Statistics Ph.D. Qualifying Exam: Part II

November 20, 2004

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 and X_2 be two independent random variables each with an exponential p.d.f.

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0, \lambda > 0.$$

Define $W = |X_1 - X_2|$ and $S = X_1 + X_2$.

(a) Find the joint p.d.f. of W and S.

(b) Find the marginal p.d.f. of W.

2. Let (X, Y, Z) be a random vector following a multinomial distribution

$$f(x, y, z) = \frac{n!}{x!y!z!(n-x-y-z)!} p^x q^y r^z (1-p-q-r)^{n-x-y-z},$$

for $x, y, z = 0, 1, 2, \cdots, n$ and $x + y + z \leq n$ where $0 \leq p, q, r, p + q + r \leq 1$.

(a) Find the moment generating function of X, Y, Z. That is, find $M(t_1, t_2, t_3) = E(e^{t_1 X + t_2 Y + t_3 Z})$.

(b) Derive the distribution of $X + Y + Z$.

(c) Derive the conditional distribution of X, Y given $Z = z$.

3. For each of the following p.d.f.s, let $X_1, X_2, \cdots, X_n$ be a random sample form that distribution, find the UMVUE of θ^r.

   (a) $f(x;\theta) = \frac{1}{\theta}, 0 < x < \theta, r < n.$

   (b) $f(x;\theta) = e^{-(x-\theta)}, x > \theta.$

4. Suppose that an random sample $X_1, ..., X_n$ taken from a normal distribution with mean θ and variance σ^2 and suppose that the prior distribution on θ is a normal distribution with mean μ and variance τ^2, where σ, μ and τ are known constants.

   (a) Find the joint pdf of $\bar{X}$ and θ.

   (b) Find the marginal pdf of $\bar{X}$.

   (c) Find the posterior pdf of θ given $\bar{X}$.

   (d) Find the Bayes estimator of θ under a square loss function.

5. Suppose that an random sample $X_1, ..., X_n$ taken from an exponential distribution with mean θ and another random sample $Y_1, ..., Y_m$ taken from exponential distribution with mean μ.

   (a) Derive the likelihood ratio test for $H_0 : \theta = \mu$ vs. $H_1 : \theta \neq \mu$.

   (b) Show the test above can be based on the statistics

   $$T = \frac{\sum X_i}{\sum Y_i}.$$

   (c) Find the sampling distribution of T when H_0 is true.

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

$$P_i = \frac{1}{1 + e^{\alpha+\beta x_i}},$$

and x_i is a fixed covariate, $i = 1, \ldots, N$.

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

(b) Find $\tilde{\alpha}, \tilde{\beta}$, the estimates of α and β that minimize

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

where $\hat{p}_i = \frac{Y_i}{n_i}$ and $l_i = \log(\frac{\hat{p}_i}{1-\hat{p}_i}), i = 1, \ldots, N$.

(c) Let $\hat{\alpha}$ and $\hat{\beta}$ represent the conditional expectations of $\tilde{\alpha}, \tilde{\beta}$ given the jointly sufficient statistics for α, β. Prove that for any constants a and b

$$\text{Mean Square}(a\hat{\alpha} + b\hat{\beta}) \leq \text{Mean Square}(a\tilde{\alpha} + b\tilde{\beta})$$

7. (a) Let $\{T_n\}$ be a sequence of random variables such that as $n \to \infty, \sqrt{n}(T_n - \theta)$ converges in distribution to $N(0, \sigma^2(\theta))$ for all $\theta \in \Omega$. Let $g(x)$ be a differentiable function with continuous derivative $g'(\theta)$ for all $\theta \in \Omega$. Show that as $n \to \infty, \sqrt{n}[g(T_n) - g(\theta)]$ converges to $N[0, \sigma^2(\theta)(g'(\theta))^2]$ for each $\theta \in \Omega$.

   (b) Let $X_1, X_2 \ldots$ be a sequence of random variables from $N(\mu, \sigma^2)$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$. Find the limiting distribution of $\sqrt{n}(S_n^2 - \sigma^2)$ and obtain a transformation $g(S_n^2)$ of S_n^2, whose distribution is independent of σ^2.

8. Consider the linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{Y}$ is $(n \times 1)$, $\boldsymbol{\epsilon}$ is $(n \times 1)$, $\boldsymbol{X}$ is $(n \times p)$, and where $E(\boldsymbol{\epsilon}) = \mathbf{0}$, $\mathrm{Cov}(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}$.

(a) Prove the Gauss Markov Theorem. That is, prove that the estimator $\boldsymbol{l}'\hat{\boldsymbol{\beta}}$ of the estimable function $\boldsymbol{l}'\boldsymbol{\beta}$ has the smallest variance among all unbiased estimators of $\boldsymbol{l}'\boldsymbol{\beta}$ that are linear functions of $\boldsymbol{Y}$. [$\hat{\boldsymbol{\beta}}$ is the least squares estimator of $\boldsymbol{\beta}$.]

(b) Extend the result in (a) to the case where $\mathrm{Cov}(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is a non-singular diagonal matrix.

9. Let $\{(X_{i,1}, \ldots, X_{i,n}), i = 1, 2\}$ be independent random samples from normal distributions with means μ_i and variance σ^2 respectively. Let $\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{i,j}, i = 1, 2$ and $S_i^2 = \sum_{j=1}^{n}(X_{i,j} - \bar{X}_i)^2, i = 1, 2$. Put $Y_i = \frac{\sqrt{n}\bar{X}_i}{\sqrt{\hat{\sigma}^2}}, i = 1, 2$, where $\hat{\sigma}^2 = (S_1^2 + S_2^2)/(2n-2)$.

   (a) Obtain the joint pdf (probability density function) of $\{Y_1, Y_2\}$ under the assumption ($\mu_i = 0, i = 1, 2$).

   (b) What is the pdf of $Y_1 - Y_2$ under the assumption $\mu_1 = \mu_2$?

10. Let $X_1, \dots, X_n$ be a random sample from the normal distribution with mean μ_1 and variance σ^2. Let $Y_1, \dots, Y_m$ be a random sample from the normal distribution with mean μ_2 and variance $4\sigma^2$. Consider the hypotheses $H_0 : \mu_1 - \mu_2 = 4$ versus $H_1 : \mu_1 - \mu_2 \neq 4$.

   (a) Derive the level-α Likelihood Ratio test for testing H_0 versus H_1.

   (b) What is the sampling distribution of your testing statistic under H_0?

11. Let $X_1, \ldots, X_n$ be a random sample from the density $f(x, \theta) = \frac{\log \theta}{\theta - 1} \theta^x, 0 < x < 1, \theta > 1$.

(a) Obtain a sufficient and complete statistic for θ.

(b) Find a function $\phi = \phi(\theta)$ of θ such that there is an unbiased estimator $\hat{\phi}$ of ϕ with variance $Var(\hat{\phi})$ achieving the Fisher-Cramer-Rao lower bound. Obtain $\hat{\phi}$.

(c) Is the estimator $\hat{\phi}$ obtained in (b) above the UMVUE (Uniformly Minimum Varainced Unbiased estimator) of ϕ? Why?

12. Consider the following regression model:

$$\begin{aligned} Y_j &= \theta_1 + \theta_2(x_j - \bar{x}) + \epsilon_j, j = 1, \ldots, n \\ \text{where} \quad \bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i, \end{aligned}$$

Assume that the ϵ_j's are independently distributed as normal random variables with means 0 and variance σ^2 and that the x_i's are non-stochastic.

(a) Assuming a non-informative prior for $\{\theta_i, i = 1, 2, \sigma^2)$ as $P(\theta_i, i = 1, 2, \sigma^2) \propto (\sigma^2)^{-1}$, derive the posterior distribution of $\{\theta_i, i = 1, 2\}$.

(b) Derive the posterior distribution of θ_2 and a $(1 - \alpha)$ % HPD (Highest Posterior Density) interval for θ_2. How is this HPD interval comparing with the $(1 - \alpha)$ % confidence interval for θ_2 ?