ECON 54 – Spring 2008

Solution for Final Exam

Q. 1.

(a) By independence,

$$L(\theta|\underset{\sim}{y}) = \prod_i p(y_i|\theta) = \theta^{-n}\exp\left[-\tfrac{1}{\theta}\textstyle\sum_i y_i\right]$$

& $\log L(\theta|y) = -n\log\theta - \frac{1}{\theta}\sum_i y_i$.

So, $(\partial \log L/\partial\theta) = -n/\theta + \frac{1}{\theta^2}\cdot n\bar{y} = 0$

$$\Rightarrow \tilde{\theta} = \bar{y} = \tfrac{1}{n}\textstyle\sum_i y_i.$$

S.O.C. : $(\partial^2\log L/\partial\theta^2) = \frac{n}{\theta^2} - \frac{2n\bar{y}}{\theta^3}$

& when $\theta = \bar{y}$, $(\partial^2\log L/\partial\theta^2) = \left(\frac{n}{\bar{y}^2} - \frac{2n}{\bar{y}^2}\right) < 0$. (Max.)

(b) LRT : $\log L(\tilde{\theta}_1) = -n\log\bar{y} - \frac{n\bar{y}}{\bar{y}} = -n(1+\log\bar{y})$

$$\log L(\tilde{\theta}_0) = -n\log(1) - \left(\tfrac{1}{1}\right)n\bar{y} = -n\bar{y}$$

So,
$$\begin{aligned} LRT &= -2\left[\log L(\tilde{\theta}_0) - \log L(\tilde{\theta}_1)\right] \\ &= -2\left[-n\bar{y} - \{-n(1+\log\bar{y})\}\right] \\ &= 2\left[n\bar{y} - n(1+\log\bar{y})\right] \\ &\xrightarrow{d} \chi^2_{(1)} \end{aligned}$$

So, reject $H_0: \theta = 1$ if LRT > critical value for $\chi^2_{(1)}$

Wald : $I = -E\{\partial^2\log L(\theta)/\partial\theta^2\}$

$$= -E\left[n/\theta^2 - \frac{2n\bar{y}}{\theta^3}\right]$$

Now, $E(y_i) = \theta$, so $E(\bar{y}) = \theta$ and

$$I = \left[-n/\theta^2 + \frac{2n\theta}{\theta^3}\right] = n/\theta^2.$$

$$IA = \lim_{n\to\infty}\left[\tfrac{1}{n}I(\theta)\right] = 1/\theta^2.$$

$$I^* = \left(n/\tilde{\theta}^2\right), \text{ so } \lim_{n\to\infty}\left[\tfrac{1}{n}I^*\right] = IA.$$

Wald statistic is $W = (\tilde{\theta}-1)^2\left(\frac{n}{\tilde{\theta}^2}\right)$; where $\tilde{\theta} = \bar{y}$.

So, $W = n(\bar{y}-1)^2/\bar{y}^2$.

Again, $W \xrightarrow{d} \chi^2_{(1)}$, so reject Ho if $W >$ crit. val.

(c) $D\log L_1 = -n/\theta + \frac{1}{\theta^2}n\bar{y}$

& $D\log L_1(\tilde{\theta}_0) = \frac{-n}{1} + \left(\frac{1}{1}\right)n\bar{y} = n(\bar{y}-1)$

So, $LM = n^2(\bar{y}-1)^2\left(\frac{\bar{y}^2}{n}\right) = n[(\bar{y}-1)\bar{y}]^2$

& when $n = 10$ & $\bar{y} = 1.5$, $LM = \cancel{40}\ 5.625$

We clearly reject $H_0: \theta = 1$ at the 10%, 5%, or 2.5% sig. levels. (Of course $n = 10$ hardly satisfies the asymptotic requirement.)

[Note: To see that $E(y_i) = \theta$:

$$E(y_i) = \frac{1}{\theta}\int_0^\infty y_i \, e^{-y_i/\theta} dy_i$$

Let $x = y_i/\theta$; $y_i = x\theta$; $dy_i = \theta dx$:

$$E(y_i) = \frac{1}{\theta}\int_0^\infty x\theta \, e^{-x}\theta dx$$

$$= \theta \int_0^\infty x e^{-x} dx = \theta \, \Gamma(2) = \theta \, \Gamma(1).1$$

$$= \theta . \quad]$$ See Q.3. for $\Gamma(.)$ integral.

Q.2.

(a) We are given the joint density, so

$$L(\beta | y) = c[v + (y - X\beta)'(y - X\beta)]^{-(n+v)/2}$$

& $\log L = \log c - (\frac{n+v}{2})\log[v + (y-X\beta)'(y-X\beta)]$

$$(\partial \log L / \partial\beta) = -(\tfrac{n+v}{2})[v + (y-X\beta)'(y-X\beta)]^{-1}.(2X'X\beta - 2X'y)$$

$$= 0$$

$$\Rightarrow 2X'X\beta - 2X'y = 0$$

$$\Rightarrow \tilde{\beta} = (X'X)^{-1}X'y \qquad (\text{OLS})$$

$\cancel{(b) \; (\partial^2 \log L / \partial\beta\partial\beta') = -(\tfrac{n+v}{2})\{2X'X[v + (y-X\beta)'(y-X\beta)]^{-1}}$

$\cancel{(2X'X\beta - 2X'y)}$

$\cancel{(2\beta'X'X - 2y'X)[v + (y-X\beta)'(y-X\beta)]^{-2}}$

(b) $(\partial^2 \log L / \partial\beta\partial\beta')$

$$= -(\tfrac{n+v}{2})\{2X'X[v + (y-X\beta)'(y-X\beta)]^{-1}$$

$$- 2(X'X\beta - X'y)[v + (y-X\beta)'(y-X\beta)]^{-2} \, 2(X'X\beta - X'y)$$

& when $\beta = \tilde{\beta}$, $X'X\tilde{\beta} = X'y$ & the 2nd term vanishes, so the evaluated Hessian is

$$-(\tfrac{n+v}{2})\, 2X'X[v + (y-X\tilde{\beta})'(y-X\tilde{\beta})]^{-1}$$

$$= -(n+v)(X'X)/[v + ns^2].$$

(c) A "Wald-type" test would use the inverse of the negative of this estimated Hessian as a consistent estimator of the information matrix, so:

$$W = (R\tilde{\beta} - q)'[R(\tfrac{v+ns^2}{n+v})(X'X)^{-1}R']^{-1}(R\tilde{\beta} - q)$$

$$= (\tfrac{n+v}{v+ns^2})(R\tilde{\beta} - q)'[R(X'X)^{-1}R']^{-1}(R\tilde{\beta} - q)$$

$$\xrightarrow{d} \chi^2_{(J)}$$, where J = number of independent restrictions.

(d) If we are ignorant about β then

$p(\beta) \propto \text{constant}$, so $p(\beta|y) \propto L(\beta|y)$.

Under a zero-one loss we take the mode of the posterior, & hence the mode of the likelihood as the Bayes estimator. So, the Bayes estimator would be just $\bar{\beta}$.

Q.3. (a) $E(X) = bc$

$$E(X^2) = b^2c(c+1)$$

So, we equate
$$\begin{cases} \frac{1}{n}\sum_i X_i = \bar{X} = \hat{b}\hat{c} & \text{(i)} \\ \frac{1}{n}\sum_i X_i^2 = \hat{b}^2\hat{c}(\hat{c}+1) & \text{(ii)} \end{cases}$$

& solve for $\hat{b}$ & $\hat{c}$:

From (i): $\hat{b} = \bar{X}/\hat{c}$

In (ii): $m_2' = (\bar{X}/\hat{c})^2(\hat{c}^2+\hat{c})$; where $m_2' = \frac{1}{n}\sum X_i^2$.

So, $m_2' = \bar{X}^2 + \bar{X}/\hat{c}$

or, $\hat{c}(m_2' - \bar{X}^2) = \bar{X}$

or, $\hat{c} = \bar{X} / [\frac{1}{n}\sum_i x_i^2 - \bar{x}^2]$

& $\hat{b} = [\frac{1}{n}\sum_i x_i^2 - \bar{x}^2]$.

(b) Both estimators are weakly consistent. Actually, as the Gamma distribution is a member of the exponential family, these MOM estimators will also be asymptotically efficient.

Q.4

(a) $\tilde{\beta} = [A+X'X]^{-1}[A\bar{\beta}+X'y]$

$$E(\tilde{\beta}) = [A+X'X]^{-1}[A\bar{\beta} + X'E(y)]$$

$$= [A+X'X]^{-1}[A\bar{\beta} + X'X\beta]$$

$\neq \beta$; hence Biased.

Bayesians don't worry about repeated sampling, so they are not concerned with what happens "on average". They just want to optimize their estimator relative to the given sample.

(b) Let $\sigma^2 A^{-1} = \sigma^2(X'X)^{-1}$

Then $\tilde{\beta} = [X'X + X'X]^{-1}[X'X\bar{\beta} + X'y]$

& $E(\tilde{\beta}) = [X'X + X'X]^{-1}[X'X\bar{\beta} + X'X\beta]$

$$= \frac{1}{2}(X'X)^{-1}(X'X)(\bar{\beta}+\beta)$$

$$= \frac{1}{2}(\bar{\beta}+\beta).$$

(c) (i) If $n \to \infty$, the estimator converges to MLE = OLS. (Normal errors.)

(ii) If $A^{-1} \uparrow \infty$, the prior information becomes diffuse as the prior variance $\uparrow \infty$. So, Bayes estimator $\to$ MLE again.

Note, $A^{-1} \uparrow \infty \Rightarrow A \to 0$

$$\Rightarrow \tilde{\beta} = [0 + X'X]^{-1}[0\bar{\beta} + X'y]$$

$$= (X'X)^{-1}X'y = \text{MLE}.$$

Q.5.

(a) $L(\beta|\underset{\sim}{y}) = \prod_{i=1}^{n} \exp(-\lambda_i)\lambda_i^{y_i}/y_i!$

$$= \frac{\exp[-\sum_i \lambda_i] \prod_i \lambda_i^{y_i}}{\prod_i y_i!}$$

where $\lambda_i = \exp(x_i'\beta)$.

& $\log L = -\sum_i \lambda_i + \sum_i y_i \log \lambda_i - \sum_i \log y_i!$

$$\frac{\partial \log L}{\partial \beta} = -\sum_i \frac{\partial \lambda_i}{\partial \beta} + \sum_i (\partial \log \lambda_i / \partial \beta)$$

$$= -\sum_i \exp(x_i'\beta)x_i + \sum_i y_i x_i$$

$$= \sum_i [y_i - \exp(x_i'\beta)]x_i$$

& setting this derivative to zero gives us the "likelihood equations" that we have to solve.

(b) The j'th. marginal effect is

$$\partial E(y_i|x_i)/\partial\beta_j$$

$$= \partial[\exp(x_i'\beta)]/\partial\beta_j$$

$$= \exp(x_i'\beta).\beta_j.$$

(c) (i) Evaluate $\exp(x_i'\beta)$ with the dummy variable set to unity & all other variables set to their sample medians (or means).

(ii) Repeat the above but with the dummy variable in question set to zero.

Take the difference between (i) & (ii) — this is the change in the conditional mean as the dummy changes, & is the desired marginal effect (estimated at a representative point in the sample).

Q.6 (a) Poisson regression because the dependent variable comprises "count" data – i.e. non-negative integers.
The model has been estimated by MLE.

(b) $H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$

vs H_A: Not H_0.

$LRT \xrightarrow{d} \chi^2_5$. The p-value is essentially zero, so we reject the null hypothesis.
That is, the covariates are jointly significant.

(c) $ME_i = \exp(x_i'\beta)\beta_j$. Evaluate at medians. So,

$$ME = \exp[(-1.122311) - (0.006751 * 81.295) + (0.232016 * 3) + 0] \times 0.232016$$
$$= 0.0875$$

So a 1-unit (\$10,000) increase in income will lead to an ↑ change of 0.09 bad credit reports. (This is neglible.)

(d) Variance > mean ⇒ over-dispersion & suggests that Poisson model should be discarded.
Perhaps use the NegBin model.

(e) NegBin – you can tell from the presence of the estimated shape parameter used to account for over-dispersion.

(f) Testing H_0: Poisson vs. H_A: NegBin.

Wald test of the single restriction that

exp(c(5)) = 0. If true, ⇒ Poisson.

The p-value of the test is 2.29%, so I'd reject H_0 & stay with the NegBin model.

N.B. LRT for Poisson vs NegBin –

$$LRT = 2[-66.09688 - (-77.83257)]$$
$$= 23.471 \xrightarrow{d} \chi^2_{(1)}$$

So easily reject H_0: Poisson again.