Econ 551b Econometrics II
Problem Set 4

Prof. John Rust

Due: March 29, 1999 (Monday)

QUESTION 1: Show that the maximum likelihood estimator $\hat{\theta}$ is *regular* in the sense that if $\{x_1, \ldots, x_N\}$ are *IID* draws from $f(x|\theta_N)$ where $\theta_N = \theta^* + \delta/\sqrt{N}$, then

$$\sqrt{N}[\hat{\theta} - \theta_N] \Longrightarrow N(0, I^{-1}(\theta^*))$$

for all $\delta \in R^k$. What conditions on θ^* and $f(x|\theta)$ did you need to assume to establish this result?

QUESTION 2: Prove the information equality

$$E[\frac{\partial \ln f(x|\theta)}{\partial \theta} \frac{\partial \ln f(x|\theta)}{\partial \theta'}] = E[\frac{\partial^2 \ln f(x|\theta)}{\partial \theta \partial \theta'}]$$

under correct specification, i.e., $f(x|\theta)$ is the true data generating process for x.

QUESTION 3: Find the asymptotic distribution of the Non-linear Least Squares estimator

$$\hat{\theta} = \arg\min \frac{1}{n} \sum_{i=1}^{N} \frac{1}{2} (Y_i - f(X_i|\theta))^2$$

Where

$$Y_i = g(X_i) + \epsilon_i$$

and X_i and ϵ_i are i.i.d., and also independent of each other. Consider both the case of correct specification, i.e., $g(X_i) = f(X_i|\theta^*)$ and that of misspecification, i.e., $g(X_i) \neq f(X_i|\theta)$ for all θ.

QUESTION 4: Extract data in file `data1.asc` in the

`pub/John_Rust/courses/econ551/regression/`

directory on `gemini.econ.yale.edu` (either ftp to `gemini.econ.yale.edu` and login as "anonymous" and `cd pub/John_Rust/courses/econ551/regression` and `get data1.asc` or click on the hyperlink in the html version of this document). This data file contains $n = 1500$ *IID* observations (y_i, x_i) that I generated on the computer from the nonlinear regression

$$y = \exp\{\beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3\} + \epsilon$$

where ϵ is normally distributed with mean zero, and the independent variable is a scalar random variable x which is also normally distributed. The sorted (y, x) observations are graphed in the file `data1.eps`, also available by clicking on the hyperlink: to `http://gemini.econ.yale.edu/jrust/econ551/exams/98/ps0/data_ex1.eps`

(a) Using the data in `data1.asc` compute the maximum likelihood estimates of the parameter vector $\theta = (\beta, \sigma^2, \mu, \delta^2)$, where β is the (4×1) vector of regression coefficients, σ^2 is the variance of ϵ, and μ is the mean of the x distribution and δ^2 is its variance. Show theoretically that the asymptotic covariance between the (μ, δ^2) parameters and the (β, σ^2) parameters is zero. Is zero also the sample estimate of this covariance from your estimation algorithm?

(b) Compute White misspecification-consistent estimates of the standard errors for your parameters and compare them to the standard estimates from an estimate of the inverse of the estimated information matrix. Are there big discrepancies between these two different estimates that would lead you to be concerned about possible misspecification of your model?

(c) What happens if you try to estimate (β, σ) by simple OLS with the log-linear specification:

$$\log(y) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon \; ?$$

Compare the OLS (or MLE) estimates of this log-linear model to those you obtained in step 1. Can you come up with a theoretical argument that the probability limits for (β, σ) are the same for the two different estimation methods? If so, write down a proof, otherwise provide an argument of why the probability limits are different.

(d) Examine the estimated residuals from the nonlinear regression model in part 1 for evidence of heteroscedasticity. What kinds of statistics could you think of to provide evidence of the possibility of heteroscedasticity? Can you think of a simple way to test the hypothesis of no heteroscedasticity, i.e. homoscedasticity?

(e) One way to test for homoscedasticity is to nest the model in part 1 in a larger model that allows for heteroscedasticity and then test the null of homoscedasticity via a likelihood ratio or Wald test (topics we will cover later in Econ 551). Restimate the model in part 1 but now allow for heteroscedasticity of the following form:

$$\sigma^2(x) = \exp\{\gamma_0 + \gamma_1 x + \gamma_2 x^2\}$$

Now what are your maximum likelihood estimates of $\theta = (\beta, \gamma, \mu, \delta^2)$? Is the asymptotic covariance between the β and γ parameters zero? Why or why not? Can you reject the hypothesis of homoscedasticity via likelihood ratio or Wald test at the 5% significance level?