University of Victoria **Department of Economics**

ECON 546: Themes in Econometrics
Lab. Exercise 2

Introduction

The purpose of this lab. exercise is to show you how to use EViews to estimate the parameters of a regression model by Maximum Likelihood, when the model is of some non-standard type. Specifically, you will learn how to estimate models of types that are not "built in" as a standard option in EViews. This will involve setting up the log-likelihood function for the model, based on the assumption of independent observations; and then maximizing this function numerically with respect to the unknown parameters.

First, to introduce the concepts and commands that are involved, we'll consider the standard linear multiple regression model with normal errors, for which we know that the MLE of the coefficient vector is just the same as the OLS estimator. This will give us a "bench-mark" against which to check our understanding of what is going on. Then we can move on to some more general models. *Only part of this handout will be covered in the lab. class itself, which is why there is plenty of EViews output embedded in this document so that you can explore the rest of the material by yourself.*

Part 1

So, suppose that we have a linear multiple regression model, satisfying all of the usual assumptions:

$$y = X\beta + \varepsilon \quad ; \qquad \varepsilon \sim N[0, \sigma^2 I_n]$$

where the regressors are non-random. The MLE for β is $\tilde{\beta} = (X'X)^{-1}X'y$, and the MLE for σ^2 is $\tilde{\sigma}^2 = (y - X\tilde{\beta})'(y - X\tilde{\beta})/n$.

(a) Open the EViews workfile, **S:\Social Sciencces\Economics\ECON546\lab2.wf1**.
(b) Estimate an OLS regression model with Y as the dependent variable, and an intercept and X as the regressors. Save the results as EQ01.

So, our simple model is

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i \quad ; \ \varepsilon_i \sim iid\ N[0, \sigma^2] \quad ; i = 1, 2, 3, \ldots, n$$

and the marginal data density for the i^{th} observation is

$$p(y_i \mid \beta_1, \beta_2, \sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)\exp\{-\frac{1}{2\sigma^2}(y_i - \beta_1 - \beta_2 x_i)^2\}\,. \qquad (1)$$

Given the independence of the data, to get the *joint* data density, and hence the Likelihood Function, we need to *multiply* each of the expressions of the form (1) together, for all '*i*':

$$L(\beta_1, \beta_2, \sigma \mid y) = p(y \mid \beta_1, \beta_2, \sigma) = \prod_{i=1}^{n} p(y_i \mid \beta_1, \beta_2, \sigma) .$$

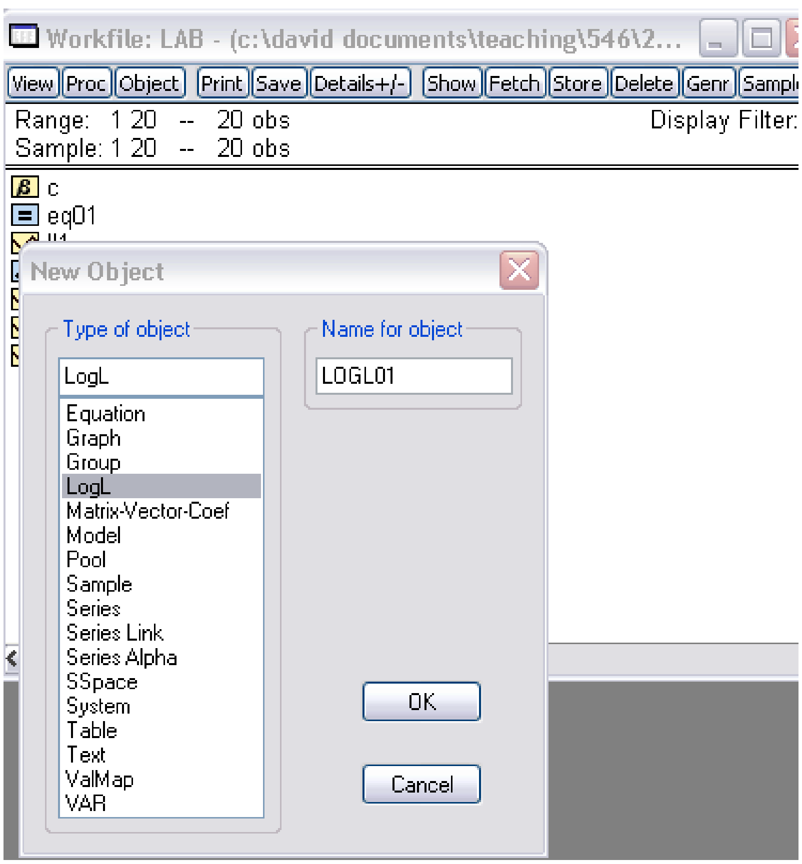
So, to get the Log-Likelihood Function, we need to *add* the logarithms of the marginal data densities:

$$\begin{aligned} &Log[L(\beta_1, \beta_2, \sigma \mid y)] = Log[p(y \mid \beta_1, \beta_2, \sigma)] \\ &= Log[\prod_{i=1}^{n} p(y_i \mid \beta_1, \beta_2, \sigma)] = \sum_{i=1}^{n} Log[p(y_i \mid \beta_1, \beta_2, \sigma)] \end{aligned} \tag{2}$$

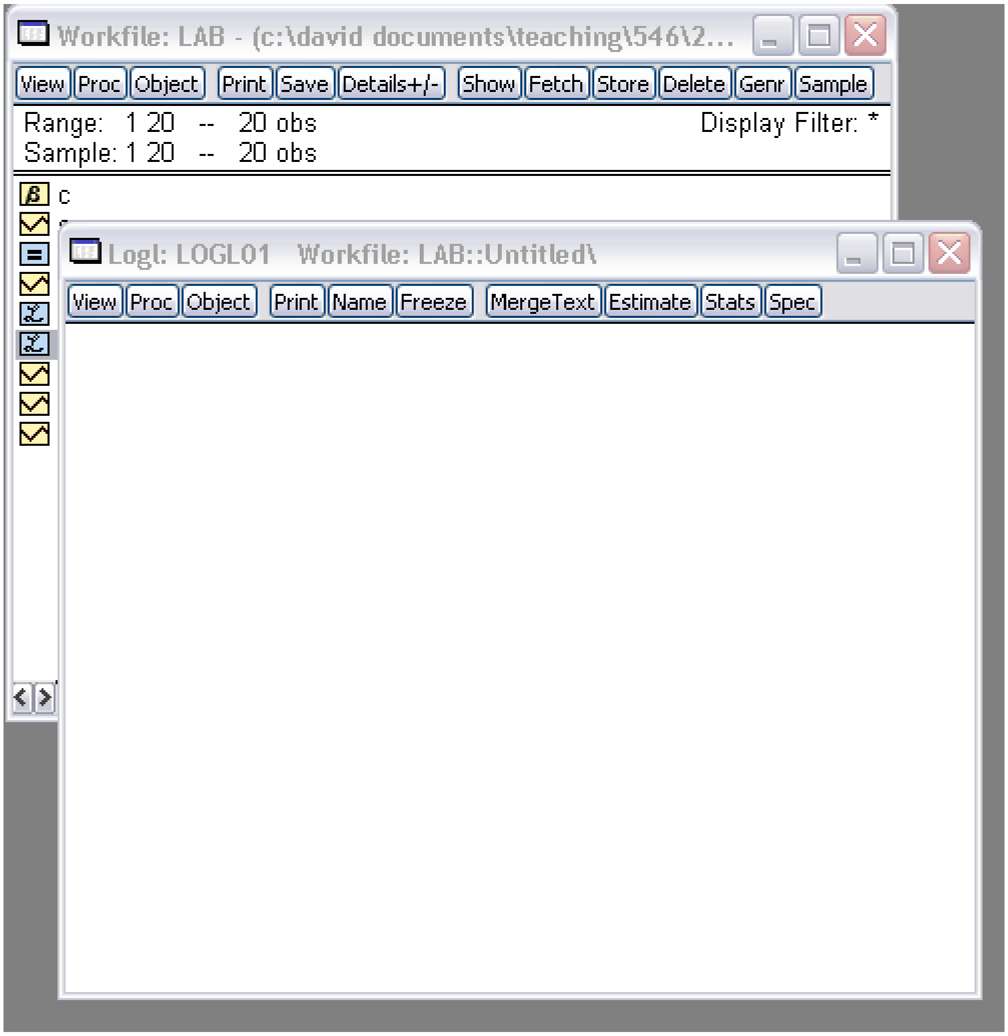
So, a typical term in the summation that appears in equation (2) is obtained by taking the logarithm of (1):

$$Log[p(y_i \mid \beta_1, \beta_2, \sigma)] = -\log(\sigma) - \log(2\pi)/2 - \frac{1}{2\sigma^2}\varepsilon_i^2 \ , \ \text{where} \ \ \varepsilon_i = (y_i - \beta_1 - \beta_2 x_i) . \ (3)$$

(c) To get EViews to perform MLE, we have to supply a typical term of the form (3). This is done via the so-called "**LOGL**" object. In your workfile, click on the **Object** button, and choose the "**New Object**" option. Then, highlight **LogL** as, shown below. You can supply a name for this object, either now, or later on when you save it – it may be a good idea to call this new object **LOGL01** at this stage.



When you click “**OK**”, this is what you will see next:

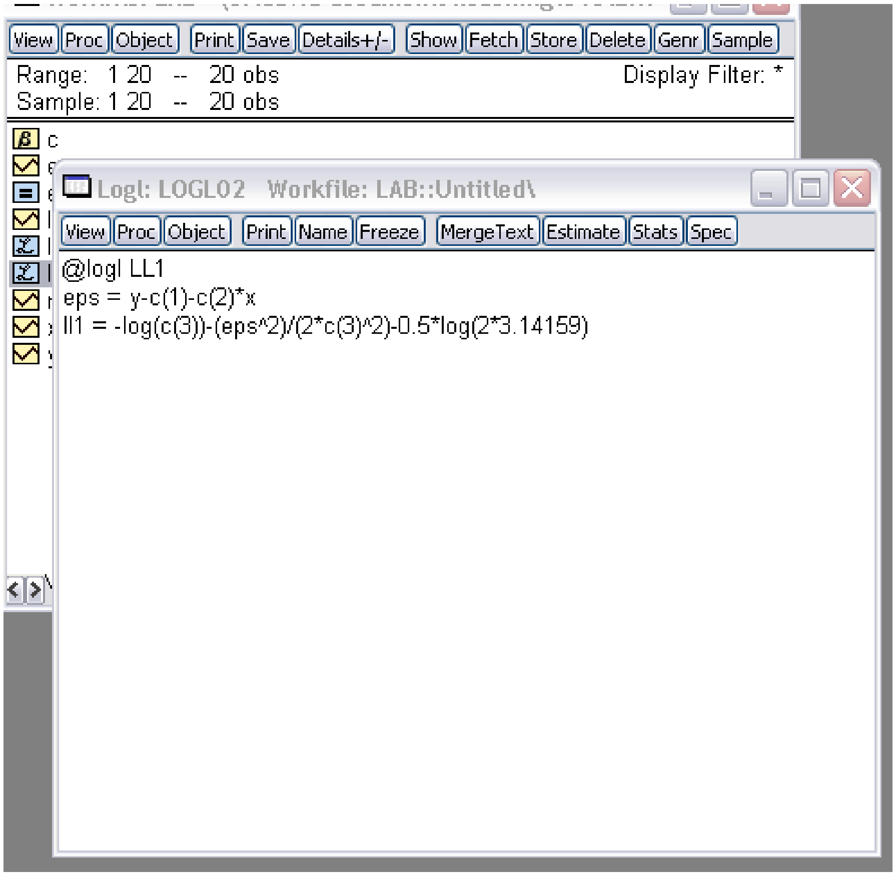


(d) You can now enter the formula for the i^{th} term of the Log-Likelihood Function into the empty Object Box:

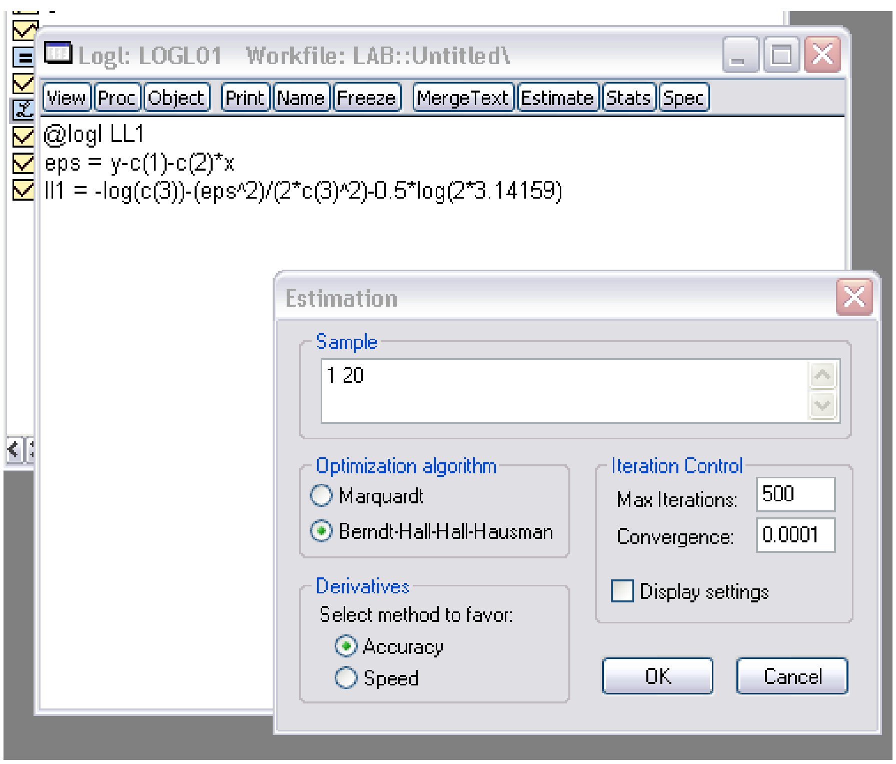
```
@logl LL1
eps = y-c(1)-c(2)*x
ll1 = -log(c(3))-(eps^2)/(2*c(3)^2)-0.5*log(2*3.14159)
```

The first line of code declares that we are constructing a log-likelihood function, and are going to call it LL1. (You can use any name you like.) The second line of code is introduced merely to make the expression in the third line a little simpler. Note that we are supplying the expression for just a single log-density. *EViews will assume that the data are independent, and do the summing that we see in equation (2) above for us.* Here, the coefficients c(1), c(2) and c(3) correspond to β_1, β_2, and σ respectively.

The object box will now look like this:



(e) Now, press the "**Estimate**" button, and this is what you will see:



Notice that you have a choice of algorithms for maximizing the Log-Likelihood Function. In evaluating the derivatives you should always choose "accuracy" over "speed". The following results then emerge when you click "**OK**":

Logl: LOGL01 Workfile: LAB2::Untitled\

View | Proc | Object | Print | Name | Freeze | MergeText | Estimate | Stats | Spec

LogL: LOGL01
Method: Maximum Likelihood (BHHH)
Date: 12/18/09 Time: 13:31
Sample: 1 20
Included observations: 20
Evaluation order: By observation
Convergence achieved after 75 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.315787	1.471689	0.214574	0.8301
C(2)	1.084213	0.110950	9.772121	0.0000
C(3)	2.413336	0.602963	4.002460	0.0001

Log likelihood	-45.99900	Akaike info criterion	4.899900
Avg. log likelihood	-2.299950	Schwarz criterion	5.049260
Number of Coefs.	3	Hannan-Quinn criter.	4.929056

The OLS results you saved as EQ01 are as follows:

Equation: EQ01 Workfile: LAB2::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: Y
Method: Least Squares
Date: 12/18/09 Time: 13:32
Sample: 1 20
Included observations: 20

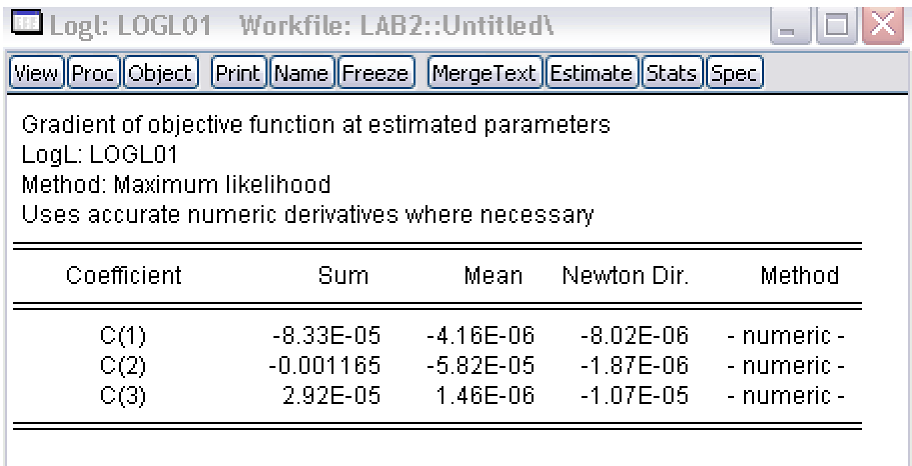
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.315789	1.181714	0.267230	0.7923
X	1.084211	0.098648	10.99075	0.0000

R-squared	0.870314	Mean dependent var	11.70000
Adjusted R-squared	0.863109	S.D. dependent var	6.875586
S.E. of regression	2.543884	Akaike info criterion	4.799901
Sum squared resid	116.4842	Schwarz criterion	4.899474
Log likelihood	-45.99901	Hannan-Quinn criter.	4.819338
F-statistic	120.7965	Durbin-Watson stat	1.753823
Prob(F-statistic)	0.000000		

(f) Why is the estimate of c(3) in the MLE output different from the "standard error of regression" in the OLS output? Why are the standard errors different? Verify that everything has actually been calculated correctly.

(g) Notice that the "Log likelihood" values are the same in each output – this is the value of the Log-Likelihood Function when the MLE's for the parameters are substituted into equation (2) above. It is the *maximized* value of the Log-Likelihood Function.

(h) Check that the Log-Likelihood Function has been properly maximized. In the LOGL01 output box, click on "**View**", "**Gradients**", then "**Summary**":



Gradient of objective function at estimated parameters
LogL: LOGL01
Method: Maximum likelihood
Uses accurate numeric derivatives where necessary

Coefficient	Sum	Mean	Newton Dir.	Method
C(1)	-8.33E-05	-4.16E-06	-8.02E-06	- numeric -
C(2)	-0.001165	-5.82E-05	-1.87E-06	- numeric -
C(3)	2.92E-05	1.46E-06	-1.07E-05	- numeric -

The gradients in each direction of the parameter space are evaluated at each point in the sample. These values are summarized by taking the mean and sum of each gradient across the sample values. We see that the gradients are essentially zero, as they should be.

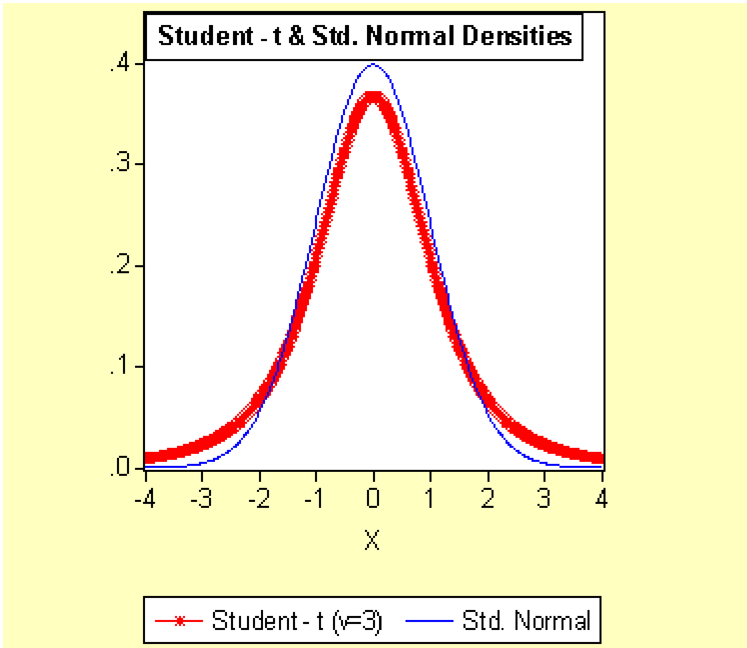
Note:

- In practice, you may need to edit the elements of the coefficient vector before you estimate a model by MLE to make sure that you don't silly starting values for the maximization algorithm. For example, in this exercise, if we had not already altered the coefficient values by running the initial OLS regression, we would have had to make sure that c(3) did not start of with the (default) value of zero – can you see why?
- If you need to modify the code for the Log-Likelihood specification in the LOGL01 object box, just select "**View**", then "**Likelihood Specification**" in that box's header bar, and then edit accordingly.

Now you are ready to estimate a non-standard model by MLE.

Part 2

Suppose that instead of assuming Normally distributed errors, you want to allow for "fat tails" (*i.e.*, a higher probability of outliers) in the error distribution. Recall that the Student-t distribution has a density function with this property if the associated degrees of freedom are relatively small. The need to allow for fatter tails in the density may arise, for example, when modeling financial returns. The Student-t distribution has a finite first moment only if $v > 1$, where v is the degrees of freedom parameter. It has a finite second moment only if $v > 2$, so probably the smallest value for the degrees of freedom that we should consider is $v = 3$.



Also, recall from a class example, that if the errors of our standard multiple linear regression model follow a *multivariate* Student-t distribution, then the MLE for the coefficient vector is just the OLS estimator. However, this result *does not* arise if the individual errors are *independent* Student-t distributed! This is the specification that we will follow next.

To set up the Log-Likelihood function we need to know the formula for the density function for a random variable that is Student-t distributed, with v degrees of freedom. This density takes the form:

$$p(\varepsilon_i \mid \nu) = const(h^{1/2})[1 + \frac{h\varepsilon_i^2}{\nu}]^{-(\nu+1)/2} \quad ; \quad -\infty < \varepsilon_i < \infty \qquad (4)$$

where 'h' is a scale parameter and '*const*' is the normalizing constant that ensures that the density is "proper" – that is, that it integrates to unity. (In the case of the normal density, this is the role that the $1/\sqrt{2\pi}$ term plays.) For the Student-t density, this normalizing constant is:

$$const = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)} \qquad (5)$$

where the "Gamma Function" is defined as:

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt\,. \qquad (6)$$

Fortunately, EViews can evaluate $\Gamma(x)$ for us via the **@gamma(x)** function. So, using (4), (5) and (6), we can build up a typical i^{th} term in the Log-Likelihood Function. Note from (4) that

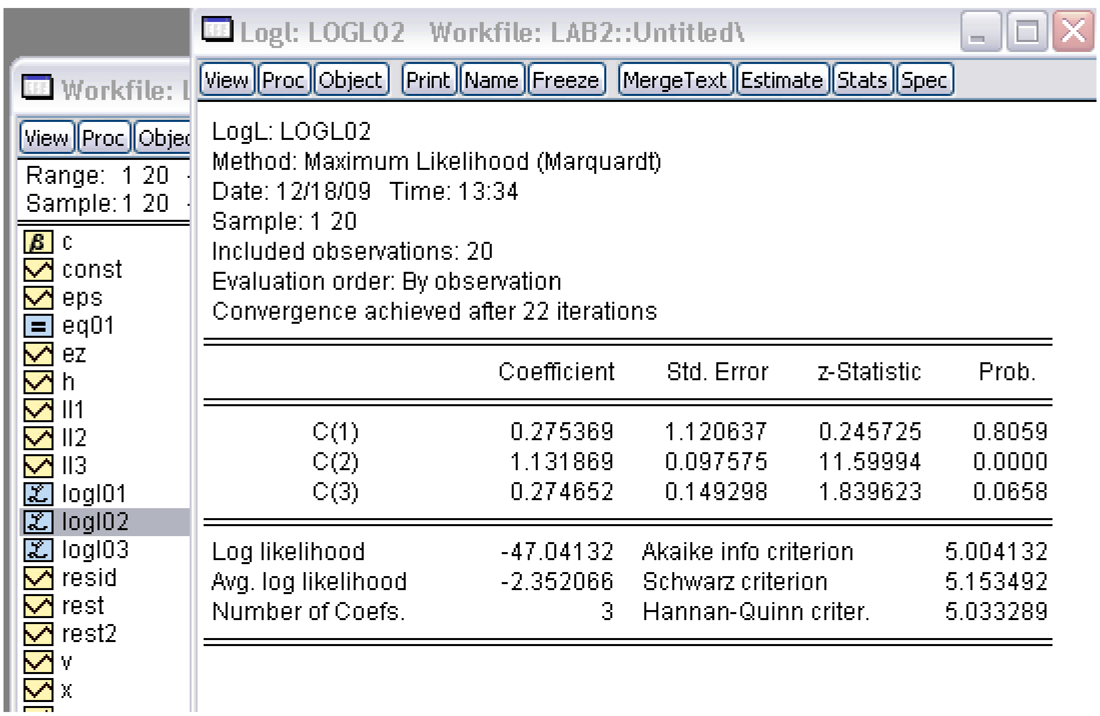
$$Log[p(\varepsilon_i \mid \nu)] = \log(const) + 0.5\log(h) - ((\nu+1)/2)\log[1 + h\varepsilon_i^2/\nu].$$

(a) Create a new object and name it LOGL02.

(b) Use the following code to set up the Log-Likelihood Function for our simple regression model with independent Student-t errors:

```
@logl LL2
eps = y-c(1)-c(2)*x
v=3
const=@gamma((v+1)/2)/(@sqrt(v*3.14159)*@gamma(v/2))
LL2 = log(const)+0.5*log(c(3))-((v+1)/2)*log(1+c(3)*(eps^2/v))
```

(c) You should now obtain the following MLE output:



LogL: LOGL02
Method: Maximum Likelihood (Marquardt)
Date: 12/18/09 Time: 13:34
Sample: 1 20
Included observations: 20
Evaluation order: By observation
Convergence achieved after 22 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.275369	1.120637	0.245725	0.8059
C(2)	1.131869	0.097575	11.59994	0.0000
C(3)	0.274652	0.149298	1.839623	0.0658

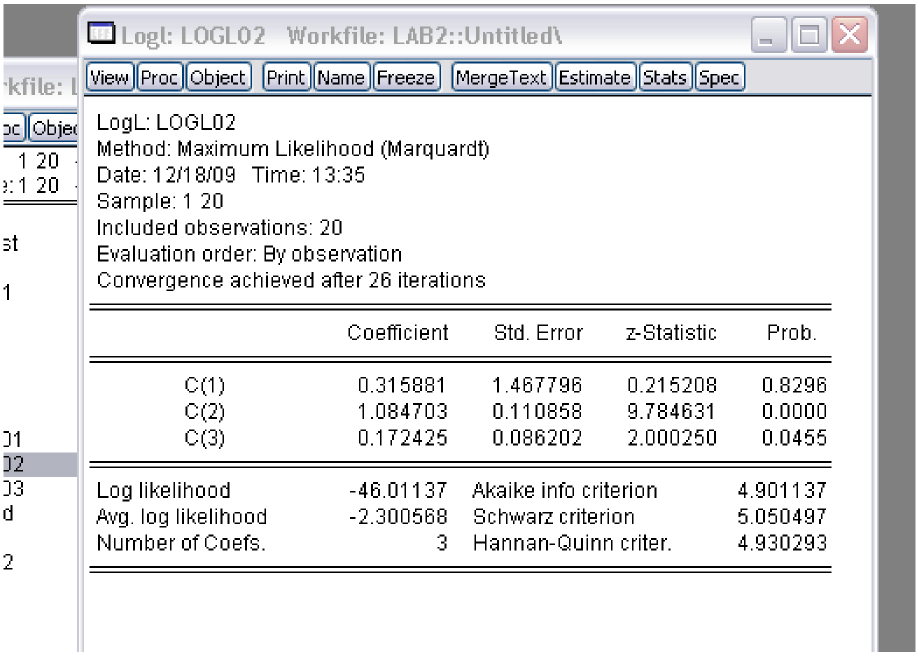
Log likelihood	-47.04132	Akaike info criterion	5.004132
Avg. log likelihood	-2.352066	Schwarz criterion	5.153492
Number of Coefs.	3	Hannan-Quinn criter.	5.033289

(d) Check the gradients to make sure that the Log-Likelihood Function has been properly maximized.

(e) Are your results at all sensitive to the choice of initial values for the coefficients?

(f) Compare the estimates of the two coefficients with those obtained when normally distributed errors were assumed.

(g) Recall that as $\nu \to \infty$ the Student-t density becomes a normal density. So, what do you think will happen if you keep increasing the value assigned to ν in the LOGL02 code? When ν = 300, you should get the following results:



LogL: LOGL02
Method: Maximum Likelihood (Marquardt)
Date: 12/18/09 Time: 13:35
Sample: 1 20
Included observations: 20
Evaluation order: By observation
Convergence achieved after 26 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.315881	1.467796	0.215208	0.8296
C(2)	1.084703	0.110858	9.784631	0.0000
C(3)	0.172425	0.086202	2.000250	0.0455

Log likelihood	-46.01137	Akaike info criterion	4.901137
Avg. log likelihood	-2.300568	Schwarz criterion	5.050497
Number of Coefs.	3	Hannan-Quinn criter.	4.930293

(h) The variance of a Student-t distribution is $(\nu / h)/(\nu - 2)$, which is defined if $\nu > 2$. Using the results of this last output, verify that the estimated error variance is approximately the same as the error *variance* (not standard deviation) estimate that was obtained with normal errors.

Part 3

Let's now suppose that we want to generalize our last model even further. As well as allowing for an error distribution with fat tails, let's suppose that we want to allow for a particular form of heteroskedasticity:

$$\text{var}(\varepsilon_i) = \exp\{\alpha_1 + \alpha_2 z_i) \quad ; \quad i = 1, 2, 3, \ldots, n$$

where α_1 and α_2 are unknown parameters, and z is another variable for which data are available. Note that the special case of homoskedastic errors arises if $\alpha_2 = 0$. If we equate this variance expression with the Student-t variance given at the top of this page, we obtain:

$$(\nu / h)/(\nu - 2) = \exp\{\alpha_1 + \alpha_2 z_i) ,$$

or,

$$h = \nu / [(\nu - 2)\exp\{\alpha_1 + \alpha_2 z_i\}] \quad ; \quad i = 1, 2, 3, \ldots, n.$$

(a) Create a new object and name it LOGL03.
(b) Use the following code to set up the Log-Likelihood Function for our simple regression model with independent but heteroskedastic Student-t errors:

```
@logl LL3
eps = y-c(1)-c(2)*x
v=3
h=v/((v-2)*@exp(c(3)+c(4)*z))
const=@gamma((v+1)/2)/(@sqrt(v*3.14159)*@gamma(v/2))
LL3 = log(const)+0.5*log(h)-((v+1)/2)*log(1+h*(eps^2/v))
```

(c) You should now obtain the following MLE output:

LogL: LOGL03
Method: Maximum Likelihood (Marquardt)
Date: 12/18/09 Time: 13:37
Sample: 1 20
Included observations: 20
Evaluation order: By observation
Convergence achieved after 28 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.595366	0.927136	0.642156	0.5208
C(2)	1.094496	0.103905	10.53358	0.0000
C(3)	1.736478	1.289447	1.346684	0.1781
C(4)	0.064108	0.105920	0.605245	0.5450

Log likelihood	-46.74678	Akaike info criterion	5.074678
Avg. log likelihood	-2.337339	Schwarz criterion	5.273825
Number of Coefs.	4	Hannan-Quinn criter.	5.113554

(d) Compare your results with those when homoskedasticity is assumed.

(e) Check the gradients to see that we have effectively maximized the Log-Likelihood Function.

(f) Looking at the last estimation results, is there any evidence of significant heteroskedasticity?