Department of Economics — Spring 2006
University of California — Prof. Woroch

Economics 140
Problem Set 1
Due: Thursday, February 2, 2006, 2:00 PM (before lecture)

1. Consider the experiment that involves rolling <u>two dice</u> and let the random outcome *M* be the sum of the dots that turn up on the two dice. *M* is a whole number running from 1 through 12.
   a) In an **EXCEL** spreadsheet, list the possible outcomes for *M* together with their associated probabilities.
   b) Using **EXCEL,** graph the "probability density function" (p.d.f.) for this experiment and the corresponding "cumulative distribution function" (c.f.d.), placing the values of *M* on the x-axis and probability values on the y-axis.
   c) Calculate the expected value and the standard deviation of the random variable *M*.
   d) What is the probability of the following outcomes?
      (i) Pr($M = 7$)
      (ii) Pr($M = 2$ or $M = 10$)
      (iii) Pr($M = 4$ or $M \neq 4$)
      (iv) Pr($M = 6$ or $M > 10$)

2. The heights of male college students in the U.S. are normally distributed with a mean of 70 inches and a standard deviation 3.5 inches, i.e., $N(70, 3.5^2)$. If you had a phone book with all U.S. male college students, if you randomly selected one and call him, what would be the probability that his height would be:
   a) taller than 6'0"?
   b) between 5'3" and 6'5"?
   c) shorter than 5'7" (the mean height of female students)?

   Now suppose that you conducted a survey that sampled a given number of male students and asked their height. Find the probability of the following events involving the sample mean, X, generated by your survey when you take different sample sizes, *n*:

   d) *Pr(X < 6'2")* when $n = 25$
   e) *Pr(X > 5'3")* when $n = 100$
   f) *Pr(5'7" < X < 6'0")* when $n = 144$

3. You are interested in investigating the relationship between the age of heads of households and weekly earnings of households. The accompanying table gives the number of occurrences grouped by age and income for 1,744 individuals. After sorting the data, you generate the accompanying table. The median of the income group of $800 and above is $1,050. When answering the questions, treat this sample as representative of the full population.

   a) Calculate the joint relative frequencies and the marginal relative frequencies (use a similar table as the one described above). Interpret one of each of these.
   b) Sketch the cumulative income distribution using **EXCEL**, placing household income by group on the x-axis and Probability or Percent on the y-axis.
   c) Calculate the conditional relative income frequencies for the two age categories 16-20, and 45-65. Calculate the mean household income for both age categories.

Joint Absolute Frequencies of Age and Income, 1,744 Households

		Age of head of household				
		X_1	X_2	X_3	X_4	X_5
Weekly Household Income		16-20	20-25	25-45	45-65	> 65
Y_1	\$0-\$200	80	76	130	86	24
Y_2	\$200-\$400	13	90	346	140	8
Y_3	\$400-\$600	0	19	251	101	6
Y_4	\$600-\$800	1	11	110	55	1
Y_5	> \$800	1	1	108	84	2

4. In accord with the last will and testament of Alfred Nobel, the wealthy manufacturer of dynamite, five Nobel Prizes are awarded each year for outstanding achievements in Chemistry, Physics, Physiology or Medicine, Literature, and Peace. In 1968, the Bank of Sweden added a 6th prize in Economic Sciences in memory of Nobel. The accompanying table lists the joint probability distribution between recipients of economics prize and the original five prizes, and the citizenship of the recipients. These probabilities are based on award of prizes over the 1969-2001 period. Treat these data as representative of the true underlying population.

   a) Compute $E(Y)$ and interpret the resulting number.
   b) Calculate and interpret $E(Y \mid X = 1)$ and $E(Y \mid X = 0)$
   c) A randomly selected Nobel Prize winner reports that he is a non-U.S. citizen. What is the probability that this genius has won the Economics Nobel Prize? A Nobel Prize in the other five disciplines?
   d) Show what the joint distribution would look like if the two categories were independent (use a similar table as below).

Joint Distribution of Nobel Prize Winners in Economics and Non-Economics Disciplines, and Citizenship, 1969-2001			
	U.S. Citizen ($Y = 0$)	Non-U.S. Citizen ($Y = 1$)	Total
Economics Nobel Prize ($X = 0$)	0.118	0.049	0.167
Physics, Chemistry, Medicine, Literature, and Peace Nobel Prize ($X = 1$)	0.345	0.488	0.833
Total	0.463	0.537	1.00

5. Go to the class website and download the data PS1data.xls. This dataset contains 100 observations. The variables are defined as:

FEMALE	1 if the individual is female and 0 if the individual is male.
AGE	Age of the individual
COLLEGE	1 if the individual has a college degree, 0 otherwise.

a) Use **EXCEL** to complete the following table:

	Total	Males	Females
Number of Obs.			
Average Age			
Median Age			
% People w/ Age>30			
Std Dev Age			
% College Degree			

b) Use **EXCEL** to compute the sample correlations between FEMALE and COLLEGE, and between AGE and COLLEGE.