Economics 20 Prof. Patricia M. Anderson

Problem Set 1
(due at the beginning of class October 9, 2000)

Part I**.** For each of the following questions, choose the one best answer. Briefly explain your reasoning.

1. If **X** and **Z** are two random variables, then E[**X**-**Z**] is
   a. E[**X**] – E[**Z**] + 2E[**XZ**]
   b. E[**X**] + E[**Z**]
   c. the same as E[**Z**-**X**]
   d. E[**X**] – E[**Z**]

2. Any linear combination of normally distributed random variables:
   a. is also normally distributed
   b. is t-distributed
   c. is F-distributed
   d. none of the above

3. What can be said about the estimated slope coefficient for a regression of Y on X, versus the estimated slope coefficient for a regression of X on Y?
   a. the slopes are reciprocals
   b. the slopes are not reciprocals
   c. the slopes are the negative of each other
   d. the slopes are identical

4. Suppose the following is the true population model of the effect of smoking on birth weight in ounces: *birth weight* = β_0 + β_1*daily cigarettes* + *u*, and that the average birth weight is 125 oz and the average mother smoked 2 cigarettes per day. If your estimate of β_1 is -2.5, what must your estimate of β_0 be?
   a. 120
   b. 125
   c. 130
   d. not enough information

5. Suppose a new sample is chosen and you estimate the following: *predicted birth weight in oz = 136 – 2.1 daily cigarettes*. What would be your estimate of β_0 if birth weight were recorded in pounds?
   a. 136
   b. 8.5
   c. 33.6
   d. 13.1

6. Using the same sample as in question 5, you now estimate the following: *birth weight in oz* = β_0 + β_1*weekly cigarettes* + *u*. Your estimate of β_1 would be:
   a. – 2.1
   b. – 3
   c. – 0.3
   d. –14.7

Part II. Stata Problems. See "Important Things to Know about Stata" on our website for more about installing and using Stata. This PS uses the Stata commands **sum**, **graph**, the "**if** qualifer" and **reg**.

1. Alabama.dta has standardized test scores from 127 public school districts in the State of Alabama in the late 1980's. You should copy it from our web site into the folder from which you are running Stata. Only the following 2 variables will be used in this problem:

score89	Average reading and math standardized test score for 8-9 grade students. (In standard deviation units).
pcy	Per capita income in the district.

a) Use the appropriate commands to find the mean and median for the test score and per capita income.

b) Create histograms showing the distribution of test scores and of per capita income. Be sure to choose an appropriate bin size in each case.

c) Use the appropriate commands to answer these questions. What is the average test score in districts with per capita income at or above the median? What is the average test score in districts with per capita income below the median? What is the difference in average per capita income between these high and low-income districts?

d) Create a two-way graph to examine the relationship between per capita income and test scores for all school districts.

e) Run a regression of test scores (the dependent variable) on per capita income (the independent variable). How do you interpret the sign and magnitude of the coefficient on income? Did this come as a surprise, given the results in c) and d)? Briefly explain your answer.

2. Copy nflqb.dta, which has information on the 1993 statistics (and 1994 salary) for 60 NFL quarterbacks, from our web site. Several variables are interesting, but for this problem we'll only use:

TDs	Number of TD passes thrown
attempts	Number of passes attempted
completions	Number of passes completed

a) Run a regression of touchdowns on pass attempts. Based on this regression, how many more touchdowns would a quarterback be expected to get if he threw 100 more passes over the season?

b) Now add completed passes to the regression. Based on this regression, what would happen, all else equal, if the quarterback threw 100 more passes? What makes this prediction different from the one in a)?

c) Based on the regression in b), how would completing 100 more passes affect the number of touchdowns a quarterback throws? (Note: Be careful about your interpretation here).

IMPORTANT: Please submit all the graphs along with the relevant portions of your log file. You should delete errant commands and output from the log file, paste in the graphs, and type in your answers and explanations to produce one concise answer sheet for the group. Do not forget to include the names of all group members!