ECON 546 – Spring 2010

Midterm Test – Solution

Q. 1

(a) The "moments" of a probability distribution are defined as $E(X^r)$; $r = 1, 2, 3, 4, \ldots$ where X is the random variable. These are also called the "raw" moments, or the moments about the origin. Often, we also talk of "moments about the mean", which are of the form $E[(X-\mu)^r]$, where $\mu = E(x)$. The moments of a distribution may or may not be defined. They won't be defined if the integrals needed to define the expectations diverge. If all of the moments of a distribution exist, they uniquely define the distribution. One way of obtaining the moments of a distribution is through differentiation of its characteristic function, $\phi_x(t) = E(e^{itx})$.

(b) The Cramér-Rao lower bound ~~s~~ provides a formula for the smallest value that ~~the~~ variance of an estimator can take, for a given estimation problem. It depends, among other things on Fisher's Information measure, which is obtained from the hessian of the log-likelihood function. If an estimator is unbiased, & its variance attains the Cramér-Rao lower bound, it must be a MLE. The <u>asymptotic</u> CRLB plays the same role in large samples. If we consider consistent (asymptotically unbiased) estimators, then if one such estimator has an asymptotic variance that attains the asymptotic CRLB, this estimator is asymptotically efficients. MLE'S have this desirable property.

(c) The Invariance property of MLE's is simply the property that the MLE of some continuous function of the parameters is just the same function of the MLE's of the individual paramets: If $\tilde{\theta}$ is the MLE of the vector θ, then $g(\tilde{\theta})$ is the MLE of $g(\theta)$, for any continuous fctn, $g(\cdot)$. For example, in the linear regression model with Normal errors, $\tilde{\beta} = (X'X)^{-1}X'y$. Let β_j be the j'th element of β. If we wanted the MLE of (β_1/β_2) we simply have to construct $(\tilde{\beta}_1/\tilde{\beta}_2)$. Being the MLE, this will be a consistent and asymptotically efficient estimator of (β_1/β_2).

Q.2. (a). By independence,

$$L(\theta \mid y_1, \ldots, y_n) \equiv p(y_1, y_2, \ldots, y_n \mid \theta)$$

$$= \prod_{i=1}^{n} p(y_i \mid \theta)$$

So, $L = \left(\theta/2\pi\right)^{n/2} e^{n\theta} \prod_i y_i^{-3/2} \exp\left[-\theta/2 \sum_i (y_i + y_i^{-1})\right]$

$$l = \log L = \frac{n}{2}\log\theta - \frac{n}{2}\log 2\pi + n\theta - \frac{3}{2}\sum_i \log y_i$$
$$- \theta/2 \sum_i (y_i + y_i^{-1})$$

$$(\partial l/\partial\theta) = \frac{n}{2\theta} + n - \frac{1}{2}\sum_i (y_i + y_i^{-1}) = 0.$$

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

$$\Rightarrow \tilde{\theta} = \left[\frac{1}{n}\sum_i (y_i + y_i^{-1}) - 2\right]^{-1}.$$

$$(\partial^2 l/\partial\theta^2 = -\frac{n}{2\theta^2} < 0$$, everywhere.

So, $\tilde{\theta}$ maximizes l.

(b) $I(\theta) = -E\left(\frac{\partial^2 l}{\partial\theta^2}\right) = E\left(\frac{n}{2\theta^2}\right) = \frac{n}{2\theta^2}.$

(c) $H_0 : \theta = 1$ vs $H_1 : \theta \neq 1$. The likelihood ratio itself is

$$\tilde{\lambda} = L(\tilde{\theta}_0) / L(\tilde{\theta}_1)$$

where $\tilde{\theta}_0 = 1$ and $\tilde{\theta}_1 = \tilde{\theta}$, from (a).

So, $L(\hat{\theta}_0) = \left(\frac{1}{2\pi}\right)^{n/2} e^{n} \prod_i y_i^{-3/2} \exp\left[-\frac{1}{2}\sum_i (y_i + y_i^{-1})\right]$

$L(\tilde{\theta}_1) = \left(\frac{\tilde{\theta}}{2\pi}\right)^{n/2} e^{n\tilde{\theta}} \prod_i y_i^{-3/2} \exp\left[-\frac{\tilde{\theta}}{2}\sum_i (y_i + y_i^{-1})\right]$

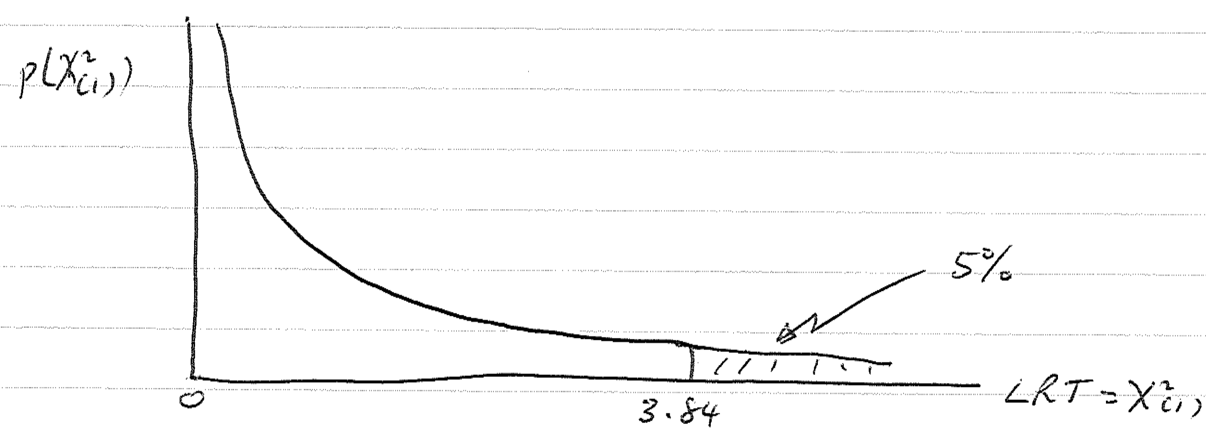
& so

$\tilde{\lambda} = L(\tilde{\theta}_0)/L(\tilde{\theta}_1) = \tilde{\theta}^{-n/2} \exp\left[(1-\tilde{\theta})\left(n - \frac{1}{2}\sum_i (y_i + y_i^{-1})\right)\right]$

The LRT statistic is $-2\log\tilde{\lambda}$, and this $\xrightarrow{d} \chi^2_{(1)}$ because H_0 is "nested" in H_1:

$LRT = n\log\tilde{\theta} - 2(1-\tilde{\theta})\left[n - \frac{1}{2}\sum_i (y_i + y_i^{-1})\right]$

(d) Assuming that n is large, $LRT \xrightarrow{d} \chi^2_{(1)}$.



Suppose we want $\alpha = 5\%$ for significance level. Then compute LRT and reject $H_0: \theta = 1$ if $LRT > 3.84$.

Note: If $\theta = 1$ is true, we would expect $\tilde{\theta}$ to be close to 1 if n large. If $\tilde{\theta} \to 1$ then LRT $\to 0$. This makes sense, because this says if H_0 is true we'll compute a small value of LRT, & this will lead us to not reject H_0.

Q. 3. (a) Again, we have independent sampling,

so $L(\theta \mid y_1, \ldots, y_n) \equiv p(y_1, \ldots, y_n \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta)$;

or $L = (\theta / 2\pi)^{n/2} \exp\left[-(\theta/2) \sum_i \left(\frac{1}{y_i}\right)\right] \prod_i y_i^{-3/2}$

and $$\ell = \log L = \frac{n}{2} \log \theta - \frac{n}{2} \log 2\pi - \theta/2 \sum_i \left(\frac{1}{y_i}\right) - \frac{3}{2} \sum_i \log y_i$$

$$(\partial \ell / \partial \theta) = \frac{n}{2\theta} - \frac{1}{2} \sum_i \left(\frac{1}{y_i}\right) = 0$$

$$\Longrightarrow \quad 1/\tilde{\theta} = \frac{1}{n} \sum_i (1/y_i)$$

$$\text{or} \quad \tilde{\theta} = \left[\frac{1}{n} \sum_i \left(\frac{1}{y_i}\right)\right]^{-1} = HM(y_i\text{'s}).$$

$$(\partial^2 \ell / \partial \theta^2) = -\frac{n}{2\theta^2} < 0 \quad \text{everywhere.}$$

So, we have maximized L.

(b) $$I = -E\left[\partial^2 l / \partial\theta^2\right] = n/2\theta^2 ,$$

$$IA = \underset{n\to\infty}{\text{Limit}}\left[\frac{1}{n} I\right] = \left(\frac{1}{2\theta^2}\right).$$

$\tilde{\theta}$ is consistent for θ, so $\left(\frac{1}{2\tilde{\theta}^2}\right)$ is consistent for $(1/2\theta^2)$, by Slutsky's Theorem.

So, $\text{plim}\left[\frac{1}{2\tilde{\theta}^2}\right] = IA$.

(c) $$\sqrt{n}\,(\tilde{\theta} - \theta) \xrightarrow{d} N\left[0, IA^{-1}\right] = N\left[0, 2\theta^2\right]$$

(d) $$\frac{\partial p(y_i)}{\partial y_i} = \sqrt{\theta/2\pi}\left\{\exp\left[-\theta/(2y_i)\right]\cdot\left(-\frac{3}{2}\right)y_i^{-5/2} + y_i^{-3/2}\exp\left[-\theta/(2y_i)\right]\cdot\frac{\theta}{2}y_i^{-2}\right\}$$

$$= 0 \quad ; \text{ for mode.}$$

So $$\frac{1}{2}\exp\left[-\theta/(2y_i)\right]\left\{\theta\, y_i^{-7/2} - 3y_i^{-5/2}\right\} = 0$$

$$\Rightarrow \quad y_m^{-1}\cdot\theta = 3$$

$$\Rightarrow \quad y_m = (\theta/3)\ ; \text{ for mode of } p(y_i|\theta).$$

By invariance, the MLE of the mode is $(\tilde{\theta}/3)$, & this will be weakly consistent and asymptotically efficient.

(e) From part (c) above, the asymptotic variance of $\tilde{\theta}$ is $2\theta^2/n$, so the asymptotic std. deviation of $\tilde{\theta}$ is $\theta\sqrt{2/n}$.

We can estimate this consistently, using $\tilde{\theta}\sqrt{2/n}$: this is the asymptotic std. error for $\tilde{\theta}$. We know that $\tilde{\theta}$ is asymptotically normal, so a 95% asymptotic confidence interval for θ is:

$$\tilde{\theta} \pm 1.96\left(\tilde{\theta}\sqrt{2/n}\right).$$

(f) $\phi_y(t) = \exp\left\{-\sqrt{-2i\theta t}\right\}$.

$$\phi_y'(t) = \exp\left\{-\sqrt{-2i\theta t}\right\}\left(-\tfrac{1}{2}\right)(-2i\theta)$$
$$= i\theta\exp\left[-\sqrt{-2i\theta t}\right].$$

$$E(Y) = \underset{t\to 0}{\text{Limit}}\left[\phi_y'(t)/i\right]$$
$$= \underset{t\to 0}{\text{Limit}}\left[\theta\exp\left[-\sqrt{-2i\theta t}\right]\right]$$
$$\to \infty.$$

Q.4 (a) Maximum likelihood estimation has been used – in OUTPUT 1, the second line says "Method: ML – ARCH (Marquardt) - - -". The Marquardt algorithm has been used to maximize the log-likelihood function. Convergence was achieved after 11 iterations. $n = 20{,}276$ (very large).

(b) The C.I. is $0.084296 \pm 1.96(0.007123)$

(because the "t-statistic" is asymptotically standard Normal).

So, the interval is 0.084296 ± 0.01396

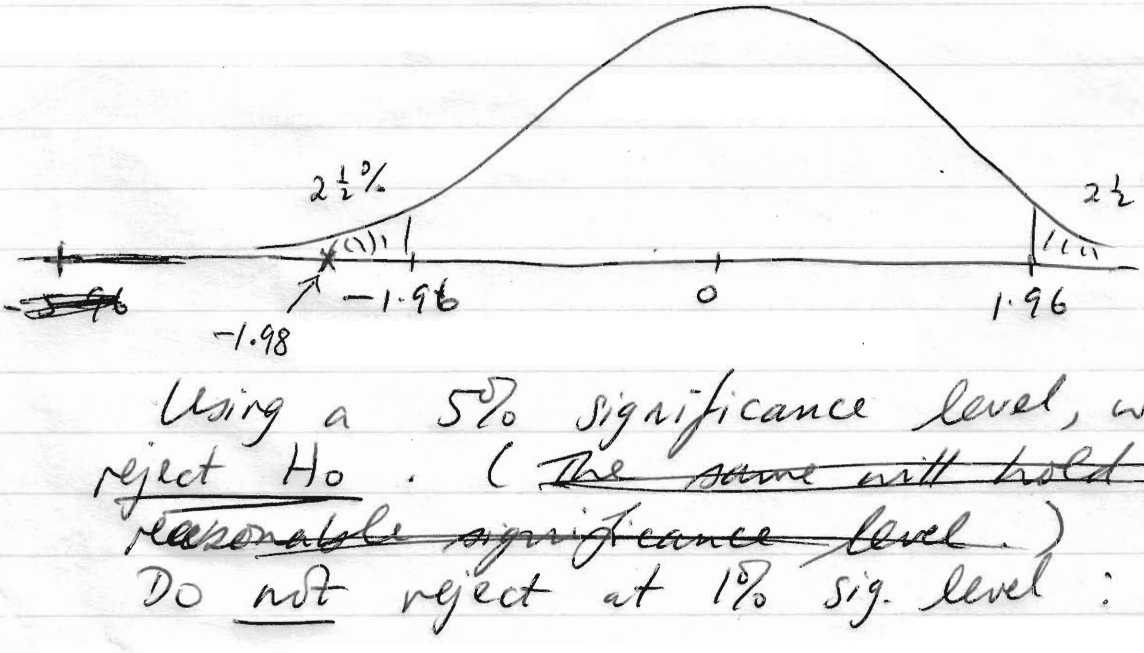
or $[0.07033, 0.0983]$

(The interval limits are unit-less, because the y-variable (RETURN) & the x-variable ($RETURN_{-1}$) have the same units, in this case.)

(c) The estimate of the D.O.F. is 6.503847. The associated asymptotic std. error is 0.250915.

$H_0: DOF = 7$ vs. $H_1: DOF \neq 7$

So, $z = \left(\dfrac{6.503847 - 7}{0.250915}\right) = -1.977$ ~~-5.9628~~.



Using a 5% significance level, we ~~clearly~~ just reject Ho. (~~The same will hold for any reasonable significance level.~~)

Do not reject at 1% sig. level : $t_c = \pm 2.575$

(d) Three restrictions have been imposed:

Coefficient of $RETURN_{-3} = 0$
Coefficient of $RETURN_{-4} = 0$
DOF = 7 (no longer estimated).

$\log \tilde{L}_1$ comes from Output 1 : $\log \tilde{L}_1 = -25685.58$

$\log \tilde{L}_0$ comes from Output 2 : $\log \tilde{L}_0 = -25690.22$

So, $\tilde{\lambda} = (\tilde{L}_0 / \tilde{L}_1)$

$$
\begin{aligned}
-2\log\tilde{\lambda} &= 2\,[\log\tilde{L}_1 - \log\tilde{L}_0] \\
&= 2\,[-25685.58 + 25690.22] \\
&= 9.28.
\end{aligned}
$$

$-2\log\tilde{\lambda} \xrightarrow{d} \chi^2_{(3)}$ if H_0 is true. We have a very large sample here. The 5% and 1% critical values for $\chi^2_{(3)}$ are 7.81 and 11.34. So, we would reject these restrictions at the 5% level, but not at the 1% level.