## Cornell University <br> Department of Economics

Econ 620 - Spring 2008
Instructor: Professor Kiefer
TA: Jae Ho Yun

## Problem Set \#6

(Due: May 2, Friday)

1. Logit: Your dependent variables are indicators $d_{i}$, ( 0 or 1 ) and your model is $\operatorname{Pr}\left(d_{i}=1\right)=\frac{1}{1+\exp \left(-\alpha-\beta x_{i}\right)} . x_{i}$ is a scalar. Suppose you fit this model by nonlinear least squares.
(a) How to get NLS estimator? Describe the procedure.
(b) What is the variance of $d_{i}$ conditioning on $x_{i}$ ?
(c) Suppose you use your estimator from a and your formula from b to "correct" for heteroskedasticity and rerun your nonlinear least squares. Is your second-round estimator "better" than your firstround estimator from a)? In what sense?
(d) We know that Maximum likelihood is the most efficient. How to get ML estimator? Describe the procedure.
2. (Final 2002) You now decide that the outcome variable of interest in the problem 1 of PS\#5 should not be the log wage, but whether or not the musician has a recording contract. You observe $d_{i}=1$ if the player has a contract, $d_{i}=0$ if not. Your specification is $\operatorname{Pr}\left(d_{i}=1\right)=F\left(x_{i}^{\prime} \beta+\delta t_{i}\right)$, where F is some function $R \rightarrow[0,1]$. t remains unobserved, but you have $t^{*}$ and z ( z includes x ).
(a) How would you estimate $\beta$ and $\delta$ ?
(b) How would you test whether measurement error in $t$ is a problem?
3. (Final 2005) The credit scoring problem. You are interested in the probability that a loan given to an applicant with characteristics x will default. You have a sample of N loans, $1 / 3$ with x values $0,1 / 3$ at 1 , and $1 / 3$ at 2. For each observation, you see $x$ and an indicator $d$ of whether the loan defaulted $(\mathrm{d}=1)$ or not. Let $F_{0}, F_{1}$ and $F_{2}$ be the default probabilities associated with each value of x . You are particularly interested in the value of $F_{1}$.
(a) How might you estimate $F_{1}$ sensibly? Give a good unbiased estimator and its variance (the variance will depend on the true value, of course).
(b) You now assume a logit model for the dependence between F and x; specifically $F_{i}=1 /\left(1+\exp \left\{-x_{i} \beta\right\}\right)$ for $\mathrm{i}=1,2,3$ and some $\beta$. How would you estimate $\beta$ ? What is the asymptotic variance of the MLE for $\beta$ ? (Assume that as n gets large, $1 / 3$ of the observations are from each of the 3 possible values of x ).
(c) What is the asymptotic variance of your MLE for $F_{1}$,specifically $1 /(1+\exp \{-\beta\})$ ?
(d) Compare this with the variance calculated in a. What are the strengths and weaknesses of the two approaches (be brief).
(e) How might you test whether your logit specification is appropriate?
4. (Final 2005) Your model for the nonnegative dependent variable $y_{i} \sim$ $c_{i} \exp \left(-y_{i}\left(x_{i} \beta\right)\right)$ (conditional density of $\left.y_{i}\right)$, where $x_{i}$ is a $1 \times K$ vector of regressors and $\beta$ is the vector of coefficients to be estimated.
(a) What is $\mathrm{c}_{i}$ ? (Note that we are given a density function)

For a sample of N independent observations ( $y_{i}, x_{i}$ ),
(b) give the loglikelihood function, the score function and the second derivative matrix.
(c) What is the asymptotic distribution of the MLE, $\beta_{M L}$ ?

Now you want to plot some "residuals" to check the specification. Hence, you need a transformation $z_{i}$ of $y_{i}$, given $x_{i}$ and $\beta$, such that the distribution of z does not depend on x and $\beta$. Consider the random variable $z_{i}=1-\exp \left(-y_{i}\left(x_{i} \beta\right)\right)$.
(d) What is the distribution of $z_{i}$ ?

Suppose you calculate z using $\beta_{M L}$ instead of $\beta$ and you plot the empirical cdf F-hat $(\mathrm{t})=\left(\# z_{i} \leq t\right) / N$.
(e) What should this plot look like if your model is correct (and the estimate of $\beta$ is good)?
5. (Final 2006) A funny regression problem (Freedman). ( $y_{i}, x_{i}$ ) are iid with means zero. Define $b$ as the minimizer of $E(y-x b)^{2}$ (so $b=E x y / E x^{2}$ ). Suppose you have a sample of size $n$ from the distribution $x \sim N(0,1)$ and $y=x^{3}$.
(a) What are the mean and sampling variance of $b^{*}$, the OLS estimator (recall that these expectations are conditional on the sample of $x$ 's realized)?
(b) What is $b$ ? (Hint: recall the mgf of a standard normal r.v.).
(c) Find a good approximation to the estimated standard error of $b^{*}$, to t-statistics (for $b=0$ ), and to $R^{2}$ (Hint: the same hint).
Despite these good results, you are suspicious of the specification so you include another regressor, $x^{2}$, to capture nonlinearities.
(d) What is the coefficient on $x^{2}$ (asymptotically)?
(e) Could you discover this misspecification with a plot of fitted vs. actual values? Provide a sketch.
6. (Final 2007) The Gauss Markov Theorem. Consider the linear model $y=X \beta+\epsilon$ with uncorrelated errors with common variance. We saw in class that the linearity requirement in the GMT could be abandoned. Now we will consider modifying the unbiasedness assumption. Linear estimators have the form $T(y)=c^{\prime} y$ for some $n \times k$ matrix $c$.
(a) Show that the OLS estimator for the linear model is a linear estimator.
(b) Give an example of a linear estimator that is biased.

The mean squared error (MSE) of an estimator $T(y)$ is $m(T(y), \beta)=$ $E(T(y)-\beta)(T(y)-\beta)^{\prime}$.
(c) What is the MSE of the OLS estimator?
(d) What is the MSE of an arbitrary linear estimator?

The MSE of an estimator $T(y)$ is bounded if there exists an $M$ such that $m(T(y), \beta)<M$ for all $\beta \in R^{k}$ where $M<\infty$. Consider an alternative optimality claim: Within the class of linear estimators with bounded MSE, the OLS estimator has minimum variance.
(e) Is this true?
(f) Offer a proof or counterexample. Interpret briefly and insightfully.
7. (Final 2005) You are a clever theorist and have a model characterizing y as a function of x by the differential equation

$$
\frac{d y(x)}{d x}=\alpha+\beta(y(x)-\gamma-\alpha x
$$

with the endpoint condition $y(0)=\gamma+1$.
(a) Solve for $y(x)$ and assume an additive error with mean zero and variance $\sigma^{2}$ (If you have forgotten how to solve these, just verify that $y(x)=\gamma+\alpha x+\exp (x \beta)$ is a solution).
(b) How would you estimate the unknown parameters $(\alpha, \beta, \gamma)$ by nonlinear squares, given a sample of observations on $y$ and $x$ ? (give the "normal equations").
(c) Assuming x is a scalar variable, find a trick to reduce the dimension of the nonlinear optimization problem to 1 .
(d) You now have a slightly more complicated model in which x is also stochastic and is modeled by the equation $x_{t}=z_{t} \delta+\varepsilon_{t}$, where $\varepsilon_{t}$ is correlated with the error term in your structural equation for $y(x) . z_{t}$ is a $k$-vector. Now how could you estimate the structural equation?
(e) Use your results for b and d to suggest a possible test for the endogeneity of $x$.

