Statistics Ph.D. Qualifying Exam: Part II

November 9, 2002

Student Name: ____________________

1. Answer 8 out of 12 problems. Mark the problems you selected in the following table.

1	2	3	4	5	6	7	8	9	10	11	12

2. Write your answer right after each problem selected, attach more pages if necessary.

3. Assemble your work in right order and in the original problem order.

1. Let $\{X_1, \ldots, X_n\}$ be a random sample from the population with density $f(x, \theta)$, where $f(x, \theta)$ is given by: $f(x, \theta) = e^{-(x-\theta)}, x > \theta; = 0$ if otherwise. Denote by $\{Y_i = i(X_{(n-i+1)} - X_{(n-i)})$ if $i = 1, \ldots, n-1, Y_n = n(X_{(1)} - \theta)$, where $X_{(r)}$ is the $r - th$ order statistic, $r = 1, \ldots, n$.

   (a) Show that $\{Y_i, i = 1, \ldots, n\}$ are independently and identically distributed random variables with density $g(y) = e^{-y}, y > 0; = 0, y \leq 0$. Use this result to derive the UMVUE (Uniformly Minimum Varianced Unbiased estimator) of θ.

   (b) Suppose the density of the prior distribution is given by: $h(\theta) = e^{-(a-\theta)}, a > \theta; = 0, a \leq \theta$, where a is a known constant. Derive the Bayese estimator of θ and compare it with the UMVUE of θ.

2. Let $\{X_1, \ldots, X_n\}$ be a random sample from the population with density $g(x; \theta, \mu_1, \mu_2) = \theta f(x; \mu_1) + (1 - \theta) f(x; \mu_2)$, where $f(x; \mu)$ is the density of $N(\mu, 1)$ and $0 < \theta < 1$. Illustrate how to derive a procedure to compute the MLE (Maximum Likelihood Estimator) of $\{\theta, \mu_1, \mu_2\}$ by using the EM-algorithm.

3. Let $\{X_1, \ldots, X_n\}$ be a random sample from the population with density $f(x; \theta_i, i = 1, 2)$, where $f(x; \theta_i, i = 1, 2)$ is given by:

$$\begin{aligned} f(x; \theta_i, i = 1, 2) &= \frac{1}{\theta_2 - \theta_1} \text{ if } \theta_1 \leq x \leq \theta_2, \\ &= 0 \text{ if for otherwise.} \end{aligned}$$

   (a) Show that the statistics $\{X_{(1)} = Min(X_1, \ldots, X_n), X_{(n)} = Max(X_1, \ldots, X_n)\}$ are sufficient and complete statistics for the parameters $\{\theta_i, i = 1, 2\}$.

   (b) Derive the UMVUE of $\theta_2 - \theta_1$.

4. Let U be a Uniform $(0,1)$ random variable. Let λ a constant, such that $0 < \lambda < 1$, and let V be a random variable with support $(0,1)$ that is independent of U.

   (a) Prove that $\min\left(\frac{U}{\lambda}, \frac{1-U}{1-\lambda}\right)$ has a Uniform$(0,1)$ distribution.

   (b) Find $P(\min\left(\frac{U}{V}, \frac{1-U}{1-V}\right) > .5)$.

   (c) If $X \sim N(0,4)$ and $Z \sim N(0,1)$ with distribution function Φ. Find a relation between the distribution functions of $2\min(\Phi(X),(1-\Phi(X))$ and $2\min(\Phi(Z),(1-\Phi(Z))$

5. Let $Y_1, \ldots, Y_N$ be independent random variable such that $Y_i \sim \text{Binomial}(n_i, p_i)$, where

$$p_i = \frac{1}{1 + e^{\beta' x_i}},$$

and x_i are vectors of fixed covariates, $i = 1, \ldots, N$.

(a) Find a set of jointly sufficient statistics for β.

(b) Suppose $\tilde{\beta}$ is the estimate of β that minimizes

$$Q = \sum_{i=1}^{N} n_i \hat{p}_i (1 - \hat{p}_i)(x_i'\beta - l(\hat{p}_i))^2,$$

where $\hat{p}_i = Y_i/n_i$, and $l(\hat{p}_i) = \log(\frac{\hat{p}_i}{1-\hat{p}_i}), i = 1, \ldots, N$. Find $\tilde{\beta}$.

(c) Let $\hat{\beta}$ be MLE's of β. Compare the two methods of parametric estimation. Which is better. Sketch a proof of your statement.

6. Let X_1 and X_2 be independently and identically distributed from the following probability mass function

$$p(x) = P(X = x) = pq^x, \ x = 0, 1, 2, \ldots, \text{ where } 0 < p < 1 \text{ and } q = 1 - p.$$

(a) Find the UMVUE for $(1 + p)^2$.

(b) Find the UMVUE for $(1 + p)^2/p$.

7. Let $Y_1, Y_2, \ldots, Y_n, X_1, X_2$ be i.i.d. with continuous distribution function F. Define $Y_{(0)} = -\infty$, $Y_{(n+1)} = +\infty$ and for $j = 1, 2, \ldots, n$ let $Y_{(j)}$ be the jth smallest Y_j.

   (a) Find $P[Y_{(j)} \leq X_1 \leq Y_{(j+1)}]$ for $j = 0, 1, \ldots, n$.

   (b) Find $P[Y_{(j)} \leq X_1 \leq Y_{(j+1)} \leq X_2 \leq Y_{(j+2)}]$ for $j = 0, 1, \ldots, n-1$.

8. Consider independent random samples $X_{i1}, \ldots, X_{in}$ $(i = 1, 2)$ from the uniform populations:

$$f(x_i) = \frac{1}{\theta_i}, \quad 0 < x_i < \theta_i, \quad \theta_i > 0, \quad i = 1, 2.$$

Let $Y_i = \max(X_{i1}, \ldots, X_{in})$, $i = 1, 2$, and $Y = \max(Y_1, Y_2)$.

(a) Show that the likelihood ratio statistic to test the hypothesis $H_0 : \theta_1 = \theta_2$ is given by

$$Z = (Y_1 Y_2 / Y^2)^n.$$

(b) Obtain the exact distribution of $-2 \log Z$ under H_0.

9. Let X denote a nonnegative integer valued random variable. The function $\lambda(n) = P\{X = n | X \geq n\}$, $n \geq 0$, is called the discrete hazard rate function.

(a) Show that $P\{X = n\} = \lambda(n) \prod_{i=0}^{n-1} (1 - \lambda(i))$.

(b) Show that we can simulate X by generating independent uniform $(0, 1)$ random variables $U_1, U_2, \ldots$ stopping at $X = \min\{n : U_n \leq \lambda(n)\}$.

(c) Apply this method to simulating a geometric random variable. Explain why it works.

(d) Suppose that $\lambda(n) \leq p < 1$ for all n. Consider the following algorithm for simulating X and explain why it works: Simulate X_i, U_i, $i \geq 1$, where X_i is geometric with mean $1/p$ and U_i is a uniform $(0, 1)$ random variable. Set $S_k = X_1 + \ldots + X_k$ and let

$$X = \min\{S_k : U_k \leq \lambda(S_k)/p\}.$$

10. Suppose that Λ is distributed as V/s_0, where s_0 is some constant and $V \sim \chi^2(k)$, and that given $\Lambda = \lambda$, $X_1, \ldots, X_n$ are i.i.d. Poisson distributed with parameter λ. Let $T = \sum_{i=1}^n X_i$.

(a) Show that $(\Lambda | T = t)$ is distributed as W/s, where $s = s_0 + 2n$ and $W \sim \chi^2(m)$ with $m = k + 2t$.

(b) Show how quantiles of the χ^2 distribution can be used to determine level $(1 - \alpha)$ upper and lower limits for λ.

11. Let $(X_i, Y_i), i = 1, 2, \cdots, n$ be n i.i.d. random vectors, where $X_i \sim N(\theta, 1)$, $P(Y_i = 1) = p, P(Y_i = 0) = 1 - p$, and X_i and Y_i are independent. Suppose that we can *only* observe the product $Z_i = X_i Y_i, i = 1, 2, \cdots, n$.

   (a) Find a minimal sufficient statistics for (p, θ).

   (b) Find the maximum likelihood estimator of p and θ.

12. Let $X_1, \ldots, X_n$ be a random sample of size n from a population with density

$$f(x;\theta) = \frac{x}{\theta^2} e^{-x^2/(2\theta^2)}, \quad x > 0.$$

Let $Y_1, \ldots, Y_n$ be a random sample of size n from a population with density

$$g(y;\mu) = \frac{1}{\mu} e^{-x/\mu}, y > 0.$$

(a) Find the uniformly post powerful test for $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ based on sample $\{X_1, \ldots, X_n\}$.

(b) Find the uniformly post powerful test for $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ based on sample $\{Y_1, \ldots, Y_n\}$.

(c) Consider a transformation $Z_i = X_i^2$ and then draw a connection between questions (a) and (b).