ECON 546 – Spring 2008

Mid-Term Test
Solution

Q.1

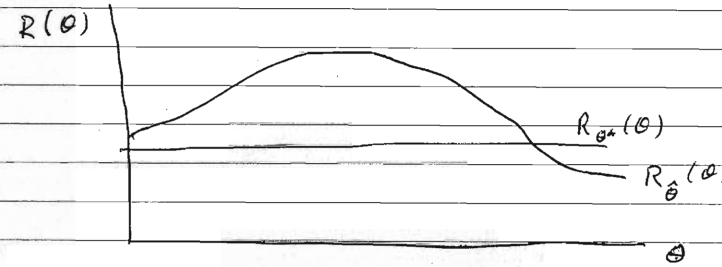
(a) In the decision-theoretic approach to estimation we "penalize" poor estimation performance by introducing a Loss function:

$$L(\hat{\theta};\theta) \geq 0 \quad ; \quad \forall\ \theta, \hat{\theta}$$
$$= 0 \quad ; \quad \text{iff } \hat{\theta} = \theta.$$

Because $\hat{\theta}$ is a random fctn. of the data, $L(\cdot)$ is also random, so it is not easy to compare the losses of different estimators. So we average over the sample values, and the result is called the Risk fctn. for $\hat{\theta}$:

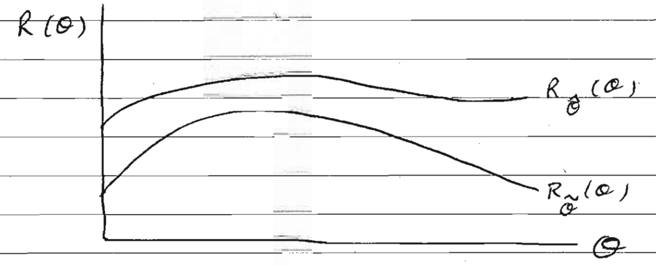
$$R_{\hat{\theta}}(\theta) = E_y\left[L(\hat{\theta}(y);\theta)\right].$$

Typically, risk will depend on the value of θ, but it could constant; e.g.:



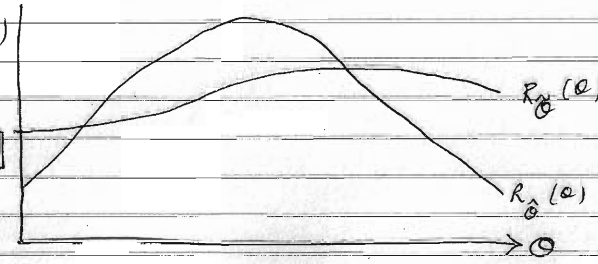
We then seek an estimator with low risk (ideally over the full parameter space). If the loss fctn. is quadratic, then risk is just the familiar MSE of the estimator.

If one estimator, $\tilde{\theta}$, has risk that is less than or equal to the risk of $\hat{\theta}$ for all θ, and strictly less for some θ, then $\tilde{\theta}$ dominates $\hat{\theta}$, and we say that $\hat{\theta}$ is inadmissible. Clearly, under the chosen loss structure, $\tilde{\theta}$ would be preferred to $\hat{\theta}$:



If we have several admissible estimators, their risk fctns. may cross. In this case we might use the mini-max estimator – the one with the smallest maximum risk:

[$\tilde{\theta}$ is mini-max]



(b) An asymptotically efficient estimator is one that is weakly consistent (so that $\text{plim}(\hat{\theta}) = \theta$) and whose asymptotic distribution has a variance (or covariance matrix) that attains the asymptotic Cramér-Rao lower bound. The latter is defined as $IA^{-1}(\theta)$, where $IA = \underset{n\to\infty}{\text{Limit}}[I(\theta)]$ and $I(\theta)$ is Fisher's Information matrix:

$$I(\theta) = -E[\partial^2 \text{Log}\, L(\theta)/\partial\theta\partial\theta']$$

and $L(\theta)$ is the Likelihood Function. The asymptotic Cramér-Rao Lower Bound, $IA^{-1}(\theta)$ is the "best" covariance matrix that any consistent estimator can have. This lower bound is attained (under the usual regularity conditions) by all MLE's — they are consistent & Best Asymptotically Normal.

(c) The N-P Lemma justifies the use of the LRT when we have a point null hypothesis & a point alternative hypothesis — e.g.:

$$H_0: \theta = 1 \quad \text{vs.} \quad H_0: \theta = 4.$$

The Likelihood Ratio is $(\tilde{L}_0/\tilde{L}_1)$, where $\tilde{L}_0$ indicates the value of the likelihood function when we insert the value of θ if H_0 is true; & similarly for $\tilde{L}_1$. The N-P Lemma tells us that if we use the following test (rule):

"Reject H_0 if $\lambda = (\tilde{L}_0/\tilde{L}_1) < k$; where k is chosen such that $\Pr.[\lambda < k \mid H_0 \text{ True}] = \alpha$, and α is the desired significance level", then this rule will lead to a Most Powerful Test. No other test for this problem (& using the same α and sample) will have greater power — greater ability to reject H_0 if it is false.

The practical difficulty of this Lemma is that to apply it for a particular situation, we would have to be able to establish the sampling distribution of λ, so we could obtain k, & hence know the critical region. This may be difficult in practice (at least in finite samples).

Q. 2

(a) By independence —

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

$$= \theta^{\frac{n}{2}}(2\pi)^{-n/2} \prod_i y_i^{-3/2} e^{-\theta \sum_i (1/2y_i)}$$

$$\text{Log}\, L = \frac{n}{2}\log\theta - \frac{n}{2}\log(2\pi) - \frac{3}{2}\sum_i \log(y_i) - \theta \sum_i (1/2y_i)$$

$$\partial \log L/\partial\theta = \frac{n}{2\theta} - \sum_i (1/2y_i) = 0 \qquad \text{(i)}$$

From (i):

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

(This is actually the Harmonic Mean of the data.)

From (i):

$$(\partial^2 \log L / \partial\theta^2) = -n/2\theta^2 < 0 \text{ everywhere,}$$

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

(b) $I(\theta) = -E\left[\frac{\partial^2 \log L}{\partial\theta^2}\right] = \frac{n}{2\theta^2}$

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

$IA(\theta)$ can be estimated consistently a $\left(\frac{1}{2\tilde{\theta}^2}\right)$, by Slutsky's Theorem.

[N.B.: I did not ask you to obtain I^*.]

(c) $\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{d} N[0, 2\theta^2]$.

(Because $p(y_i|\theta)$ satisfies the regularity conditions.)

The above is all that is needed. Note, however, that $2\theta^2$ is the asymptotic variance of $\sqrt{n}(\tilde{\theta} - \theta)$, so the asy. var $(\tilde{\theta})$ is $2\theta^2/n$.

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

$\phi_y'(t) = [\partial\phi_y(t)/\partial t]$

$= \exp\{-\sqrt{-2i\theta t}\}(-1)(\tfrac{1}{2})(-2i\theta)(-2i\theta t)^{-\frac{1}{2}}$

$= \dfrac{i\theta \exp\{-\sqrt{-2i\theta t}\}}{\sqrt{-2i\theta t}}$

Then, $E(Y) = \left[\phi_y'(t)/i\right]_{t=0}$

$= \underset{t\to 0}{\text{Limit}}\left\{\dfrac{\theta \exp\{-\sqrt{-2i\theta t}\}}{\sqrt{-2i\theta t}}\right\}$

The numerator $\to \theta$ (times 1)
The denominator $\to 0$

So $E(Y) \to \infty$.

As $Var(Y) = E(Y^2) - [E(Y)]^2$

& the second term is infinite, the variance is likely to be infinite too. (In fact this turns out to be correct.)

Q.3.

(a) By independence ~

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

$$= \prod_i y_i \; \theta^{-2n} \exp\left[-\frac{1}{2\theta^2} \sum_i y_i^2\right]$$

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

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

$$= -\frac{2n}{\theta} + \sum_i y_i^2 / \theta^3 \quad = 0$$

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

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

(Must be +ve root, as $\theta > 0$.)

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

& when $\theta = \tilde{\theta}$, this Hessian is

$$\left(\frac{4n^2}{\sum_i y_i^2}\right) - 3 \left(\frac{4n^2}{\sum_i y_i^2}\right)$$

$$= -\left(8n^2 / \sum y_i^2\right) < 0.$$

We have maximized the L.F.

(b) $E(Y) = \mu_1' = \theta \, 2^{\frac{1}{2}} \, \Gamma(1 + \frac{1}{2})$

$$= \theta \sqrt{2}\, \Gamma(\tfrac{1}{2}) . \tfrac{1}{2}$$

$$= \theta \sqrt{2} \sqrt{\pi} . \tfrac{1}{2}$$

$$= \theta \sqrt{\pi} / \sqrt{2} = \theta \sqrt{\pi/2} .$$

$$E(Y^2) = \mu_2' = \theta^2 \, 2\, \Gamma(2) = \theta^2 \, 2\, \Gamma(1) . 1$$

$$= 2\theta^2$$

So, $Var(Y) = \mu_2' - (\mu_1')^2 = 2\theta^2 - \theta^2(\pi/2)$

$$= \theta^2 (4 - \pi)/2 .$$

The MLE for $E(Y) = 2\tilde{\theta}^2$, by invariance

The MLE for $Var(Y) = \tilde{\theta}^2(4-\pi)/2$, " " "

& the MLE for $S.D.(Y) = \tilde{\theta}\, {}^{+}\sqrt{(4-\pi)/2}$, " " "

The estimators will each be Best Asymptotically Normal (& weakly consistent).

(c) This is a case of Nested Hypotheses, so

$$LRT = -2\log(\tilde{\lambda}) = -2 \log\left[\tilde{L}_0 / \tilde{L}_1\right]$$

$$= 2\left[\log \tilde{L}_1 - \log \tilde{L}_0\right)$$

$$\xrightarrow{d} \chi^2_{(1)} \;; \text{ if } H_0 \text{ is true.}$$

$$\log \tilde{L}_0 = \sum_i \log y_i - 0 - \tfrac{1}{2} \sum_i y_i^2 .$$

$$\log \tilde{L}_1 = \sum_i \log y_i - 2n \log(\tilde{\theta}) - \frac{1}{2\tilde{\theta}^2} \sum y_i^2 .$$

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

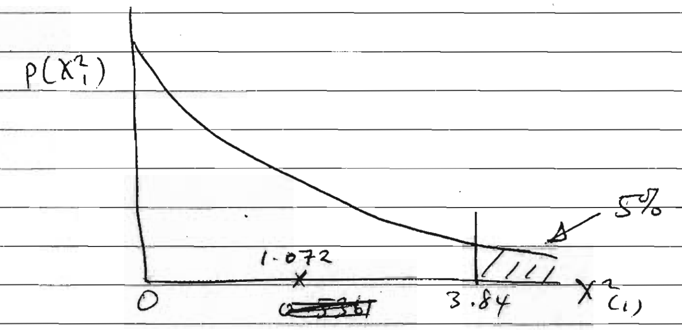
To apply the test, reject H_0 if $LRT > crit$, where $Pr.[\chi^2_{(1)} > crit] = \alpha$, & α is the desired significance level. This test will be asymp. valid.

(d) $\tilde{\theta} = +\sqrt{\frac{1}{2n}\sum_i y_i^2} = +\sqrt{180/200} = 0.94868$

$LRT = 2\{-200\log(0.9487) - (180/1.8) + 90\}$

$= \cancel{0.5361}\ 1.072$

The 5% critical value for $\chi^2_{(1)}$ is 3.84. So, we cannot reject H_0 ~~homoskedasticity~~ at the 5% level, or even at the 10% level (crit. = 2.71).



Q.4. This is from one of your lab. classes, but at that stage we had not covered the LRT.

(a) From Output 1, $\log\tilde{L}_0 = -135.3922$.
From Output 2, $\log\tilde{L}_1 = -135.2106$

So, $LRT = 2[-135.2106 + 135.3922] = 0.3622$

$H_0: c(5) = 0$ vs $H_1: c(5) \neq 0$.

(or, $H_0: \delta_1 = 0$ vs. $H_1: \delta_1 \neq 0$)

$LRT \xrightarrow{d} \chi^2_{(1)}$, so we once again do not reject H_0. We cannot reject homoskedasticity, at any reasonable significance level.

(b) Using the z-statistic for c(5) in Output 2, the p-value is correct, as we have a 2-sided alternative. So, we would not reject the null hypothesis of homoskedasticity at any reasonable significance level, as $p = 0.8154$.

(c) From Output (1):

$0.764673 \pm 1.96(0.218204)$

$\Rightarrow [0.33699\ ,\ 1.1924]$

(1.96 from N(0,1) Table)

From Output 2 :

$$0.659183 \pm 1.96\,(0.450980)$$

$$\Rightarrow \left[-0.2247 \quad , \quad 1.54311\right]$$

The two intervals overlap – in fact the first is covered entirely by the second. The first $\hat{c}_3$ is within the 2nd interval, & the second $\hat{c}_3$ is in the first interval. This suggests there is the same c_3.

(d) We have only 18 observations !! All of the above is based on asymptotic results when $n \to \infty$. We would have to treat our results with <u>great caution</u> !