BOSTON COLLEGE
Department of Economics
EC771: Econometrics
Spring 2004
Prof. Baum, Ms. Uysal

PROBLEM SET 1: DUE THURSDAY 5 FEBRUARY 2004 AT CLASSTIME

Note: you may use Stata in producing the answers to any of the questions below if you wish. If you use Stata (or another computer program) hand in your output.

1. Are the following quadratic forms positive for all values of $\mathbf{x}$?
(a) $y = x_1^2 - 28x_1x_2 + (11x_2)^2$.
(b) $y = 5x_1^2 + x_2^2 + 7x_3^2 + 4x_1x_2 + 6x_1x_3 + 8x_2x_3$.

2. Compute the characteristic roots of

$$A = \begin{pmatrix} 2 & 4 & 3 \\ 4 & 8 & 6 \\ 3 & 6 & 5 \end{pmatrix}.$$

3. If x has a normal distribution with mean 1 and standard deviation 3, what are the following?
(a) $Prob\,[|x| > 2]$.
(b) $Prob\,[x > -1 |\, x < 1.5]$.

4. If x has a normal distribution with mean μ and standard deviation σ, what is the probability distribution of $y = e^x$?

5. The following sample is drawn from a normal distribution with mean μ and standard deviation σ:

$$\mathbf{x} = 1.3, 2.1, 0.4, 1.3, 0.5, 0.2, 1.8, 2.5, 1.9, 3.2.$$

Using the data, test the following hypotheses:

$$\mu \geq 2.0,$$

$$\mu \leq 0.7,$$

$$\sigma^2 = 0.5.$$

Using a likelihood ratio test, test the hypothesis

$$\mu = 1.8, \sigma^2 = 0.8.$$

6. A common method of simulating random draws from the standard normal distribution is to compute the sum of 12 draws from the uniform [0,1] distribution and subtract 6. Can you justify this procedure?

7. The random variable x has a continuous distribution $f(x)$ and cumulative distribution function $F(x)$. What is the probability distribution of the sample maximum? [Hint: In a random sample of n observations, $x_1, x_2, \ldots, x_n$, if z is the maximum, then every observation in the sample is less than or equal to z. Use the cdf.]

8. *Testing for normality.* One method that has been suggested for testing whether the distribution underlying a sample is normal is to refer the statistic

$$L = n[skewness^2/6 + (kurtosis - 3)^2/24]$$

to the chi-squared distribution with two degrees of freedom. Using the data in Exercise 5, carry out the test.

9. *Mixture distribution.* Suppose that the joint distribution of the two random variables x and y is

$$f(x, y) = \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}, \beta, \theta > 0, y \geq 0, x = 0, 1, 2, \ldots$$

Find the maximum likelihood estimators of β and θ and their asymptotic joint distribution.

10. *Empirical exercise.*

Using the "canned" dataset `discrim`, which you may access from within Stata with the command

```
.use http://fmwww.bc.edu/ec-p/data/wooldridge/discrim
```

(a) How many of the observations are from Pennsylvania? (hint: `describe` may be useful)
(b) What is the average starting wage (first wave)? What are the min and max of this variable?
(c) Test the hypothesis that average incomes in NJ and PA are equal. (hint: `help ttest`)
(d) Test the hypothesis that the average price of an entree (first wave) was $1.39.
(e) Write a Stata program, using the `ml` syntax, which will estimate the parameters of a k-variable linear regression model with a constant term, including the σ^2 parameter, via maximum likelihood. Use your program to estimate the model:

$$pfries_i = \beta_0 + \beta_1 income_i + \beta_2 prpblck_i + \epsilon_i$$

over the whole sample, and
(f) only over the New Jersey observations (those for which `state` equals 1).
(g) Use Stata's `regress` command to estimate the same linear regression, and use `bootstrap` to generate bootstrap standard errors for the $\hat{\beta}$ parameters. Discuss how the bootstrap confidence intervals for `income` and `prpblk` compare with the conventional confidence intervals computed by `regress`.

Turn in a printout of your program illustrating the results of each estimation.