

**PROBLEM SET 2 – SOLUTIONS**

**Part I – Analytical Questions**

Problem 1: Consider the following stationary data generation process for a random variable  $y_t$

$$y_t = \beta y_{t-1} + e_t \quad e_t \sim N(0, 1) \text{ i.i.d.}$$

with  $|\beta| < 1$ , and  $y_0 \sim N(0, (1 - \beta^2))$ .

(a) Obtain the population mean, variance, autocovariances and autocorrelations.

$$E(y_t) = E\left(\sum_{i=0}^{t-1} \beta^i e_{t-i}\right) = 0$$

$$\sigma_y^2 = E(\beta^2 y_{t-1}^2 + 2\beta y_{t-1} e_t + e_t^2) = \beta^2 \sigma_y^2 + 1 = \frac{1}{(1 - \beta^2)}$$

since  $\{y_t\}$  is stationary. Further, notice that  $E(e_t^4) = 3$ . Noting only the nonzero terms:

$$E(y_t^4) = \beta^4 E(y_{t-1}^4) + 6\beta^2 E(y_{t-1}^2 e_t^2) + E(e_t^4) = \frac{3 + 6\beta^2 \sigma_y^2}{1 - \beta^4}$$

Finally,

$$COV(y_t, y_{t-1}) = E(y_t, y_{t-1}) = \beta E(y_{t-1}^2) + E(y_{t-1} e_t) = \beta \sigma_y^2$$

$$E(y_t, y_{t-k}) = \beta E(y_t, y_{t-k+1})$$

(b) Derive

$$(i) E\left[T^{-1} \sum_{t=2}^T y_{t-1} e_t\right] = T^{-1} \sum_{t=2}^T E(y_{t-1} e_t) = 0$$

$$(ii) E\left[T^{-1} \sum_{t=1}^T y_t^2\right] = T^{-1} \sum_{t=1}^T E(y_t^2) = T^{-1} \sum_{t=1}^T \frac{1}{1 - \beta^2} = \frac{1}{1 - \beta^2} = \sigma_y^2$$

$$(iii) E\left[T^{-1} \sum_{t=2}^T y_t y_{t-1}\right] = T^{-1} \sum_{t=2}^T E(y_t, y_{t-1}) = T^{-1} \sum_{t=2}^T \frac{\beta}{1 - \beta^2} \approx \frac{\beta}{1 - \beta^2} \beta \sigma_y^2$$

$$\begin{aligned}
 \text{(iv) Note: } E(y_t^4) &= \frac{3 + 6\beta^2\sigma_y^2}{1 - \beta^4}. \text{ Derive} \\
 V\left[T^{-1} \sum_{t=1}^T y_t^2\right] &= E\left[\left(T^{-1} \sum_{t=1}^T y_t^2 - \sigma_y^2\right)^2\right] = T^{-2} \sum_{t=1}^T \sum_{s=1}^{t-1} E(y_t^2 y_s^2) - \sigma_y^4 \\
 &= T^{-2} \sum_{t=1}^T E(y_t^4) + 2T^{-2} \sum_{t=1}^T \sum_{s=1}^{t-1} E(y_t^2 y_s^2) - \sigma_y^4 \\
 &= 3T^{-1} \sigma_y^2 - \sigma_y^2 + 2T^{-2} \sum_{t=1}^T \sum_{s=1}^{t-1} (1 + 2\beta^{2(t-s)})^2 \sigma_y^2 \\
 &= \sigma_y^4 \left(\frac{3}{T} - 1\right) + 2 \frac{\sigma_y^4}{T^2} \left(\frac{T(T-1)}{2} + \frac{2\beta^2}{(1-\beta^2)} \left[T - \sum_{t=1}^T \beta^{2(t-1)}\right]\right) \\
 &= \sigma_y^4 \left[3T^{-1} - T^{-1} + 4T^{-1} \beta^2 (1 - \beta^2)^{-1}\right] + O(T^{-2}) \\
 &\approx 2T^{-1} \sigma_y^4 \frac{1 + \beta^2}{1 - \beta^2}
 \end{aligned}$$

(c) Derive the limiting distribution of the sample mean. *Hint:  $\{y_t\}$  is not i.i.d. but it is ergodic for the mean.*

Since  $y_t$  is stationary, the only difficulty in the derivation of the distribution of the sample mean is that  $\{y_t\}$  is not an i.i.d. process, although it is ergodic. A general strong law of large numbers and a central limit theorem can be applied to such processes so that the sample mean converges *a.s.* to the population mean, which is zero, and is normally distributed around zero (i.e. using the central limit theorem for martingale difference sequences we discussed in class). However, the problem can also be solved from first principles as follows:

$$\begin{aligned}
 E(\bar{y}) &= E\left(T^{-1} \sum_{t=1}^T y_t\right) = T^{-1} \sum_{t=1}^T E(y_t) = 0; \\
 V(\bar{y}) &= E\left[\left(T^{-1} \sum_{t=1}^T y_t\right)^2\right] = T^{-2} \sum_{t=1}^T \sum_{s=1}^T E(y_t y_s) \\
 &= \sigma_y^2 T^{-2} \sum_{t=1}^T \sum_{s=1}^T \beta^{|t-s|} = \sigma_y^2 T^{-1} \left(1 + 2 \sum_{j=1}^{T-1} \left(1 - \frac{j}{T}\right) \beta^j\right) \\
 &\approx T^{-1} \sigma_y^2 \frac{1 + \beta}{1 - \beta}
 \end{aligned}$$

As  $T \rightarrow \infty, V(\bar{y}) \rightarrow 0$  so  $p \lim_{T \rightarrow \infty} \bar{y} = 0$ . However

$$V(\sqrt{T}\bar{y}) \rightarrow \sigma_y^2 \frac{1 + \beta}{1 - \beta} = \frac{1}{(1 - \beta)^2}$$

Next,  $\sqrt{T}\hat{y}$  is a linear function of independent, normally distributed components  $\{e_t\}$ , with a finite, non-zero variance in large samples, so that:

$$\sqrt{T}\hat{y} \xrightarrow{d} N\left(0, \frac{1}{(1-\beta^2)}\right)$$

(d) Obtain the limiting distribution of the least squares estimator of  $\beta$ . *Hint: calculate  $\sqrt{T}(\hat{\beta} - \beta)$  and then derive its distribution.*

Using the above results and Slutsky's theorem:

$$p \lim_{T \rightarrow \infty} (\hat{\beta} - \beta) = \frac{p \lim_{T \rightarrow \infty} T^{-1} \sum_{t=2}^T y_{t-1} e_t}{p \lim_{T \rightarrow \infty} T^{-1} \sum_{t=2}^T y_{t-1}^2} = \frac{0}{\sigma_y^2} = 0$$

Next, from Mann-Wald

$$\frac{1}{\sqrt{T}} \sum_{t=2}^T y_{t-1} e_t \xrightarrow{d} N\left(0, \frac{1}{1-\beta^2}\right)$$

Using Cramer's theorem, we thus have

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, (1-\beta^2))$$

Problem 2: Consider the AR(1) model of the previous exercise but suppose instead that  $e_t \sim N(0, \sigma^2)$ . The conditional density for observation  $t$  is therefore

$$\log f(y_t | y_{t-1}; \beta, \sigma^2) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{(y_t - \beta y_{t-1})^2}{2\sigma^2}$$

Let  $(\hat{\beta}, \hat{\sigma}^2)$  be the unrestricted *MLE* estimate of  $\theta = (\beta, \sigma^2)'$  and let  $(\tilde{\beta}, \tilde{\sigma}^2)$  be the restricted *MLE* estimate subject to the constraint  $R\beta = c$ , where  $R$  and  $c$  are known constants. Also, let

$$\hat{\Sigma} = \begin{bmatrix} \frac{1}{\hat{\sigma}^2} \frac{1}{T} \sum_{t=2}^T y_{t-1}^2 & 0 \\ 0 & \frac{1}{2(\hat{\sigma}^2)^2} \end{bmatrix}; \tilde{\Sigma} = \begin{bmatrix} \frac{1}{\tilde{\sigma}^2} \frac{1}{T} \sum_{t=2}^T y_{t-1}^2 & 0 \\ 0 & \frac{1}{2(\tilde{\sigma}^2)^2} \end{bmatrix}$$

- (a) Verify that  $\hat{\beta}$  minimizes the sum of squared residuals so that if it is the same as the OLS estimator.

This problem is a lot simpler if we proceed with the conditional likelihood, which is

$$L(\theta) = -\frac{T-1}{2} \log(2\pi) - \frac{T-1}{2} \log(\sigma^2) - \frac{\sum_{t=2}^T (y_t - \beta y_{t-1})^2}{2\sigma^2}$$

and can be viewed as a QMLE problem. The first order conditions are,

$$\frac{\partial L}{\partial \beta} = \frac{-\sum_{t=2}^T (y_t - \beta y_{t-1}) y_{t-1}}{\sigma^2} = 0, \Rightarrow \hat{\beta} = \frac{\sum_{t=2}^T y_t y_{t-1}}{\sum_{t=2}^T y_{t-1}^2}$$

which is the OLS estimator

- (b) Verify that  $\tilde{\beta}$  minimizes the sum of squared residuals subject to  $R\beta = c$  so that it is the restricted LS estimator.

The restricted likelihood is very simple and takes the form,

$$\tilde{L}(\theta) = -\frac{T-1}{2} \log(2\pi) - \frac{T-1}{2} \log(\sigma^2) - \frac{\sum_{t=2}^T (y_t - \beta y_{t-1})^2}{2\sigma^2} + \frac{\lambda}{\sigma^2} (R\beta - c)$$

where the lagrange multiplier is rescaled by  $\sigma^2$  for convenience and without loss of generality. The first order conditions are,

$$\frac{\partial \tilde{L}}{\partial \beta} = -\frac{\sum_{t=2}^T (y_t - \beta y_{t-1}) y_{t-1}}{\sigma^2} + \frac{\lambda R}{\sigma^2} = 0$$

from which

$$\tilde{\beta} = \hat{\beta} - \frac{\lambda R}{\sum_{t=2}^T y_{t-1}^2}$$

Of course, in this example the constraint can be trivially imposed in the problem and it would not require any estimation. The restricted least squares estimator would be calculated by minimizing,

$$\min_{\beta, \lambda} \sum_{t=2}^T (y_t - \beta y_{t-1})^2 + 2\lambda(R\beta - c)$$

whose first order conditions can be shown to coincide with those calculated via restricted maximum likelihood.

- (c) What assumptions did you make on  $f(y_t | \beta, \sigma)$  assuming  $y_0$  is unobserved? Discuss the importance of any simplifying assumptions and the importance that the parameter  $\beta$  has if its value were unrestricted.

We have been using the conditional likelihood for simplicity. However, if we had used the exact likelihood, the sensible assumption regarding  $f(y_t | \beta, \sigma)$  is that it is distributed  $N(0, \sigma^2 / (1 - \beta^2))$ . Note, however, that the exact likelihood is only valid if  $|\beta| < 1$ , otherwise, the log-likelihood becomes unbounded for  $\beta = 1$ .

- (d) Let

$$Q_T(\theta) = \frac{1}{T} \sum_{t=2}^T \log f(y_t | y_{t-1}; \beta, \sigma^2)$$

Show that

$$Q_T(\hat{\theta}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2} - \frac{1}{2}\log\frac{SSR_U}{T}$$

$$Q_T(\tilde{\theta}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2} - \frac{1}{2}\log\frac{SSR_R}{T}$$

Where  $SSR_U \equiv \sum_{t=2}^T (y_t - \hat{\beta}y_{t-1})^2$  and  $SSR_R \equiv \sum_{t=2}^T (y_t - \tilde{\beta}y_{t-1})^2$ . *Hint: Show that  $\hat{\sigma}^2 = \frac{SSR_U}{T}$  and  $\tilde{\sigma}^2 = \frac{SSR_R}{T}$ .*

To show,  $\hat{\sigma}^2 = \frac{SSR_U}{T}$  and  $\tilde{\sigma}^2 = \frac{SSR_R}{T}$  all that you need is to compute the first order conditions of the likelihood and the restricted likelihood with respect to  $\sigma^2$ .

Substitution of  $\hat{\sigma}^2 = \frac{SSR_U}{T}$  and  $\tilde{\sigma}^2 = \frac{SSR_R}{T}$  into  $Q_T(\theta)$  trivially delivers the two expressions for  $Q_T(\hat{\theta})$  and  $Q_T(\tilde{\theta})$ .

(e) Verify that the  $\hat{\Sigma}$  given above, although not the same as  $-\frac{1}{T}\sum_{t=2}^T H(y_t; \hat{\theta})$  is consistent for  $-E[H(y_t; \theta_0)]$ . Verify that  $\tilde{\Sigma}$ , although not the same as  $-\frac{1}{T}\sum_{t=2}^T H(y_t; \tilde{\theta})$ , is consistent for  $-E[H(y_t; \theta_0)]$ . *Hint: You may assume that  $\hat{\theta}$  is consistent for  $\theta_0$  and  $\tilde{\theta}$  is consistent for  $\theta_0$  under the null. This should make proving consistency of  $\hat{\sigma}^2$  and  $\tilde{\sigma}^2$  easier.*

The easiest approach is to calculate the hessian for the conditional and the restricted likelihoods. Using the hint, it then becomes trivial to show that both

$$-\frac{1}{T}\sum_{t=2}^T H(y_t; \hat{\theta}) \rightarrow E[H(y_t; \theta_0)] \text{ and}$$

$$-\frac{1}{T}\sum_{t=2}^T H(y_t; \tilde{\theta}) \rightarrow E[H(y_t; \theta_0)] \text{ and}$$

$$\hat{\Sigma} \rightarrow E[H(y_t; \theta_0)] \text{ and } \tilde{\Sigma} \rightarrow E[H(y_t; \theta_0)]$$

(f) Show that the Wald, LM, and LR statistics, using  $\hat{\Sigma}$  and  $\tilde{\Sigma}$  can be written as

$$W = T \frac{(R\hat{\beta} - c)^2}{(R^2 y_{t-1}^2) SSR_U}$$

$$LR = T \left\{ \log \left( \frac{SSR_R}{T} \right) - \log \left( \frac{SSR_U}{T} \right) \right\}$$

$$LM = T \frac{(Y_t' - \tilde{\beta} Y_{t-1}') Y_{t-1} (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1} (Y_t' - \tilde{\beta} Y_{t-1}')}{SSR_R}$$

No secrets here, just direct application of the formulas.

(g) Show that these three statistics can also be written as

$$W = T \frac{SSR_R - SSR_U}{SSR_U}$$

$$LR = T \log \left( \frac{SSR_R}{SSR_U} \right)$$

$$LM = T \frac{SSR_R - SSR_U}{SSR_R}$$

This involves simple algebraic manipulations.

Problem 3: Consider the stochastic process  $\{x_t\}$  which describes the number of trades per interval of time of a particular stock. Thus,  $x_t$  is integer-valued and non-negative. The Poisson distribution is commonly used to describe this type of process. Its density has the form:

$$f(x_t = j | \Omega_{t-1}; \theta) = \frac{e^{-\lambda_t} \lambda_t^j}{j!}$$

with conditional mean

$$E(x_t | x_{t-1}, \dots) = \lambda_t = \exp(\alpha + \beta x_{t-1})$$

Answer the following questions:

- (a) Under what parametric restrictions of  $\alpha$  and  $\beta$  will  $\{x_t\}$  be stationary?

This is a tricky question since we have a specification for the conditional mean directly rather than for  $\{x_t\}$ . However, it can be shown that the model is explosive for  $\beta > 0$ .

- (b) Write down the log-likelihood function for this problem conditional on  $x_1$ .

$$L(\theta) = \sum_{t=2}^T \{x_t (\alpha + \beta x_{t-1}) - \exp(\alpha + \beta x_{t-1}) - \ln x_t!\}$$

- (c) Compute the first order conditions and obtain the estimators for  $\alpha$  and  $\beta$ .

$$\frac{\partial L}{\partial \alpha} = \sum_{t=2}^T \{x_t - \exp(\alpha + \beta x_{t-1})\} = 0$$
$$\frac{\partial L}{\partial \beta} = \sum_{t=2}^T \{x_t [(\alpha + \beta x_{t-1}) - \exp(\alpha + \beta x_{t-1})]\} = 0$$

- (d) Suppose  $x_t$  is not Poisson distributed. Are  $\alpha$  and  $\beta$  consistently estimated by the conditional log-likelihood in (b)?

The Poisson distribution is an example of *linear exponential function* whose QMLE properties have been established by Gourieroux, Momfort and Trognon (1984). Check Cameron's (1998) book for the references.

- (e) Compute the one-step ahead forecast  $x_{t+1|t}$ . Next, describe the Monte Carlo exercise that would allow you to compute the two-step ahead forecast. Finally, describe what approach would you take if the data were not Poisson distributed but you wanted to produce multi-step ahead forecasts with the conditional mean estimated in (b).

The one step ahead forecast  $x_{t+1|t}$  is  $\lambda_{t+1} = \exp(\alpha + \beta x_t)$ .

To compute the two-step ahead forecast note that we need to compute  $E(x_{t+2} | x_t)$ . To do this, we will draw from a Poisson distribution with mean parameter  $\lambda_{t+1} = \exp(\alpha + \beta x_t)$ . Suppose you draw  $n$  times from this Poisson, you will have  $n$  values for  $\tilde{x}_{t+1}^j$  which you can then plug into the expression of the conditional mean to compute the forecast as:

$$x_{t+2|t} = \frac{1}{n} \sum_{i=1}^n \exp(\alpha + \beta \tilde{x}_{t+1}^i)$$

Computing the  $h$ -step ahead forecast would consist on drawing  $n$  times from a Poisson distribution whose conditional mean is given by the  $(h - 1)$ -step ahead forecast and calculating the average as is done above.

Note that as long as the conditional mean is correctly specified, multi-step ahead forecasts computed as described above would be consistent (albeit not efficient if the distribution is not Poisson).