Give a point estimate for the wage of a worker with 17.7 years of experience. Give a point estimate for the wage of a worker with 22.6 years of experience.

d. Suppose that the average 45-year old man earns a wage which is \$6.76 higher than the average wage of a 45-year old woman. What fraction of this difference is explained by differing levels of full-time work experience, according to the results?

Problem 5 A criminologist is concerned about the number of crimes in a particular town that involve guns. He randomly chooses 600 crime reports and finds that 420 of the crimes involve a gun.

a. Develop a 95% confidence interval for the proportion of crimes that involve a gun.

b . What is the error of estimation for this interval estimate?

c. How many reports would the criminologist need to review in order to reduce the margin of error to .01?