STAT 613 Final Test (2 hours) **May 10th, 2010**

NAME: **Total number of Marks: /50**

Answer all the questions in the exam (questions are on both sides of the paper).

Marks will be given for clarity of the solution. One line answers in general not get marks. Also solutions which appear to 'fish' for marks will not get marks.

As with many estimators that arise in statistics, in some of the questions below, there may not be explicit forms for the estimators.

Write your solutions in the question paper.

Good Luck, and have a relaxing break!

(1) Suppose that $\{Y_t\}$ are independent random variables where the regressors x_t have some influence on Y_t. Suppose we fit the Gumbel distribution, with the density functions,

$$f_t(y) = e^{-(y-\beta' x_t)} \exp(-e^{-(y-\beta' x_t)}) \qquad -\infty < y \infty$$

to $\{Y_t\}$. Let $\hat{\beta}$ be an estimator of β. Define a transformation of Y_t, such that if the above model were correct, then the transformed $\{Y_t\}$ would be close to independent, uniformly distributed random variables. [5]

(2) Let us suppose that $\hat{\theta}_T$ is an estimator of θ_0, and $\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{D} \mathcal{N}(0, V)$.

(i) Give an estimator of θ_0^3.

(ii) Using a Taylor expansion of $g(\theta)$ about θ_0, obtain the asymptotic distribution of the above estimator.

[5]

(3) Let $\{X_t\}_{t=1}^T$ be independent, identically distributed normal random variables with distribution $\mathcal{N}(\mu, \sigma^2)$, where $\mu > 0$. Suppose that μ and σ^2 are unknown.

A non-central t-distribution with 11 degrees of freedom

$$f(x; a) = C(11)\Big(1 + \frac{(x-a)^2}{11}\Big)^{-(11+1)/2},$$

where $C(\nu)$ is a finite constant which only depends on the degrees of freedom, is mistakenly fitted to the observations. [8]

(i) Suppose we construct the likelihood using the t-distribution with 11 degrees of freedom, to estimate a. In reality, what is this MLE actually estimating?

(ii) Denote the above ML estimator as $\hat{a}_T$. Assuming that standard regularity conditions are satisfied, what is the approximate distribution of $\hat{a}_T$?

(4) The survivial time of disease A follow an exponential distribution, where the distribution function has the form $f(x) = \lambda^{-1}\exp(-x/\lambda)$. Suppose that it is known that at least one third of all people who have disease A survive for more than 2 years. [8]

(i) Based on the above information derive a lower bound for λ.

(ii) Suppose that it is known that $\lambda \geq \lambda_0$. What is the maximum likelihood estimator of λ.

(iii) Derive the sampling properties of maximum likelihood estimator of λ, for the cases $\lambda = \lambda_0$ and $\lambda > \lambda_0$.

(5) Suppose that $\{Y_t\}$ are independent random variables with the canonical exponential distribution, whose logarithm satisfies

$$\log f(y;\theta_t) = \frac{y\theta_t - \kappa(\theta_t)}{\phi} + c(y;\phi),$$

where ϕ is the dispersion parameter. Let $E(Y_t) = \mu_t$. Let $\eta_t = \beta' x_t = \theta_t$ (hence the canonical link is used), where x_t are regressors which influence Y_t. [14]

(a) (i) Obtain the log-likelihood of $\{(Y_t, x_t)\}_{t=1}^{T}$.

(ii) Denote the log-likelihood of $\{(Y_t, x_t)\}_{t=1}^{T}$ as $\mathcal{L}_T(\beta)$. Show that

$$\frac{\partial \mathcal{L}_T}{\partial \beta_j} = \sum_{t=1}^{T} \frac{(Y_t - \mu_t)x_{t,j}}{\phi} \quad \text{and} \quad \frac{\partial^2 \mathcal{L}_T}{\partial \beta_k \partial \beta_j} = -\sum_{t=1}^{T} \frac{\kappa''(\theta_t)x_{t,j}x_{t,k}}{\phi}.$$

(b) Let Y_t have Gamma distribution, where the log density has the form

$$\log f(Y_t;\mu_t) \quad = \quad \frac{-Y_t/\mu_t - \log\mu_t}{\nu^{-1}} + \Big\{ -\frac{1}{\nu^{-1}}\log\nu^{-1} + \log\Gamma(\nu^{-1})\Big\} + \Big\{\nu^{-1} - 1\Big\}\log Y_t$$

$\mathrm{E}(Y_t) = \mu_t$, $\mathrm{var}(Y_t) = \mu_t/\nu^2$ and $\nu_t = \beta' x_t = g(\mu_t)$.

(i) What is the canonical link function for the Gamma distribution and write down the corresponding likelihood of $\{(Y_t, x_t)\}_{t=1}^T$.

(ii) Suppose that $\eta_t = \beta' x_t = \beta_0 + \beta_1 x_{t,1}$. Denote the likelihood as $\mathcal{L}_T(\beta_0, \beta_1)$. What are the first and second derivatives of $\mathcal{L}_T(\beta_0, \beta_1)$?

(iii) Evaluate the Fisher information matrix at β_0 and $\beta_1 = 0$.

(iv) Using your answers in (ii,iii) and the mle of β_0 with $\beta_1 = 0$, derive the score test for testing $H_0 : \beta_1 = 0$ against $H_A : \beta_1 \neq 0$.

(6) Let us suppose that $\{X_t\}$ are iid exponentially distributed random variables with density $f(x) = \frac{1}{\lambda}\exp(-x/\lambda)$. Suppose that we only observe $\{X_t\}$, if $X_t > c$ (else X_t is not observed). [10]

(i) Show that the sample mean $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$ is a biased estimator of λ.

(ii) Suppose that λ and c are unknown, derive the log-likelihood of $\{X_t\}_{t=1}^{T}$ and the maximum likelihood estimators of λ and c.