ECON 546 - Spring 2010
Solution for Assignment 1

Q. 1.

(a) $p(y_1, ---, y_n \mid \theta) = \left(\frac{1}{\theta^n} \exp^{-1/\theta \sum y_i}\right)(1)$

So, $\sum_i y_i$ is sufficient for θ, & hence $\bar{y} = 1/n \sum_i y_i$ is also sufficient.

(b) $\phi_y(t) = E(e^{ity}) = \int_0^\infty e^{ity} \frac{1}{\theta} e^{-y/\theta} dy$

$= \frac{1}{\theta} \int_0^\infty e^{-y(\frac{1}{\theta} - it)} dy$

$= \frac{1}{\theta} \left[\frac{1}{\theta} - it\right]^{-1} (-1) \left[e^{-y(\frac{1}{\theta} - it)}\right]_0^\infty$

$= -\frac{1}{\theta} \left[\frac{1}{\theta} - it\right]^{-1} [0 - 1]$

$= \frac{1}{\theta(1/\theta - it)} = (1 - it\theta)^{-1}.$

(c) To get the raw moments we could either repeatedly differentiate $\phi_y(t)$ w.r.t. "t", & then evaluate at $t = 0$, or use the Taylor series expansion. Here, the latter will converge only if $|it\theta| < 1$. This will hold as $t \to 0$, but not otherwise. You will get the correct answer if you use this approach. However, to be safe, I'm going to differentiate the c.f. to get the moments:

$\phi_y(t) = (1 - it\theta)^{-1}$

$\phi_y'(t) = i\theta(1-it\theta)^{-2}$

$\phi_y''(t) = 2i^2\theta^2(1-it\theta)^{-3}$

So, $\mu_1' = \frac{1}{i} \lim_{t \to 0} [\phi_y'(t)] = \theta.$

$\mu_2' = \frac{1}{i^2} \lim_{t \to 0} [\phi_y''(t)] = 2\theta^2.$

So, $\begin{cases} E(Y) = \mu_1' = \theta \\ var.(Y) = \mu_2' - (\mu_1')^2 = (2\theta^2 - \theta^2) = \theta^2. \end{cases}$

(d) $\mu_3' = \frac{1}{i^3} \lim_{t \to \infty} [\phi_y^{(3)}(t)] = 6\theta^3$

$\mu_4' = \frac{1}{i^4} \lim_{t \to \infty} [\phi_y^{(4)}(t)] = 24\theta^4.$

Skew $= \mu_3 / (\mu_2)^{3/2}$

& $\mu_3 = \mu_3' + 2(\mu_1')^3 - 3\mu_1'\mu_2'$
$= 2\theta^3$

$\mu_2 = var(Y) = \theta^2$

& So Skew$(Y) = 2\theta^3/(\theta^2)^{3/2} = 2.$

Kurtosis$(Y) = \mu_4 (\mu_2)^2 = \mu_4/\theta^4$

& $\mu^4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$
$= 9\theta^4$

So, $\text{Skew}(Y) = (9\theta^4/\theta^4) = 9$

& Excess kurtosis $= (9 - 3) = 6$.

(e) $L(\theta|\underset{\sim}{y}) = p(\underset{\sim}{y}|\theta) = \theta^{-n} \exp\left[-\frac{\sum y_i}{\theta}\right]$

& $\log L = -n \log\theta - \frac{1}{\theta} n\bar{y}$.

$$\frac{\partial \log L}{\partial \theta} = \frac{-n}{\theta} + \frac{n\bar{y}}{\theta^2} = 0$$

$$\Rightarrow \tilde{\theta} = \bar{y}.$$

$$\left(\partial^2 \log L / \partial\theta^2\right) = \frac{n}{2\theta^2} - \frac{2n\bar{y}}{\theta^3}$$

When $\theta = \tilde{\theta}$, $\frac{\partial^2 \log L}{\partial \theta^2} = \left[\frac{n}{2\bar{y}^2} - \frac{2n\bar{y}}{\bar{y}^3}\right]$

$$= \frac{n}{2\bar{y}^2} - \frac{2n}{\bar{y}^2} < 0.$$

Hence, we have <u>maximized</u> $\log L$.

(f) We have shown that $\bar{y}$ is sufficient, & $\tilde{\theta} = \bar{y}$, so $\tilde{\theta}$ is a (trivial) function of a sufficient statistic. (It is also unique)

(g) By the invariance of MLE — ☺

$\widetilde{\text{Skew}} = 2$; $\widetilde{\text{Ex. kurt}} = 6$

because neither measure depends on θ!

Q.2. (a) Using independence –

$$L(\theta|\underset{\sim}{y}) = p(\underset{\sim}{y}|\theta) = \prod_i p(y_i|\theta)$$

$$= \prod_{i=1}^{n} y_i \exp[-y_i^2/(2\theta^2)]/\theta^2$$

$$= (\prod_i y_i) \exp[-\frac{1}{2\theta^2}\sum_i y_i^2]\,\theta^{-2n}$$

$$\Rightarrow \log L = \sum_i \log y_i - \frac{1}{2\theta^2}\sum_i y_i^2 - 2n\log\theta$$

$$(\partial \log L/\partial\theta) = (-\tfrac{1}{2})(-2)\theta^{-3}\sum_i y_i^2 - (2n/\theta)$$

$$= (\sum y_i^2/\theta^3) - (2n/\theta)$$

$= 0$ for max.

$$\Rightarrow 2n\tilde{\theta}^3 = \tilde{\theta}\sum_i y_i^2$$

$$\Rightarrow \tilde{\theta} = +\sqrt{\frac{1}{2n}\sum_i y_i^2}$$

(must be $+\sqrt{\ }$ because $y_i > 0$, & $\theta\sqrt{\pi/2} = E(y_i)$)

Check S.O.C. :

$$(\partial^2 \log L/\partial\theta^2) = (-3\sum_i y_i^2/\theta^4) - (2n)(-1)\theta^{-2}$$

$$= -3\sum y_i^2/\theta^4 + 2n/\theta^2$$

When $\theta = \tilde{\theta}$, this becomes

$$-3\sum_i y_i^2 \Big/ \left(\frac{1}{2n}\sum_i y_i^2\right)^2 + 2\Big/\left(\frac{1}{2n}\sum y_i^2\right)$$

$$= (-6n^2/\sum y_i^2) + (4n^2/\sum y_i^2) = (-2n^2/\sum y_i^2) < 0.$$

(MAX)

(b) Using the Invariance of MLE's —

the MLE of $\theta\sqrt{\pi/2}$ is $\tilde{\theta}\sqrt{\pi/2}$ or

$$\left(\sqrt{\pi/2}\right)\,{}^{+}\!\sqrt{\sum y_i^2/2n} = {}^{+}\!\sqrt{\pi\sum y_i^2/4n}.$$

$$V(y_i) = E(y_i^2) - [E(y_i)]^2$$
$$= 2\theta^2 - (\theta\sqrt{\pi/2})^2 = 2\theta^2 - \theta^2\pi/2$$
$$= \theta^2(2 - \pi/2)$$

So, $\tilde{V}(y_i) = \tilde{\theta}^2(2-\pi/2) = \left(\sum_i y_i^2/2n\right)(2-\pi/2)$
(Re-arrange & simplify if you wish.)

(c) To obtain the mode —

$$\partial p(y_i)/\partial y_i = \frac{1}{\theta^2}\left[\exp\left[-y_i^2/(2\theta^2)\right] + y_i\exp\left[-y_i^2/(2\theta^2)\right]\cdot(-2y_i/2\theta^2)\right]$$

$$= \frac{1}{\theta^2}\exp\left[-y_i^2/2\theta^2\right]\left\{1 - y_i^2/\theta^2\right\} = 0$$

$$\Rightarrow 1 - y_i^2/\theta^2 = 0, \text{ or } y_i = \theta.$$

So, the mode is at $y_i = \theta$, & the MLE of the mode is $\tilde{\theta}$, given in (a) above.

(6)

Question 3

(a) $c(1) = \log(\gamma)$; $c(2) = \nu$; $c(3) = \rho$; $c(4) = \delta$; starting values were 1.0, 0.8, 0.6, 0.4, respectively.

Dependent Variable: LOG(Q)
Method: Least Squares
Date: 01/30/07 Time: 09:58
Sample: 1 25
Included observations: 25
Convergence achieved after 4 iterations
LOG(Q)=C(1) - (C(2)/C(3))*LOG(C(4)*K^(-C(3)) + (1-C(4))*L^ (-C(3))
)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	2.432348	0.112039	21.70984	0.0000
C(2)	0.822206	0.050101	16.41107	0.0000
C(3)	0.604258	0.259669	2.327036	0.0300
C(4)	0.406370	0.127327	3.191540	0.0044

R-squared	0.957110	Mean dependent var	4.375734
Adjusted R-squared	0.950983	S.D. dependent var	0.365243
S.E. of regression	0.080864	Akaike info criterion	-2.046453
Sum squared resid	0.137318	Schwarz criterion	-1.851433
Log likelihood	**29.58067**	Durbin-Watson stat	2.483055

(b) With starting values: 1, 1, 1, 1 you get the same results as in (a). With 0.1, 0.1, 0.1, 0.1 you get:

Dependent Variable: LOG(Q)
Method: Least Squares
Date: 01/30/07 Time: 10:06
Sample: 1 25
Included observations: 25
Convergence achieved after 117 iterations
WARNING: Singular covariance - coefficients are not unique
LOG(Q)=C(1) - (C(2)/C(3))*LOG(C(4)*K^(-C(3)) + (1-C(4))*L^ (-C(3))
)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	3.989236	NA	NA	NA
C(2)	0.098080	NA	NA	NA
C(3)	-36.26204	NA	NA	NA
C(4)	-219.3733	NA	NA	NA

R-squared	0.039186	Mean dependent var	4.375734
Adjusted R-squared	-0.098073	S.D. dependent var	0.365243
S.E. of regression	0.382734	Akaike info criterion	1.062695
Sum squared resid	3.076195	Schwarz criterion	1.257715
Log likelihood	**-9.283682**	Durbin-Watson stat	1.644684

Gradients of the objective function at estimated
parameters
Equation: EQ01
Method: Least Squares
Specification: LOG(Q)=C(1) - (C(2)/C(3))*LOG(C(4)*K^(
-C(3)) + (1-C(4))*L^ (-C(3)))
Computed using analytic derivatives

Coefficient	Sum	Mean	Newton Dir.
C(1)	-8.57E-08	-3.43E-09	2.46E-05
C(2)	3.49E-07	1.39E-08	0.000186
C(3)	2.45E-12	9.79E-14	2415.882
C(4)	1.05E-12	4.20E-14	79275.40

Clearly, we have not obtained the global maximum of the likelihood function – compare the maximized log-likelihood values in cases (a) and (b). In (b), the standard errors are not available because the estimated covariance matrix is singular – we have reached a saddlepoint. Remember that zero gradients occur at a max., a min., or a saddlepoint (a point of inflexion). Some care needs to be taken with the starting values.

(c)
Gradient summary for original model in part (a)

Gradients of the objective function at estimated
parameters
Equation: EQ01
Method: Least Squares
Specification: LOG(Q)=C(1) - (C(2)/C(3))*LOG(C(4)*K^(
-C(3)) + (1-C(4))*L^ (-C(3)))
Computed using analytic derivatives

Coefficient	Sum	Mean	Newton Dir.
C(1)	1.07E-11	4.29E-13	-9.34E-12
C(2)	1.22E-10	4.89E-12	-4.93E-11
C(3)	-2.57E-08	-1.03E-09	-2.51E-10
C(4)	-2.12E-08	-8.47E-10	6.31E-12

The gradients are essentially zero in all directions, and for all observation values. It seems that we have located a maximum – no guarantee, still, that it is the **global maximum**.

(d)
We need to test if $c(3) = 0$. The "t-statistic" will be asymptotically std. normal under the null. The statistic's value is 2.327 – so we reject the null (against a 2-sided alternative) at any reasonable significance level. We would come to the same conclusion if we used the Wald test – the asymptotic Chi-Square statistic is 5.415, and the p-value is 0.02.

(e)
Test if [1 / (1+c(3)] = 0.5. This is a non-linear restriction on the parameter, and the Wald test is **not** invariant to the form of the non-linearity:

Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	1.494416	(1, 21)	0.2351
Chi-square	**1.494416**	**1**	**0.2215**

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-0.5 + 1/(1 + C(3))	0.123341	0.100895

Delta method computed using analytic derivatives.

We cannot reject the null hypothesis (p-value = 0.22)

Now test if 0.5*(1+c(3)) = 1:

Test Statistic	Value	df	Probability
F-statistic	2.322644	(1, 21)	0.1424
Chi-square	**2.322644**	**1**	**0.1275**

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-1 + 0.5*(1 + C(3))	-0.197871	0.129835

Restrictions are linear in coefficients.

We cannot reject the null hypothesis (p-value = 0.13) - same conclusion.

Some people asked why we don't get the same MLE results if we set up the log-likelihood equation using the @LOG object. In fact we **do** get the same answer if we do it properly – see below:

```
@logl LL1
eps = LOG(Q)-( C(1) - (C(2)/C(3))*LOG(C(4)*K^(-C(3))  +  (1-C(4))*L^ (-C(3))) )
ll1 = -log(c(5))-(eps^2)/(2*c(5)^2)-0.5*log(2*3.14159)
```

Note that c(5) is the error's standard deviation – maybe you called it c(3) as in the lab. material?

9

Then we get:

LogL: LOGL01
Method: Maximum Likelihood (BHHH)
Date: 01/30/07 Time: 10:33
Sample: 1 25
Included observations: 25
Evaluation order: By observation
Convergence achieved after 1 iteration

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	2.432346	0.142782	17.03537	0.0000
C(2)	0.822206	0.049604	16.57539	0.0000
C(3)	0.604265	0.300171	2.013067	0.0441
C(4)	0.406367	0.154433	2.631345	0.0085
C(5)	0.074113	0.016251	4.560513	0.0000

Log likelihood	29.58068	Akaike info criterion	-1.966454
Avg. log likelihood	1.183227	Schwarz criterion	-1.722679
Number of Coefs.	5	Hannan-Quinn criter.	-1.898841

This matches the results in part (a).

Question 4

(a) If you look at the output in Q3 (a) you will see that the std. error of estimate was 0.0808. So, rather than setting the errors' variance to be unity, we are scaling it here in the code to be 0.08, so it is comparable with that in the earlier application. After this in the code, we are generating an artificial series for the dependent variable, log(Q), and then estimating the non-linear model by OLS (which we showed earlier in Question 2, is equivalent to MLE given that we are assigning the errors to be normally distributed).

(b) The % bias that is being calculated in the original code is for the coefficient c(1), which is log (γ). To get the % bias for the MLE of γ itself, we need to alter part of the code, as follows:

```
mle(!i)=exp(c(1))
next
smpl 1 1000
scalar bias=@mean(mle)-exp(b1)
scalar percentbias=100*bias/exp(b1)
```

The % bias of the MLE of γ is 0.4903% (compared with -0.058% for the MLE of log(γ)).

Altering the code to:

```
mle(!i)=c(2)
next
smpl 1 1000
scalar bias=@mean(mle)-b2
scalar percentbias=100*bias/b2
```

the % bias of the MLE of ν is 0.364%.

Altering the code to:

```
mle(!i)=c(3)
next
smpl 1 1000
scalar bias=@mean(mle)-b3
scalar percentbias=100*bias/b3
```

the % bias of the MLE of ρ is 0.459%.

Altering the code to:

```
mle(!i)=c(4)
next
smpl 1 1000
scalar bias=@mean(mle)-b4
scalar percentbias=100*bias/b4
```

the % bias of the MLE of δ is 1.884%.

The substitution elasticity is $\sigma = 1 / (1+\rho)$, so altering the code to:

```
mle(!i)=1/(1+c(3))
next
smpl 1 1000
scalar bias=@mean(mle)-(1/(1+b3))
scalar percentbias=100*bias/(1/(1+b3))
```

and exploiting the invariance of MLE's, the % bias of the MLE of σ is 2.501%.

(c) We now need to use the latest code, which should look like this:

```
' MONTE CARLO EXPERIMENT TO VERIFY THE BIAS OF THE MLE ESTIMATOR IN A NON-
LINEAR MODEL
rndseed 123456
smpl 1 25
scalar b1=2.43
scalar b2=0.82
scalar b3=0.60
scalar b4=0.41
scalar subelast=1/(1+b3)
scalar nrep=1000
vector(nrep) mle
for !i=1 to nrep
' NOTE THAT THE ERROR'S STD. DEVIATION = 0.08
series eps=0.08*@rnorm
series y=b1-(b2/b3)*log(b4*k^(-b3)+(1-b4)*l^(-b3))+eps
equation eq1.ls y=c(1)-(c(2)/c(3))*log(c(4)*k^(-c(3))+(1-c(4))*l^ (-c(3)) )
' CALCULATE AND STORE THE SUBSTITUTION ELASTICITY
mle(!i)=1/(1+c(3))
next
smpl 1 1000
scalar bias=@mean(mle)-(1/(1+b3))
scalar percentbias=100*bias/(1/(1+b3))
```

We look at a range of different values for the parameter b3 – that is, we alter the row in bold a third of the way down the code. **In each case the initial values were set to the values suggested in the question.** I have used b3 = -.0.5; -0.25, 0.25, 0.5, 0.75, 1, 1.25, 1.5. The plot is below: