ECON 4130: Practice Problem set 6

1-3: Rice Edition 3: Chapter 4: Problem 75, 78, 80.

4: Lecture notes to Problem set 5: Exercise 1

5: Here is another practice problem, which appears in last year's supplementary exercises. You should be able to solve it given our discussion on Chapter4.

Let (X,Y) be two random variables with some joint distribution, $f(x,y)$, expected values, μ_X, μ_Y, and variances, σ_X^2, σ_Y^2 respectively. Suppose that the regression of Y with respect to X is linear and homoscedastic, i.e.,

$$(1) \quad \begin{aligned} &E(Y \mid x) = E(Y \mid X = x) = \alpha + \beta x && \text{where } \alpha, \beta \text{ are constants} \\ &\operatorname{var}(Y \mid x) = \operatorname{var}(Y \mid X = x) = \sigma^2 && \text{(constant)} \end{aligned}$$

a. Show, using the law of double expectations (and the corresponding one for variances), that (1) implies

$$(2) \qquad \mu_Y = E(Y) = \alpha + \beta\mu_X$$

$$(3) \qquad \sigma_Y^2 = \operatorname{var}(Y) = \sigma^2 + \beta^2\sigma_X^2$$

b. Show that

$$(4) \qquad \operatorname{cov}(X,Y) = \beta\sigma_X^2 \qquad \left(\text{or } \beta = \frac{\operatorname{cov}(X,Y)}{\sigma_X^2}\right)$$

[**Hint:** Note that $E(XY) = E\left[E(XY \mid X)\right] = E\left[X \cdot E(Y \mid X)\right]$, which follows since the inner expectation, $E(XY \mid X)$, is of the form, $h(X)$, where $h(x)$ is determined by the conditional distribution of XY when X is fixed to the value x. I.e., $h(x) = E(XY \mid x) = E(XY \mid X = x) = E(xY \mid X = x) = xE(Y \mid X = x) = xE(Y \mid x)$, since x is a constant in the expectation. Hence, replacing x by the r.v. X, we get $h(X) = X \cdot E(Y \mid X) = X(\alpha + \beta X)$ and so on ….]

c. Show that (3) and (4) imply that

$$(5) \qquad \sigma^2 = \sigma_Y^2(1-\rho^2)$$

where $\rho = \rho(X,Y) = \text{cov}(X,Y)/(\sigma_X \sigma_Y)$ is the correlation between X and Y.

[**Hint:** Solve (4) for β and substitute in (3).]

[**Note** that solving (5) with respect to ρ^2 gives an alternative interpretation of ρ: Interpreting σ^2 as measuring the part of Y which *is not* explained by the regression relation in (1), then $\rho^2 = (\sigma_Y^2 - \sigma^2)/\sigma_Y^2$ measures the part of the variation of Y, (i.e., σ_Y^2) which *is* explained.]

d. The model (1) may be reformulated as follows: Write

(6) $$Y = \alpha + \beta X + u$$

where the "error term" u is simply defined as $u = Y - \alpha - \beta X$. Show that (1) implies:

(7) $$\text{E}(u \mid x) = 0 \quad \text{and} \quad \text{var}(u \mid x) = \sigma^2$$

and therefore also

(8) $$\text{E}(u) = 0 \quad \text{and} \quad \text{var}(u) = \sigma^2$$

Show (as in **b.**) that (7) also implies that u and X are uncorrelated, i.e.,

(9) $$\text{cov}(u, X) = 0$$

e. Show the other way round, i.e. that (6) and (7) imply (1).