Dr. Markus Haas
Department of Empirical Research and Econometrics

Exercise Session - Problem Set 2

Fundamentals of Probability Theory

Problem 1 (Based on Greene, Example 7, p. 96)

Suppose that the random variables Y and X have the following joint probability distribution.

		X = 0	1	2
Y	0	0.05	0.1	0.03
	1	0.21	0.11	0.19
	2	0.08	0.15	0.08

a) What are the marginal distributions of the two random variables?

b) Compute the following probabilities:

   (1) $Prob[Y < 2]$

   (2) $Prob[Y < 2, X > 0]$

   (3) $Prob[Y = 1, X \geq 1]$

c) Calculate the unconditional mean of X and the unconditional mean of Y. Calculate also $Var(X)$ and $Var(Y)$.

d) Calculate $Cov(X, Y)$ and $Corr(X, Y)$. Under which assumption the two random variables X and Y would be independent? Does independence mean uncorrelatedness? Explain. What is about the converse? Does zero correlation between two random variables imply that they are independent? Explain.

e) What are the conditional distributions of Y given $X = 2$ and of X given $Y > 0$?

f) Find the expected value of Y conditional on X, $E(Y|X)$.

Problem 2 (Based on Greene, Example 8, p. 96)

Minimum mean square predictor and minimum mean square linear predictor.
For the joint distribution in Problem 1, compute $E(y - E(y|x))^2$. Now find the a and b that minimize the function $E(y - a - bx)^2$. Given the solutions for a and b, verify and interpret the following inequality $E(y - E(y|x))^2 \leq E(y - a - bx)^2$.

Problem 3

Partial effects in the conditional expectation function.
Consider the following conditional expectation functions:

1) $E(y|x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \beta_4 x_3$
2) $E(log(y)|x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \beta_4 x_3$,

with y being the expected value of hourly wage and $log(y)$ being the logarithm of hourly wage, respectively. The explanatory variables are years in education (x_1), years in the workforce (x_2), and a gender dummy (x_3).
Are the models in 1) and 2) linear in the explanatory variables and in the population parameters? What are the partial effects with respect to x_1, x_2, and x_3? Give an interpretation of the partial effects on the two conditional expectations based on the wage equation example. Explain the differences between 1) and 2).