ECON 546 - Spring 2009
Solution to Final Exam

Q.1.

(a) $$p(y_i | \beta, \rho) = \frac{\beta^\rho}{\Gamma(\rho)} y_i^{\rho-1} \exp\{-\beta y_i\}$$

$$L = \prod_i p(y_i|\beta,\rho) = \beta^{n\rho} [\Gamma(\rho)]^{-n} \prod_i y_i^{\rho-1} \cdot \exp[-\beta \sum_i y_i]$$

$$\log L = n\rho \log\beta - n\log\Gamma(\rho) + (\rho-1)\sum_i \log y_i - \beta \sum_i y_i$$

$$(\partial \log L / \partial\beta) = \frac{n\rho}{\beta} - \sum_i y_i = 0 \quad ; \text{ for max.}$$

So, $$\tilde{\beta} = \frac{n\rho}{\Sigma y_i} = (\rho / \bar{y}) \quad ; \quad \bar{y} = \frac{1}{n}\sum_i y_i$$

S.O.C. :

$$(\partial^2 \log L / \partial\beta^2) = -\frac{n\rho}{\beta^2} < 0 \text{ , everywhere.} \checkmark$$

(b) $$W = (\tilde{\beta} - 1)^2 I^*(\tilde{\beta}) \quad ; \quad R = 1 \; ; \; r = 1.$$

$$I = -E[\partial^2 \log L / \partial\beta^2] = n\rho/\beta^2$$

$$IA = \lim_{n\to\infty} \left[\frac{1}{n} I\right] = \rho/\beta^2$$

$$I^*(\tilde{\beta}) = (n\rho/\tilde{\beta}^2) \quad , \quad \text{so} \quad \text{plim}\left[\frac{1}{n} I^*\right] = IA.$$

So, $$W = (\tilde{\beta} - 1)^2 n\rho / \tilde{\beta}^2 \quad ; \quad \text{where } \tilde{\beta} = \frac{n\rho}{\sum_i y_i}$$

Calculate W and reject H_0 if W exceeds the $\chi^2_{(1)}$ critical value, for the chosen significance level. This will be valid if $n \to \infty$.

(c) $$LM = D\log L_1(\tilde{\psi}_0)' \, I^*(\tilde{\psi}_0)^{-1} \, D\log L_1(\tilde{\psi}_0)$$

$$= \left(\frac{n\rho}{1} - \sum_i y_i\right)\left(\frac{1^2}{n\rho}\right)\left(\frac{n\rho}{1} - \sum_i y_i\right)$$

$$= n(\rho - \bar{y})^2/\rho$$

Again, reject H_0 if LM exceeds the appropriate $\chi^2_{(1)}$ critical value. It will be valid if $n \to \infty$.

(d) $$W = \left(\tfrac{3}{2} - 1\right)^2 \left(\frac{100}{(3/2)^2}\right) = 11\tfrac{1}{9}$$

$$LM = 100\left(1 - \tfrac{2}{3}\right)^2 = 11\tfrac{1}{9}.$$

As it happens, $LM = W$ in this example. The 1% critical value for $\chi^2_{(1)} = 6.63$. In each case, we reject H_0 at the 1%, 5% or 10% levels. (NOT sensitive to the significance level.)

Q. 2.

(a) $$p(y_i \mid \theta, \lambda) = (2\lambda)^{-1} \exp\{-|y_i - \theta|/\lambda\}$$

$$L = \prod_i p(y_i \mid \theta, \lambda) = (2\lambda)^{-n} \exp\{-\sum_i |y_i - \theta|/\lambda\}$$

$$l = \log L = -n \log(2\lambda) - \sum_i |y_i - \theta|/\lambda$$

To <u>maximize</u> l w.r.t. θ (for any λ) we need to <u>minimize</u> $\sum_i |y_i - \theta|$. So, from the introduction to the question, this immediately implies that $\tilde{\theta} = M$, the sample median.

Also, $$(\partial l / \partial \lambda) = -n\left(\frac{1}{2\lambda}\right) \cdot 2 + \frac{1}{\lambda^2} \sum_i |y_i - \theta| \overset{=0}{}$$

$$\Rightarrow \tilde{\lambda} = \frac{1}{n} \sum_i |y_i - \tilde{\theta}| \quad ; \quad \tilde{\theta} = M.$$

(b) $$\text{Var.}(y_i) = E(y_i^2) - [E(y_i)]^2 = 2\lambda^2 + \theta^2 - \theta^2 = 2\lambda^2$$

So, $$CV = \theta / \sqrt{2\lambda^2} = \theta / (\lambda\sqrt{2}).$$

MLE's are weakly consistent, so by Slutsky's Theorem, a consistent estimator of CV is

$$\widetilde{CV} = \tilde{\theta} / (\tilde{\lambda}\sqrt{2}),$$

where $\tilde{\theta}$ & $\tilde{\lambda}$ are defined in (a).
By <u>invariance</u>, it is the MLE of CV, so it is asymptotically efficient & asymptotically normal.

(c) Note that in this case, θ is known, & $\lambda > 0$. So, the prior for λ that represents ignorance is

$$p(\lambda) \propto \frac{1}{\lambda}.$$

So, by Bayes' Theorem:

$$p(\lambda \mid y, \theta) \propto p(\lambda) L(\lambda \mid \theta, y)$$

$$\propto \lambda^{-1} (2\lambda)^{-n} \exp\left[-\frac{1}{\lambda} \sum_i |y_i - \theta|\right]$$

$$\propto \lambda^{-(n+1)} \exp\left[-\frac{1}{\lambda} \sum_i |y_i - \theta|\right]$$

$$\propto \lambda^{-(n+1)} \exp\left[-c/\lambda\right],$$

where $c = \sum_i |y_i - \theta|$ is known.

For a 0-1 loss function, we estimate λ as the mode of the posterior pdf:

$$\partial p(\lambda \mid \theta, y) / \partial\lambda = -(n+1)\lambda^{-(n+2)} e^{-c/\lambda}$$

$$+ \lambda^{-(n+1)} (-c)(-1) \lambda^{-2} \stackrel{=0}{}$$

$$\Rightarrow \hat{\lambda}^{-(n+2)} (n+1) e^{-c/\hat{\lambda}} = c \hat{\lambda}^{-(n+3)} e^{-c/\hat{\lambda}}$$

$$\Rightarrow (n+1)\hat{\lambda} = c$$

$$\Rightarrow \hat{\lambda} = \left(\frac{c}{n+1}\right) = \frac{1}{n+1} \sum_i |y_i - \theta|$$

Note that the MLE & the Bayes estimator differ only in the use of 'n' or '$n+1$' in the denominator. Both estimators will converge in probability to λ at the same rate — the usual rate of $n^{-1/2}$.

Q.3.

(a) For $i = 1, \cdots, n_1$,

$$p(y_1, \cdots, y_{n_1}) = \prod_{i=1}^{n_1} \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_1^2}(y_i - \beta x_i)^2\right)$$

For $i = n_1+1, \cdots, n$,

$$p(y_{n_1+1}, \cdots, y_n) = \prod_{i=n_1+1}^{n} \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_2^2}(y_i - \beta x_i)^2\right)$$

So,

$$L(\beta, \sigma_1, \sigma_2 \mid \underset{\sim}{y}) = \prod_{i=1}^{n} p(y_i \mid \beta, \sigma_1, \sigma_2)$$

$$= \prod_{i=1}^{n_1} \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_1^2}(y_i - \beta x_i)^2\right] \cdot \prod_{i=n_1+1}^{n} \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_2^2}(y_i - \beta x_i)^2\right]$$

$$= \sigma_1^{-n_1}\sigma_2^{-(n-n_1)}(2\pi)^{-n/2} \exp\left[-\frac{1}{2\sigma_1^2}\sum_1^{n_1}(y_i - \beta x_i)^2\right] \cdot \exp\left[-\frac{1}{2\sigma_2^2}\sum_{i=n_1+1}^{n}(y_i - \beta x_i)^2\right]$$

$$= \sigma_1^{-n_1}\sigma_2^{-(n-n_1)}(2\pi)^{-n/2} \cdot \exp\left[-\frac{1}{2}\left\{\sigma_1^{-2}\sum_1^{n_1}(y_i - \beta x_i)^2 + \sigma_2^{-2}\sum_{i=n_1+1}^{n}(y_i - \beta x_i)^2\right\}\right]$$

(b)

$$\log L = -\frac{n_1}{2}\log\sigma_1^2 - \frac{(n-n_1)}{2}\log\sigma_2^2 - \frac{n}{2}\log(2\pi) - \frac{1}{2}\left\{\sigma_1^{-2}\sum_{i=1}^{n_1}(y_i - \beta x_i)^2 + \sigma_2^{-2}\sum_{i=n_1+1}^{n}(y_i - \beta x_i)^2\right\}$$

$$\frac{\partial \log L}{\partial \beta} = -\frac{1}{2}\left\{\sigma_1^{-2}\sum_{i=1}^{n_1} 2(-x_i)(y_i - \beta x_i) + \sigma_2^{-2}\sum_{i=n_1+1}^{n} 2(-x_i)(y_i - \beta x_i)\right\}$$

$$= 0 \qquad \text{(i)}$$

$$\frac{\partial \log L}{\partial \sigma_1^2} = \frac{-n_1}{2\sigma_1^2} - \left(\tfrac{1}{2}\right)(-1)(\sigma_1^{-4})\sum_{i=1}^{n_1}(y_i-\beta x_i)^2$$

$$= \tfrac{1}{2}\left[-n_1\sigma_1^{-2} + \sigma_1^{-4}\sum_{i=1}^{n_1}(y_i-\beta x_i)^2\right]$$

$$= 0$$

$$\Rightarrow \hat{\sigma}_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(y_i-\hat{\beta}x_i)^2 \quad \#.$$

Similarly, taking $\frac{\partial \log L}{\partial \sigma_2^2}$,

$$\hat{\sigma}_2^2 = \frac{1}{n-n_1}\sum_{i=n_1+1}^{n}(y_i-\hat{\beta}x_i)^2 \quad \#$$

Now, from (i)

$$\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1}x_i(y_i-\beta x_i) + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n}x_i(y_i-\beta x_i) = 0$$

So, $\frac{1}{\sigma_1^2}\left[\sum_{i=1}^{n_1}x_iy_i - \beta\sum_{i=1}^{n_1}x_i^2\right]$

$$+ 1/\sigma_2^2\left[\sum_{i=n_1+1}^{n}x_iy_i - \beta\sum_{i=n_1+1}^{n}x_i^2\right] = 0$$

or, $$\hat{\beta}\left\{\frac{\sum_{i=1}^{n_1}x_i^2}{\sigma_1^2} + \frac{\sum_{i=n_1+1}^{n}x_i^2}{\sigma_2^2}\right\} = \left\{\frac{\sum_{i=1}^{n_1}x_iy_i}{\sigma_1^2} + \frac{\sum_{i=n_1+1}^{n}x_iy_i}{\sigma_2^2}\right\}$$

or, $$\hat{\beta} = \frac{\left[\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1}x_iy_i + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n}x_iy_i\right]}{\left[\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1}x_i^2 + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n}x_i^2\right]}$$

(This is "pooled" least squares)

(c) $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_A : \sigma_1^2 \neq \sigma_2^2$

$$\log \tilde{L}_0 = -\tfrac{n_1}{2}\log\hat{\sigma}_0^2 - \left(\tfrac{n-n_1}{2}\right)\log\hat{\sigma}_0^2 - n/2\log 2\pi$$

$$-1/2\left\{\frac{1}{\hat{\sigma}_0^2}\sum_{i=1}^{n_1}(y_i-\hat{\beta}_0x_i)^2 + \frac{1}{\hat{\sigma}_0^2}\sum_{i=n_1+1}^{n}(y_i-\hat{\beta}_0x_i)^2\right\}$$

$$= \tfrac{n}{2}\log\hat{\sigma}_0^2 - n/2\log 2\pi - \frac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(y_i-\hat{\beta}_0x_i)^2.$$

$$\log \tilde{L}_1 = -\tfrac{n_1}{2}\log\hat{\sigma}_1^2 - \left(\tfrac{n-n_1}{2}\right)\log\hat{\sigma}_2^2 - n/2\log 2\pi$$

$$-1/2\left[\frac{1}{\hat{\sigma}_1^2}\sum_{i=1}^{n_1}(y_i-\hat{\beta}x_i)^2 + \frac{1}{\hat{\sigma}_2^2}\sum_{i=n_1+1}^{n}(y_i-\hat{\beta}x_i)^2\right]$$

NB : $\hat{\beta}$ is the "unrestricted" estimator obtained in (b) above ; $\hat{\beta}_0$ is $\sum_{i=1}^{n}x_iy_i / \sum_{i=1}^{n}x_i^2$; &
$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{\beta}_0x_i)^2$.

$$LRT = -2\log\left(\frac{\tilde{L}_0}{\tilde{L}_1}\right)$$
$$= 2\left(\log\tilde{L}_1 - \log\tilde{L}_0\right)$$
$$= -n_1\log\hat{\sigma}_1^2 - (n-n_1)\log\hat{\sigma}_2^2 - n\log 2\pi$$
$$- \left[\frac{1}{\hat{\sigma}_1^2}\sum_{i=1}^{\hat{}}(y_i - \hat{\beta}x_i)^2 + \frac{1}{\hat{\sigma}_2^2}\sum_{i=n_1+1}^{\hat{}}(y_i - \hat{\beta}x_i)^2\right]$$
$$- n\log\hat{\sigma}_0^2 + n\log 2\pi + \frac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^{\hat{}}(y_i - \hat{\beta}_0 x_i)^2$$

(we can cancel the "$n\log 2\pi$" terms).

(d) $LRT \xrightarrow{d} \chi^2_{(1)}$. Reject H_0 if LRT exceeds $\chi^2_{(1)}$ crit. value.

Q.4.

(a) $$p(\theta|x) \propto p(\theta)p(x|\theta)$$
$$\propto \theta^{a-1}(1-\theta)^{b-1}\,{}^nC_x\,\theta^x(1-\theta)^{n-x}$$
$$\propto \theta^{a+x-1}(1-\theta)^{n+b-x-1}$$

Note that this is a Beta distribution with parameters $(a+x)$ and $(n+b-x)$.

(b) $$E(\theta|x) = \left(\frac{a+x+1-1}{a+x+n+b-x+1-1}\right)E(\theta^0|x)$$
$$= \left(\frac{a+x}{a+n+b}\right)E(1|x)$$
$$= \left(\frac{a+x}{a+n+b}\right)$$

So, this Bayes estimator is
$$\hat{\theta} = \left(\frac{a+x}{a+n+b}\right)$$

(c) $p(\theta) = k\,\theta^{a-1}(1-\theta)^{b-1}$

$$\partial p(\theta)/\partial\theta = k\left[(a-1)\theta^{a-2}(1-\theta)^{b-1} - (b-1)\theta^{a-1}(1-\theta)^{b-2}\right] = 0$$

$$\Rightarrow (a-1)\theta^{a-2}(1-\theta)^{b-1} = (b-1)\theta^{a-1}(1-\theta)^{b-2}$$

$$\Rightarrow (a-1)(1-\theta) = (b-1)\theta$$
$$\Rightarrow (a-1) - (a-1)\theta = (b-1)\theta$$
$$\Rightarrow a-1 = (a+b-2)\theta$$
$$\Rightarrow \tilde{\hat{\theta}} = \left(\frac{a-1}{a+b-2}\right),$$

& this is the mode of the prior.

Note that $0 \le \theta \le 1$, so the mode of $p(\theta)$ cannot be negative. So, we require that

$$\begin{cases} a > 1 \\ a+b > 2 \end{cases}$$

(note that we already know that $a, b > 0$.)

(d) We need the mode of the posterior in this case. Replace 'a' by $(a+x)$ and replace 'b' by $(n+b-x)$. So the Bayes estimator is $\hat{\theta} = \left[\frac{a+x-1}{(a+n+b-2)}\right]$.

(e) $E(\hat{\theta}) = \dfrac{a + E(x)}{a+n+b}$

$= \dfrac{a+n\theta}{a+n+b}$

& $\text{Bias}(\hat{\theta}) = \left[E(\hat{\theta}) - \theta\right] = \dfrac{a+n\theta - a\theta - n\theta - b\theta}{a+n+b}$

$= \left[\dfrac{a(1-\theta) - b\theta}{a+n+b}\right]$

$\text{Bias}(\hat{\hat{\theta}}) = \left[E(\hat{\hat{\theta}}) - \theta\right] =$

$= \left[\dfrac{a+n\theta - 1}{a+n+b-2}\right] - \theta$

$= \left[\dfrac{a+n\theta - 1 - a\theta - n\theta - b\theta - 2\theta}{a+n+b-2}\right]$

$= \left[\dfrac{a(1-\theta) - b\theta - 2\theta - 1}{a+n+b-2}\right]$

(Note that both biases $\to 0$ as $n \to \infty$, as expected as Bayes' estimators are consistent.

Bayesians aren't interested in what happens "on average" in repeated sampling. So properties such as unbiasedness are of no concern. They are interested in an estimator that is optimal for a given sample.

Q.5 (a) Here, we have "count" data – the dependent variable clearly takes only non-negative integer values: 0, 1, 2, ---. etc. If we used OLS, this is just MLE assuming Normal errors, & the data cannot be normal. So, OLS would be MLE with the wrong data density, & hence the wrong likelihood fctn. OLS will be inconsistent, etc.

(b) Obvious candidates are the Poisson or NegBin regression models. The former requires that the mean and variances of the data are the same (in the population). The latter allows for "over-dispersion". In our sample, the mean is 1.17 and the variance is $(1.51)^2$. We have over-dispersion & we might anticipate that the NegBin model will be preferred to the Poisson model.

(c) We see that the marginal effects associated with each dummy variable are positive – the signs of the marginal effects will be same as the signs (but not values) of the coefficients. More crises occur under "hard peg" & "floating" currency regimes than under other types. High or upper-middle income countries are more likely to experience currency crises than lower-income countries. These effects are all statistically significant in this Poisson model.

(d) LR is the LRT statistic for testing that $H_0: C(2) = c(3) = c(4) = 0$. That is, that there is no model here. The p-value is zero, so we clearly reject H_0. That is, we have a significant model.

(e) Now the NegBin model has been used instead of the Poisson model. The results are very similar to those obtained in Output 1. The NegBin model has been used to allow for any over-dispersion in the CRISIS data, as discussed in (b).

(f) Here we are conducting a formal test of over-dispersion, or more correctly of equi-dispersion. A Wald test is being used, as the equi-dispersion case is "nested" within the over-dispersion case. The null hypothesis is equi-dispersion, so if we cannot reject H_0 we would use the Poisson model. The p-value is 0.98, so we do not reject the Poisson model. (The absence of over-dispersion explains why the Poisson & NegBin results were so similar.)

(g) So now we know why OUTPUT 4 is based on the Poisson model! In the case of a dummy variable, we calculate its marginal effect by looking at the difference in the prediction of the conditional mean when DUM = 1, & when DUM = 0. If you think about it, this is exactly what happens with a dummy variable in an OLS model. So, here we could just take the difference between the mean values of LAM1 & LAM0: that is, $(1.429133 - 1.007608) = 0.421525$. Of course, this is <u>positive</u> as expected from the +ve coefficient of DUMINCHI in OUTPUT 1. The interpretation of this number is as follows: on average, high or upper-middle income countries had 0.42 more currency crises (over the period in the sample) than did low or lower-middle income countries, <u>ceteris paribus</u>.

<u>Note</u>: Although CRISIS takes integer values, its mean won't be expected to be an integer — having a value of 0.42 is quite sensible.