Cornell University
Department of Economics

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

Suggested Solutions for Midterm Exam 2008

1. First, denote the parameter estimators from the unrestricted model ($y_i = \alpha + \beta x_i + \epsilon_i$) by $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\sigma}^2$, and $\widehat{S}^2$ ($\widehat{\sigma}^2 = \frac{\widehat{\epsilon}'\widehat{\epsilon}}{N}, \widehat{S}^2 = \frac{\widehat{\epsilon}'\widehat{\epsilon}}{N-2}$). Similarly, the estimators from the restricted one ($y_i = \beta x_i + \epsilon_i$) are $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\sigma}^2$, and $\widetilde{S}^2$($\widetilde{\sigma}^2 = \frac{\widetilde{\epsilon}'\widetilde{\epsilon}}{N}, \widetilde{S}^2 = \frac{\widetilde{\epsilon}'\widetilde{\epsilon}}{N-1}$).

   (a) F-statistics:
   Note that R=(1 0) and r=0. F-statistics is same as the square of t-statistics for $\widehat{\alpha}$ since the number of restriction is one in this case.

   $$F(1, N-2) = \left(\frac{\widehat{\alpha}}{\widehat{S}\sqrt{(X'X)^{-1}_{(1,1)}}} \right)^2$$

   where $X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$, and $(X'X)^{-1}_{(1,1)}$ is (1,1)th element of $(X'X)^{-1}$ matrix, which is $\frac{\sum x_i^2}{N\sum x_i^2 - (\sum x_i)^2}$.

   (b) Wald-statistics:
   The Wald statistics is same as the F-statistics except that $\widehat{S}$ is replaced by $\widehat{\sigma}^2$, and the statistics follows $\chi^2(1)$ under the null hypothesis.

   $$Wald\ statistics = \left(\frac{\widehat{\alpha}}{\widehat{\sigma}\sqrt{(X'X)^{-1}_{(1,1)}}} \right)^2 \sim \chi^2(1).$$

   (c) Score-statistics:
   Score statistics is

   $$\frac{1}{N} S_0' i_o^{-1} S_o \sim \chi^2(1).$$

After taking derivative of log likelihood function and evaluating it at the restricted estimator, we can have

$$S_0 = \begin{bmatrix} \frac{1}{\widetilde{\sigma}^2}\sum \widetilde{\epsilon}_i \\ \frac{1}{\widetilde{\sigma}^2}\sum \widetilde{\epsilon}_i x_i \\ -\frac{N}{2}\frac{1}{\widetilde{\sigma}^2} + \sum \frac{\widetilde{\epsilon}^2}{2\widetilde{\sigma}^4} \end{bmatrix} = \begin{bmatrix} \frac{1}{\widetilde{\sigma}^2}\sum \widetilde{\epsilon}_i \\ 0 \\ 0 \end{bmatrix}.$$

The inverse of information matrix evaluated at the restricted estimator (In fact, the matrix below is 2×2 upper-left submatrix of the inverted information matrix) is

$$i_o^{-1} = N\widetilde{\sigma}^2 (X'X)^{-1}.$$

Hence, the resulting score statistics is

$$\textit{Score statistics} = \frac{1}{\widetilde{\sigma}^2}\left(\sum \widetilde{\epsilon}_i\right)^2 (X'X)^{-1}_{(1,1)} \sim \chi^2(1).$$

Also, after some manipulation, we can show that the score statistics is same as

$$\left(\frac{\widehat{\alpha}}{\widetilde{\sigma}\sqrt{(X'X)^{-1}_{(1,1)}}}\right)^2,$$

which replace $\widehat{\sigma}$ by $\widetilde{\sigma}$ in the Wald statistics formula.

(d) Likelihood ratio:

$$2(l(\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2) - l(\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\sigma}^2)),$$

where

$$\begin{aligned} l(\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2) &= -\frac{N}{2}\ln(2\pi\widehat{\sigma}^2) - \frac{1}{2\widehat{\sigma}^2}\widehat{\epsilon}'\widehat{\epsilon}, \text{ and} \\ l(\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\sigma}^2) &= -\frac{N}{2}\ln(2\pi\widetilde{\sigma}^2) - \frac{1}{2\widetilde{\sigma}^2}\widetilde{\epsilon}'\widetilde{\epsilon}. \end{aligned}$$

Using $\widehat{\sigma}^2 = \widehat{\epsilon}'\widehat{\epsilon}/N$ and $\widetilde{\sigma}^2 = \widetilde{\epsilon}'\widetilde{\epsilon}/N$, we have

$$2(l(\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2) - l(\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\sigma}^2)) = N\ln\left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\right),$$

which follows a chi-square distribution with degree of freedom 1.

2.

(a) Note that error term ϵ_i is i.i.d. and uncorrelated with the explanatory variable, x_i for each i. Clearly, the OLS estimator is unbiased, and the BLUE (Best Linear Unbiased Estimator).

(b) Find 2 differerent y_i values for $x_i = 1$, call the largest $y_h(1)$ (and note that the lower is $y_l(1) = y_h(1) - 3$); similarly obtain $y_h(0)$. Then, $\alpha = y_h(0) - 2$ and $\beta = y_h(1) - 2 - \alpha$. Since these α and β are true values, they are unbiased, consistent. Furthermore, their variances are all zero.

(c) Even though the OLS is the best unbiased linear estimator, it may give the very unlikely estimator. It's because the error term has a very special structure, where it can take a value from the finite set. Note that the OLS is the best among the "linear" unbiased estimator. However, the method used in the part (b) is not a linear method.

3.

(a) By taking log, we can have a following model:

$$\begin{aligned} \log y_i &= \log\beta + \log x_i + \log\epsilon_i, \quad i = 1, \ldots, N. \\ \Longrightarrow & \log\frac{y_i}{x_i} = \log\beta + \log\epsilon_i. \end{aligned}$$

The last is the model where the constant is the only explanatory variable. Note that $E(\log\epsilon_i) \neq 0$. We can estimate the model by the usual OLS, obtain the estimator for $\log\beta + E(\log\epsilon_i)$ (call this $\widehat{\gamma}$), and again obtain a consistent estimator for β by transformation, $\exp(\widehat{\gamma} - E(\log\epsilon_i))$. We know that $E(\log\epsilon_i)$ is approximately -0.577 (Euler's number) as given in the above.

(b) Distribution of Y for each observation is as follows:

$$\Pr(Y_i \le y_i) = \Pr(x_i\beta\epsilon_i \le y_i) = \Pr(\epsilon \le \frac{y_i}{x_i\beta}) = 1 - e^{-\frac{y_i}{x_i\beta}}.$$

Since the density is the first derivative of CDF, the density for y is

$$f(y_i) = \frac{1}{x_i\beta}e^{-\frac{y_i}{x_i\beta}}.$$

Let us construct the loglikelihood function:

$$l(\beta) = \sum_{i=1}^{N}\log f(y_i) = \sum_{i=1}^{N}\left(-\ln x_i\beta - \frac{y_i}{\beta x_i}\right).$$

The score function is

$$\frac{\partial l(\beta)}{\partial\beta} = \sum_{i=1}^{N}\left(-\frac{1}{\beta} + \frac{y_i}{\beta^2 x_i}\right).$$

The MLE for β that makes the above zero is $\widehat{\beta}_{ML} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{x_i}$

The expected hessian (for individual observation), j_0, is

$$\begin{aligned}
E\left(\frac{1}{N}\frac{\partial^2 l(\beta)}{\partial\beta^2}\right) &= E\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\beta^2} - 2\frac{y_i}{\beta^3 x_i}\right)\right) \\
&= E\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\beta^2} - 2\frac{x_i\beta\epsilon_i}{\beta^3 x_i}\right)\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\beta^2} - 2\frac{1}{\beta^2}\right) \\
&= -\frac{1}{\beta^2}
\end{aligned}$$

Note that $E(\epsilon_i) = 1$. Therefore, the information matrix i_0 is $\frac{1}{\beta_0^2}$ by $i_0 = -j_0$.
The resulting asymptotic distribution of $\widehat{\beta}$ is as follows:

$$\sqrt{N}(\widehat{\beta} - \beta_0) \xrightarrow{d} N(0, \beta_0^2).$$