ECON 546 – Spring 2007

Final Exam – Solution

Q.1.: In terms of advantages, Bayes estimators share the consistency & asymptotic efficiency of MLE. They don't share the invariance property. On the other hand, Bayes estimators are admissible (if a "proper" prior is used), whereas MLE's are often inadmissible. Both approaches require that we specify the form of the likelihood, & hence the underlying distribution. In addition, Bayesians also have to specify a prior for the parameters. In both cases, problems arise if the likelihood is mis-specified.

Q.2.

(a) The likelihood function is

$$L = \prod_{i=1}^{n} K y_m^{K} / (y_i^{K+1})$$

$$= (K^n y_m^{nK} / \prod_i y_i^{K+1}).$$

The prior for 'K' is:

$$p(K) = K^{\alpha-1} e^{-K/\theta} / [\theta^{\alpha} \Gamma(\alpha)]$$

$$\propto K^{\alpha-1} e^{-K/\theta} \theta^{-K}.$$

So, by Bayes' Theorem, the posterior for θ is:

$$p(K|\underset{\sim}{y}) \propto K^{\alpha-1} e^{-K/\theta} \theta^{-\alpha} . K^n y_m^{nK} / \prod_i y_i^{K+1}$$

or,

$$p(\kappa \mid \underset{\sim}{y}) \propto \frac{\kappa^{n+\alpha-1} y_m^{n\kappa} e^{-\kappa/\theta}}{\exp[\log \prod_i y_i^{\kappa+1}]}$$

$$\propto \kappa^{(n+\alpha)-1} y_m^{n\kappa} e^{-\kappa/\theta} e^{-(\kappa+1)\sum_i \log y_i}$$

$$\propto \kappa^{(n+\alpha)-1} y_m^{n\kappa} e^{-\kappa\theta} e^{-\kappa \sum_i \log y_i}$$

$$\propto \kappa^{(n+\alpha)-1} e^{-\kappa/\theta} \exp[\log y_m^{n\kappa}] e^{-\kappa \sum_i \log y_i}$$

$$\propto \kappa^{(n+\alpha)-1} e^{-\kappa/\theta} e^{\kappa \sum_i \log y_m} e^{-\kappa \sum \log y_i}$$

$$\propto \kappa^{(n+\alpha)-1} e^{-\kappa/\theta} e^{-\kappa \sum_i \log(y_i/y_m)}$$

$$\propto \kappa^{(n+\alpha)-1} e^{-\kappa[1/\theta + \sum \log(y_i/y_m)]}$$

& this is just the kernel of a Gamma density with parameters $[n+\alpha]$ and $[1/\theta + \sum \ln(y_i/y_m)]^{-1}$.

(b) Under quadratic loss we use the mean of the posterior as the Bayes estimator. That is, we use

$$\hat{\kappa} = (n+\alpha)\left(1/\theta + \sum_i \log(y_i/y_m)\right)^{-1}.$$

Under a 0-1 loss function we use the mode of the posterior:

$$\hat{\kappa} = (n+\alpha-1)\left(1/\theta + \sum_i \log(y_i/y_m)\right)^{-1}.$$

(c) $\log L = n\log k + nk\log y_m - (k+1)\sum_i \log y_i$

$$\frac{\partial \log L}{\partial k} = \frac{n}{k} + n\log y_m - \sum_i \log y_i = 0$$

$$\Rightarrow n/\tilde{k} = \sum_i \log y_i - n\log y_m$$

$$= \sum_i \log y_i - \sum_i \log y_m$$

$$= \sum_i \log(y_i/y_m)$$

$$\text{\& so } \tilde{k} = n / \left[\sum_i \log(y_i/y_m)\right].$$

$$\frac{\partial^2 \log L}{\partial k^2} = -n/k^2 < 0$$; so we have a global maximum at $\tilde{k}$.

(d) Suppose that $\theta \to \infty$ and $\alpha \to 0$. Then

$$\hat{k} \to (n+0)\left[\frac{1}{\infty} + \sum_i \log(y_i/y_m)\right]^{-1}$$

$$\to n / \sum_i \log(y_i/y_m) = \tilde{k}.$$

& ~~$\hat{k} \to (A)$~~

This makes sense because in this case the variance of the prior $\to \infty$ and the mean is indeterminate. That is, we then have a diffuse prior, so Bayes $\Leftrightarrow$ MLE.

Q.3. (a) From Question 2,

$$\log L = n \log K + nK \log y_m - (K+1) \sum_i \log y_i$$

The unrestricted maximized log-likelihood is

$$\log \tilde{L}_u = n \log \tilde{K} + n\tilde{K} \log y_m - (\tilde{K}+1) \sum_i \log y_i$$

and the restricted maximized log-likelihood is

$$\log \tilde{L}_R = n \log(2) + 2n \log y_m - 3 \sum_i \log y_i$$

The LRT statistic is

$$LRT = 2 [\log \tilde{L}_u - \log \tilde{L}_R]$$

$$= 2 [n [\log(\tilde{K}/2)] + n \log y_m (\tilde{K} - 2) + (2-\tilde{K}) \sum_i \log y_i]$$

Given the data values, $\tilde{K} = \dfrac{100}{25 - 0} = 4$; &

$$LRT = 2 [100 \log(2) + 0 + (2-4) \cdot 25]$$
$$= 2 (69.3147 - 50)$$
$$= 38.629. \xrightarrow{d} \chi^2(1)$$

So, we easily reject H_0. So, I'd conclude that K is significantly greater than 2 (as $\tilde{K} = 4$) & so the variance exists.

(b) The information matrix is $I = -E\left[\frac{\partial^2 \log L}{\partial k^2}\right]$,

so $I = -E\left[-n/k^2\right] = n/k^2$, and

$IA = \lim_{n\to\infty}\left(\frac{1}{n} I\right) = 1/k^2$. So,

I^* is such that $\text{plim}(1/n\, I^*) = IA$, so that

$I^* = (n/\tilde{k}^2)$, where $\tilde{k}$ is the MLE for k.

The LM test statistic uses $D \log L_u(\tilde{k}_R)$, which is $(n/2 + n \log y_m - \sum_i \log y_i)$.

So, $LM = \left(\frac{n}{2} + n \log y_m - \sum_i \log y_i\right)(n/4)^{-1}$

& for the given data values, $LM = 25$.

We again strongly reject $H_0: k=2$. As $\tilde{k}=4$ it seems that $k > 2 > 1$ & so the mean exists.

(c) By invariance, the MLE is

$$\left(y_m 2^{1/\tilde{k}}\right)/y_m = 2^{1/\tilde{k}} = 2^{1/2} = 1.414$$

This estimator will be consistent & asymptotically efficient.

Q. 4.

(a) In a regression model,

$$y_i = x_i'\beta + \varepsilon_i \quad ; \quad i = 1, \ldots, n$$

& $E(y_i \mid x_i) = x_i'\beta$

The (conditional) mean of the data is a function of the covariates. In the Poisson model the mean is μ_i, so we set μ_i as a function of the covariates. The exponential fcn. ensures $\mu_i > 0$ (the mean of non-negative values cannot be negative). It also ensures that the variance ($= \mu_i$) is positive.

(b) $$\mathcal{L} = \prod_i \exp\left[-\exp(x_i'\beta)\right]\left[\exp(x_i'\beta)\right]^{y_i} / y_i!$$

~~log L~~ $$= \frac{\exp\left[-\sum_i \exp(x_i'\beta)\right] \prod_i \exp(x_i'\beta)^{y_i}}{\prod_i y_i!}$$

$$\log L = -\sum_i \exp(x_i'\beta) + \sum_i y_i x_i'\beta + \text{constant}.$$

$$\frac{\partial \log L}{\partial \beta} = -\sum_i x_i(\exp(x_i'\beta)) + \sum_i y_i x_i = 0$$

$$= \sum_i \left[y_i - \exp(x_i'\beta)\right] x_i = 0.$$

(Remember that y_i & $\exp(x_i'\beta)$ are <u>scalars</u>.)

(c) Recall that $\exp(x_i'\beta) = \mu_i$. So when we estimate β we have an estimate of the conditional mean, μ_i. The difference between the actual y_i & the estimated mean of the y_i's corresponds to $(y_i - \hat{y}_i)$ in the linear regression model. i.e. the i'th residual.

(d) Consider the first-order condition:

$$\sum_i [y_i - \exp(x_i'\beta)]x_i = 0.$$

If there is an intercept in the model, one element of the x_i vector is '1'. So the condition becomes:

$$\sum_i (y_i - \exp(x_i'\beta)) \begin{bmatrix} 1 \\ x \\ x \\ x \\ x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

When we replace β with $\tilde{\beta}$, we get

$$\sum_i (y_i - \exp(x_i'\tilde{\beta})) \begin{pmatrix} 1 \\ x \\ x \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

∴ from the first row, $\sum_i$ residuals $= 0$.

(e) $$\frac{\partial E(y_i | x_i)}{\partial x_{ij}} = \frac{\partial \exp(x_i'\beta)}{\partial x_{ij}} = \exp(x_i'\beta) \cdot \beta_j.$$

(f) $\frac{1}{n}\sum_i \exp(x_i'\tilde{\beta})\tilde{\beta}_j$

$$= -\tilde{\beta}_j \frac{1}{n}\sum_i\{[y_i - \exp(x_i'\beta)] - y_i\}$$

$$= -\{\tilde{\beta}_j \frac{1}{n}(0 - \sum_i y_i)\}$$

if an intercept is in the model.

So, $\frac{1}{n}\sum_i \exp(x_i'\tilde{\beta})\tilde{\beta}_j = \tilde{\beta}_j \frac{1}{n}\sum_i y_i = \tilde{\beta}_j \bar{y}$.

Q.5.

(a) A system of equations with exogenous regressors but where the errors across the equations are contemporaneously correlated:

$$\begin{aligned} y_1 &= X_1\beta_1 + \varepsilon_1 \\ y_2 &= X_2\beta_2 + \varepsilon_2 \\ &\vdots \\ y_m &= X_m\beta_m + \varepsilon_m \end{aligned} \qquad \text{or} \qquad y = X\beta + \varepsilon$$

& $V(\varepsilon) = \Omega = \Sigma \otimes I_T$.

If we allow for the error covariance structure and use the SUR/GLS estimator this will yield more efficient estimates than OLS. (Exceptions appear in part (b).) Typically we will have to obtain a consistent estimator of Σ first.

(b) (i) If $X_1 = X_2 = \cdots = Z$, say.

(ii) If Σ is diagonal (no covariances <u>across</u> equations' errors.)

(c) Suppose that <u>all</u> of the columns of X_1 are included among the columns of X_2:

$$X_2 = (X_1 , W) \text{ say.}$$

Then $X_1 = (X, W)\begin{pmatrix} I \\ 0 \end{pmatrix} = X_2 S$, say.

In this case,

$$X_2(X_2'X_2)^{-1}X_2'X_1 = X_2(X_2'X_2)^{-1}X_2'X_2 S$$
$$= X_2 S \quad = X_1$$

& the condition is satisfied.

Q.6

(a) The budget shares must sum to unity: $\sum_i w_i = 1$, for <u>any</u> data values.

So, $\sum_i \alpha_i + (\sum_i \beta_i)\log(M/P) + \sum_i\left[\sum_j \gamma_{ij} \log P_j\right] + \sum_i \varepsilon_i$

$= 0$.

This can happen for <u>any</u> data values only if

$$\sum_i \alpha_i = 1 \; ; \; \sum_i \beta_i = 0 \; ; \; \sum_i \gamma_{ij} = 0 \quad (\forall i)$$

(b) From (a), we also need $\sum_i \varepsilon_i = 0$. But this implies that the covariance matrix for the errors (across the equations) must be singular. We can drop any equation so that rank$(\Sigma) = (n-1)$ = number of equations. Which equation we delete won't matter.

(c) Iterative SUR (~~estim~~ iterated feasible GLS) has been used. It is appropriate as we have a system of equations, as long as the regressors are exogenous. Usually income (M) and prices are taken as <u>given</u>. So this is O.K. It would not be appropriate if M were endogenous. Usually SUR collapses to OLS if we have the same regressors. But that is for the <u>linear</u> SUR model. This model is non-linear in the parameters. OLS differs from SUR. The OLS estimates will be asymptotically inefficient.

(d) $\sum_i \tilde{\beta}_i = 0$; $\tilde{\beta}_1 = 2.78 \times 10^{-5}$, $\tilde{\beta}_2 = -2.45 \times 10^{-5}$

so $\tilde{\beta}_3 = (0 - \tilde{\beta}_1 - \tilde{\beta}_2) = -0.33 \times 10^{-5}$

$\sum_i w_i = 1$ so $\sum_i \bar{w}_i = 1$

& $\bar{w}_3 = (1 - \bar{w}_1 - \bar{w}_2)$
$= (1 - 0.91249 - 0.06567)$
$= 0.02184.$

The estimated elasticity is

$$\left[1 + \frac{-0.33 * 10^{-5}}{0.02184}\right] = 0.99985.$$

(It is a normal but inferior good, which is a little surprising.)

(e) $W = 18.52047$; $p = 0.0001$.
We strongly reject H_0 (homogeneity). This is not good news in terms of the economic sense of the model. The test may have low power as $n = 31$. (Note that there is no invariance problem with the Wald test here — the model is non-linear, but the restrictions are linear.)

(f)
$$\begin{aligned} LRT &= 2\left[\log\tilde{L}_u - \log\tilde{L}_R\right] \\ &= 2\left[(T/2)\left(\log|\tilde{\Sigma}_R| - \log|\tilde{\Sigma}_u|\right)\right] \\ &= 31\left[\log(3.04 * 10^{-11}) - \log(1.91 * 10^{-11})\right] \\ &= 31\left[\log 3.04 - \log 1.91\right] \\ &= 14.407. \end{aligned}$$

& $LRT \xrightarrow{d} \chi^2_{(1)}$ so we easily reject H_0.