Department of Economics
Winter 2007

Econometrics 120B
Professor Berman

Problem Set #1
Due Thursday, January 18

Please answer the questions on this sheet (or photcopy) in pen.

1. **Point Spreads and Hypothesis Testing**

Gambling is despicable practice with a rich history in statistical analysis. If you are compulsive by nature or given to addiction please avoid the temptation to gamble away your college funds. Nevertheless, gambling is useful to help us think about probability, statistics, and hypothesis testing.

Gambling on sports often takes the form of betting on the "spread" which is the difference in points scored between two teams, in say, a football game. For instance if P_c are points scored by the San Diego Chargers and P_o are points scored by their opponent then the actual spread is $A = P_c - P_o$.

Organized gamblers predict the spread for all NFL games and accept bets. Call their predicted spread S. People bet on whether the favored team "beats the spread" i.e., if A > S. You have to bet \$110 to win \$100.
(I.e., if you bet \$110 you have \$210=(\$100+\$110) if A>S and lose \$110 if $A \leq S$.)

Imagine recording A (from the game) and S (from some particular gambling website before the game) for a sample of N Chargers' games and calculating $Y = A - S$ for each game.

a. Why would someone who gambles on Chargers' games in this way be interested in these data?

b. Consider a simple statistical model $Y = \mu + \epsilon$ where μ is an unknown parameter and ϵ is a random variable with $E(\epsilon) = 0$. What is the expected value of Y?

c. Suppose that Y is independently **normally** distributed with variance 9 and $\mu = 0$. Suppose you have a random sample $Y_1, Y_2, \ldots Y_N$ drawn from the distribution for Y. If you bet \$110 on the Chargers beating the spread for each of N games, what is the expected value of your winnings?

d. Suppose that Y is independently **normally** distributed with variance 9 and $\mu=0$. Suppose you draw random samples, each of size N from the distribution for Y, each time calculating the sample mean, *M*. What is the distribution of the random variable *M* ?

What are the mean and variance of *M* in the population?

e. Suppose that Y is independently **normally** distributed with variance 9. Suppose you want to test the null hypothesis $\mu = 0$ against the alternative that $\mu \neq 0$. If the sample size is 101 and you use a .05 significance level, what values of the sample mean $\overline{Y}$ would lead you to reject the null hypothesis?

f. Suppose that Y is independently **normally** distributed with variance 9, $\mu=0$ and S= 4.5 for the next game. What is the probability that San Diego wins that next game?

g. What are the advantages of assuming that Y has an independent normal distribution in part (e)?

h. Suppose that Y is independently distributed with variance 1 and $\mu=0$ but the distribution is not necessarily normal. Can you answer part (e)? If so, show your answer.

i. What is a **disadvantage** of assuming that Y has an independent normal distribution in part (f)?

2. **Data gathering exercise**. (*Submit a page or two with two graphs, stapled to this question sheet.)*

a. Gather data from a *controlled experiment*. Report the values of one variable (X) that you controlled (or randomly assigned) and another (Y) that was influenced by X. You should have at least 15 observations. (The "demand for lunch" survey performed in class was an experiment in which X was price and Y was quantity demanded.)

Report the source of the data, the sample mean of each variable and an X-Y graph (scatterplot) using Stata.

b. Gather *nonexperimental* data on two variables X and Y that may be related. You should have at least 15 observations. (GRE score and an indicator for "attended GRE prep course" are an example.)

Report the source of the data, the sample mean of each variable and an X-Y graph (scatterplot) using Stata.

Notes: Experimental data are hard to come by. You may choose to simply conduct an experiment yourself.

Nonexperimental data are available from many published sources, including USA Today and the Baseball Digest. On the Internet try www.nber.org/data or www.census.gov.