ECON 221

Assignment 6

Due: November 8, 2012

1. {Goldberger 25.1} Suppose that x and y are bivariate-normally distributed with $E(y|x) = \alpha + \beta x$, $V(y|x) = \sigma^2$, and $V(x) = \sigma_x^2$. In random sampling, sample size n from this population, let b be the sample slope and let s_x^2 be the sample variance of x. Let

$$\begin{aligned} z =& \sqrt{n}(b - \beta)/(\sigma/s_x) \\ w =& n s_x^2/\sigma_x^2 \\ u =& \sqrt{(n-1)}(b - \beta)/(\sigma/\sigma_x) \end{aligned}$$

   (a) Show that $z \sim N(0, 1)$, that $w \sim \chi^2(n-1)$, and that z and w are independent.
   (b) Show that $u \sim t(n-1)$.
   (c) Explain how the result in (b) completely specifies the marginal distribution of the sample slope in terms of parameters and sample size.

2. {S.W. 11.10} Suppose that a random variable Y has the following probability distribution: $\Pr(Y = 1) = p$, $\Pr(Y = 2) = q$, and $\Pr(y = 3) = 1 - p - q$. A random sample of size n is drawn from this distribution and the random variables are denoted $Y_1, \ldots, Y_n$.

   (a) Derive the likelihood function for the parameters p and q.
   (b) Derive formulas for the MLE of p and q.
   (c) Derive the variance for the score variable.

3. Consider again model (1) from Assignment 3. Once again using the **Growth** data set, but this time including the data for Malta, compare the coefficient estimates when using OLS and using LAD for the variable $TradeShare$. Now exclude Malta and once again compare the OLS and LAD estimators. Including Malta, which estimator might be more appropriate? What about excluding Malta? Should Malta be excluded? Explain.

4. {Greene Ex 23.1} On ICON, you will find FLP.txt, data on female labor force participation. Estimate the following model

$$\Pr[LFP = 1] = F(constant,\ age,\ age^2,\ family\ income,\ education,\ kids)$$

for $F(\mathbf{x}, \beta) = \mathbf{x}'\beta$ (linear probability model), $F(\mathbf{x}, \beta) = \Phi(\mathbf{x}'\beta)$ (Probit), and $F(\mathbf{x}, \beta) = \frac{e^{\mathbf{x}'\beta}}{1+e^{\mathbf{x}'\beta}}$ (Logit).

(a) Compare the marginal effects from the three models at averages

(b) Find the expected probability of labor force participation for a woman with average values of each variable (rounded up to the nearest integer, with age^2 the squared value of the average of age) using the LPM and Logit models. What does this say about their comparative effectiveness?

(c) Using the Probit model,

i. Interpret the marginal effects (and their statistical and economic significance)

ii. Does age seem to have a nonlinear effect on labor force participation? At what age is a female most likely to participate in the labor force?