Economics 471: Econometrics

Department of Economics, Finance and Legal Studies
University of Alabama

Problem Set #1

1. You are given the following information on 12 individuals' years of education and yearly income

Education	10	18	12	16	16	20	21	13	9	11	6	13
Income (in thousands)	12	56	23	60	41	89	111	39	19	25	18	46

   (a) Find the average and median level of years of education and yearly income. What difference do you notice between the median and average levels?

   (b) Plot yearly income vs. years of education. Does the relationship look linear?

   (c) Draw a line through the cluster of points that you think best represents the actual relationship.

   (d) Calculate the covariance and the correlation coefficient of years of education and yearly income.

   (e) Suppose yearly income is measured in \$10,000 instead of \$1,000 (hint: multiply each yearly income by ten). Calculate the covariance and the correlation coefficient of years of education and yearly income now.

   (f) What do the findings from (e) say about using covariance as a measure of association compared to the correlation coefficient?

2. Prove that $\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} x_i (y_i - \overline{y})$ and prove that $\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} (x_i - \overline{x}) y_i$.

3. You own a business that sells cheat sheets. Your production function is

$$\#cheatsheets = 1 + 2\sqrt{hours}$$

   (a) Plot out this relationship.

   (b) What is your marginal product?

   (c) Suppose your true production function was linear instead. What would your marginal product be. What is the difference between your marginal product when you have a linear production function versus a square root production function?

   (d) Which one seem 'economically' more reasonable? What does this suggest about representing every relationship in this class linearly?

4. A random variable X is defined to be the difference between the higher value and the lower value when two dice are thrown. If they have the same value, X is defined to be zero.

   (a) Is X a discrete or continuous random variable?

   (b) Find the probability density and cumulative distribution for the random variable.

   (c) What do the cdf and pdf look like if X is defined to be the larger of the two values when the dice are thrown or the value of either one when they roll up the same?

   (d) Find the expected value of X in (a) and (c).

5. Show that $V(X) = E\left(X^2\right) - [E(X)]^2$.

6. The following table gives the college GPA of students versus their corresponding high school GPA

		High School GPA			
		2.0	3.0	4.0	Total
College GPA	2.0	12	3	3	18
	3.0	6	9	6	21
	4.0	2	21	7	30
	Total	20	33	16	69

   (a) Convert the above table into one giving the joint probability distribution.

   (b) Compute $p(College\ GPA = 2.0 \mid High\ School\ GPA = 2.0)$, $p(College\ GPA = 2.0 \mid High\ School\ GPA = 3.0)$ and $p(College\ GPA = 2.0 \mid High\ School\ GPA = 4.0)$.

   (c) Compute $E(College\ GPA \mid High\ School\ GPA = 2.0)$, $E(College\ GPA \mid High\ School\ GPA = 3.0)$ and $E(College\ GPA \mid High\ School\ GPA = 4.0)$.

   (d) Plot the conditional expectation of College GPA versus High School GPA.

   (e) Do the conditional expectation you discovered in part (c) agree with *a priori* expectations about the relationship between high school performance and college performance? How do those who receive a 4.0 in high school perform in college? Should this result be surprising?

7. Is unbiasedness a necessary or sufficient condition (or neither) for consistency of an estimator?