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

SOLUTION KEY FOR PROBLEM SET 1

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) $y = 5x_1^2 + x_2^2 + 7x_3^2 + 4x_1x_2 + 6x_1x_3 + 8x_2x_3$.

The first may be written $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -14 \\ -14 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. The determinant of the matrix is 11-196=-85, so it is not positive definite. Thus, the first quadratic form need not be positive. The second uses the matrix $\begin{bmatrix} 5 & 2 & 3 \\ 2 & 1 & 4 \\ 3 & 4 & 7 \end{bmatrix}$. There are several ways to check the definiteness of a matrix. One way is to check the signs of the principal minors, which must be positive. The first two are 5 and 5(1)-2(2)=1, but the third, determinant is -34. Therefore, the matrix is not positive definite. Its three characteristic roots are 11.1, 2.9, and -1. It follows, therefore, that there are values of x_1, x_2, and x_3 for which the quadratic form is negative.

2. Compute the characteristic roots of

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

The roots are determined by $|\mathbf{A}\text{-}\lambda\mathbf{I}| = 0$. For the matrix above, this is

$$\begin{aligned} |\mathbf{A}\text{-}\lambda\mathbf{I}| &= (2-\lambda)(8-\lambda)(5-\lambda) + 72 + 72 - 9(8-\lambda) - 36(2-\lambda) - 16(5-\lambda) \\ &= -\lambda^3 + 15\lambda^2 - 5\lambda = -\lambda(\lambda^2 - 15\lambda + 5) = 0. \end{aligned}$$

One solution is obviously zero.(This might have been apparent. The second column of the matrix is twice the first, so it has rank no more than two, and therefore no more than two nonzero roots.) The other two roots are $(15 \pm \sqrt{205})/2$=.341 and 4.659.

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

Using the normal table,

$$(a) Prob\,[|x| > 2] \quad = \quad 1 - Prob\,[|x| \leq 2]$$

$$
\begin{aligned}
&= 1 - Prob\,[-2 \leq x \leq 2] \\
&= 1 - Prob\,[(-2-1)/3 \leq z \leq (2-1)/3] \\
&= 1 - [F(1/3) - F(-1)] = 1 - .6306 + .1587 = .5281.
\end{aligned}
$$

$$
\begin{aligned}
(b) Prob\,[x > -1 |\, x < 1.5] &= Prob\,[-1 < x < 1.5]\,/Prob\,[x < 1.5] \\
Prob\,[-1 < x < 1.5] &= Prob\,[(-1-1)/3 < z < (1.5-1)3] \\
&= Prob\,[z < 1/6] - Prob\,[z < -2/3] \\
&= .5662 - .2525 = .3137.
\end{aligned}
$$

The conditional probability is .3137/.5662=.5540.

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

If $y = e^x$, then $x = lny$ and the Jacobian is $dx/dy = 1/y$. Making the substitution,

$$f(y) = \frac{1}{\sigma y \sqrt{2\pi}} e^{\frac{-1}{2}[(lny-\mu)/\sigma]^2}$$

This is the density of the lognormal distribution.

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:
(a) $\mu \geq$2.0,
(b) $\mu \leq$0.7,
(c) $\sigma^2 = 0.5$.
(d)Using a likelihood ratio test, test the hypothesis

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

Mean, variance and standard deviation of the sample are as follows,

$$
\begin{aligned}
\bar{x} &= \frac{\sum_{i=1}^{n} x_i}{n} = 1.52, \\
s^2 &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = .9418, \\
s &= .97
\end{aligned}
$$

(a) We would reject the hypothesis if 1.52 is too small relative to the hypothesized value of 2. Since the data are sampled from a normal distribution, we may use a t-test to test the hypothesis. The t-ratio is

$$t[9] = (1.52 - 2)/[.97/\sqrt{10}] = -1.472.$$

The 95% critical value from the t-distribution for a one tailed test is -1.833. Therefore, we would not reject the hypothesis at a significance level of 95%.
(b) We would reject the hypothesis if 1.52 is excessively large relative to the hypothesized mean of .7. The t-ratio is

$$t[9] = (1.52 - .7)/[.97/\sqrt{10}] = 2.673.$$

Using the same critical value as in the previous problem, we would reject this hypothesis.
(c) The statistic $(n-1)s^2/\sigma^2$ is distributed as χ^2 with 9 degrees of freedom. This is 9(.94)/.5=16.920. The 95% critical values from the chi-squared table for a two tailed test are 2.70 and 19.02. Thus we would not reject the hypothesis.
(d) The log-likelihood for a sample from a normal distribution is

$$\ln L = -(n/2)\ln(2\pi) - (n/2)\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

The sample values are $\hat{\mu} = \bar{x} = 1.52$, $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n} = .8476$.
The maximized log-likelihood for the sample is -13.363. A useful shortcut for computing the log-likelihood at the hypothesized value is $\sum_{i=1}^{n}(x_i - \mu)^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2$. For the hypothesized value of $\mu = 1.8$, this is $\sum_{i=1}^{n}(x_i - 1.8)^2 = 9.26$. The log-likelihood is $-5\ln(2\pi) - 5\ln .8 - (1/1.6)9.26 = -13.861$. The likelihood ratio statistic is -2($\ln L_r - \ln L_u$)=.996. The critical value for a chi-squared with 2 degrees of freedom is 5.99, so we would not reject the hypothesis.

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?

The uniform distribution has mean 2 and variance 1/12. Therefore, the statistic $12(\bar{x} - 1/2) = \sum_{i=1}^{12} x_i - 6$ is equivalent to $z = \sqrt{n}(\bar{x} - \mu)/\sigma$. As $n \to \infty$, this converges to a standard normal variable. Experience suggests that a sample of 12 is large enough to approximate this result. However, more recently developed random number generators usually use different procedures based on the truncation error which occurs in representing real numbers in a digital computer.

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.]

If z is the maximum, then every sample observation is less than or equal to z. The probability of this is Prob[$x_1 \leq z, x_2 \leq z, \ldots, x_n \leq z$]=$F(z)F(z)\ldots F(z) = [F(z)]^n$. The density is the derivative, $n[F(z)]^{n-1}f(z)$.

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

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

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

The skewness coefficient is .14192 and the kurtosis is 1.8447. (These are the third and fourth moments divided by the third and fourth power of the sample standard deviation.) Inserting these in the expression in the question produces L= $10\{.14192^2/6 + (1.8447 - 3)^2/24\} = .59$. The critical value from the chi-squared distribution with 2 degrees of freedom (95%) is 5.99. Thus, the hypothesis of normality cannot be rejected.

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.

The log-likelihood is
$\ln L = n \ln \theta - (\beta + \theta) \sum_{i=1}^n y_i + \ln \beta \sum_{i=1}^n x_i + \sum_{i=1}^n x_i log y_i - \sum_{i=1}^n log(x_i!)$
The first and second derivatives are

$$\begin{aligned}
\partial L/\partial\theta &= n/\theta - \sum_{i=1}^n y_i \\
\partial \ln L/\partial\beta &= -\sum_{i=1}^n y_i + \sum_{i=1}^n x_i/\beta \\
\partial^2 \ln L/\partial\theta^2 &= -n/\theta^2 \\
\partial^2 \ln L/\partial\theta^2 &= -\sum_{i=1}^n x_i/\beta^2 \\
\partial^2 \ln L/\partial\beta\partial\theta &= 0.
\end{aligned}$$

Therefore, the maximum likelihood estimators are $\hat{\theta} = 1/\bar{y}$ and $\hat{\beta} = \bar{x}/\bar{y}$ and the asymptotic covariance matrix is the inverse of E $\begin{bmatrix} n/\theta^2 & 0 \\ 0 & \sum_{i=1}^n x_i/\beta^2 \end{bmatrix}$. In order to complete the derivation, we will require the expected value of $\sum_{i=1}^n x_i = nE[x_i]$. In order to obtain E$[x_i]$, it is necessary to obtain the marginal distribution of x_i, which is $f(x) = \int_0^\infty \theta e^{-(\beta+\theta)y}(\beta y)^x/x!\, dy = \beta^x(\theta/x!) \int_0^\infty e^{-(\beta+\theta)y} y^x\, dy$. This is $\beta^x(\theta/x!)$ times a gamma integral. This is $f(x) = \beta^x(\theta/x!)[\Gamma(x+1)]/(\beta+\theta)^{x+1}$. But, $\Gamma(x+1) = x!$, so the expression reduces to

$$f(x) = [\theta/(\beta + \theta)][\beta/(\beta + \theta)]^x.$$

Thus, x has a geometric distribution with parameter $\pi = \theta/(\beta+\theta)$. (This is the distribution of the number of tries until the first success of independent trials each with success probability 1-π.) Finally, we require the expected value of x_i, which is

$$E[x] = [\theta/(\beta+\theta)] \sum_{x=0}^{\infty} x[\beta/(\beta+\theta)]^x] = \beta/\theta.$$

Then, the required asymptotic covariance matrix is

$$\begin{bmatrix} n/\theta^2 & 0 \\ 0 & n(\beta/\theta)/\beta^2 \end{bmatrix}^{-1} = \begin{bmatrix} n/\theta^2 & 0 \\ 0 & (\beta\theta)/n \end{bmatrix}.$$

10. *Empirical exercise.*

Using the "canned" dataset `discrim`, which you may access 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.