Statistics Ph.D. Qualifying Exam: Part II

November 19, 2005

Student Name: ______________________________

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

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

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. Consider the following transformation

$$X_1 = \sqrt{-2\,\ln(U_1)}\cos(2\pi\,U_2), \quad X_12 = \sqrt{-2\,\ln(U_1)}\sin(2\pi\,U_2),$$

where U_1, U_2 are i.i.d. random variables with $U(0,1)$ distribution. Prove that X_1 and X_2 are i.i.d. random variables with $N(0,1)$ distribution.

2. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a $U(\theta, 2\theta)$ distribution.

   (a) Find the method of moments estimator of θ.

   (b) Find the MLE of θ, $\hat{\theta}$, and find a constant k such that $E(k\hat{\theta}) = \theta$.

   (c) Compare the estimators found above.

3. Let $X_1, X_2, ..., X_n$ be i.i.d. random variables with p.d.f.

$$f(x;\theta) = \begin{cases} \frac{\theta x^{\theta-1}}{(1+x)^{\theta+1}}, & x > 0 \\ 0 \quad , & \text{elsewhere,} \end{cases}$$

where $\theta > 0$ is an unknown parameter.

(a) Find the distribution of $Y_1 = \frac{X_1}{1+X_1}$.

(b) Find the distribution of $Z_1 = -\ln(Y_1) = -\ln\left(\frac{X_1}{1+X_1}\right)$.

(c) Find an UMVUE (Uniformly Minimum Varianced Unbiased Estimator) for θ^{-1}.

4. Let $X_1, \ldots, X_m$ be a random sample from $N(\mu_1, a^2\sigma^2)$ and $Y_1, \ldots, Y_n$ a random sample from $N(\mu_2, b^2\sigma^2)$ where a^2 and b^2 are known positive numbers.

   (a) Obtain maximum likelihood estimators of μ_1, μ_2, and σ^2.

   (b) Derive the likelihood ratio test for testing $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$. What is the sampling distribution of your test statistic under H_0?

   (c) Obtain a 95% confidence interval for σ^2. $[z_{0.05} = 1.645, z_{0.025} = 1.960, z_{0.01} = 2.326, z_{0.005} = 2.576]$

5. Let $X_1, X_2, \ldots, X_n$ be be a random sample of size n from a Poisson distribution with the probability distribution function

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!}, x = 0, 1, \ldots$$

(a) Find the UMVUE (Uniformly Minimum Varianced Unbiased Estimator) of $\theta^2 e^{-\theta}$.

(b) Under either $H_0 : \theta = 20$ or $H_1 : \theta = 10$, explain why we do not require a large n to permit a reasonable normal approximation for $\sum_{i=1}^n X_i$.

(c) Given $H_0 : \theta = 20$ vs. $H_1 : \theta = 10$, find n to guarantee type I and type II error probabilities are less than 0.05. (i.e. $\alpha, \beta \leq 0.05$) $[z_{0.05} = 1.645, z_{0.025} = 1.960, z_{0.01} = 2.326, z_{0.005} = 2.576]$

6. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be Poisson random variables with X's parameter $\theta\lambda$ and Y's parameter λ. Let $X = \sum_{i=1}^{n} X_i$ and $Y = \sum_{i=1}^{m} Y_i$.

   (a) Find the MLE's of θ and λ.

   (b) Find the conditional distribution X given $X + Y = N$.

   (c) Using a sample of size n from the above conditional distribution, calculate the MLE of θ.

   (d) Compare the MLE's of θ obtained in (b) and (c), by their bias and mean square error. Which will you prefer?

7. Let $Y_1, \ldots, Y_n$ be independent random variables such that

$$Y_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j + \epsilon_i,$$

$i = 1, \ldots, n$, and $p < n$. Let $\mathbf{X}$ be the $n \times p$ matrix of x_{ij}'s, $\mathbf{1}$, the $n \times 1$ vector of 1's, β the $p \times 1$ vector of β_j's and ϵ, the vector of ϵ_i's.

(a) Derive an expression for $\hat{\beta}$, the least squares estimate of β.

(b) Prove that $\mathbf{c}'\hat{\beta}$ achieves the lowest variance among all unbiased estimators of $\mathbf{c}'\beta$ that are linear functions of $\mathbf{Y}$

8. Let $X_1, \ldots, X_n$ be a random sample from a Normal $(\mu, 1/\tau)$ population. Assume the following prior specifications on μ and τ: $\mu|\tau \sim N(\mu_0, \frac{1}{\lambda_0 \tau})$, $\tau \sim Gamma(\alpha_0, \beta_0)$.

   (a) Show that this prior specification is conjugate for this problem.

   (b) Find the posterior distribution of $\sqrt{\frac{\lambda_0 \alpha_0}{\beta_0}}(\mu - \mu_0)$.

9. Let $\{(X_{i,1}, \ldots, X_{i,n_i}), i = 1, 2 (n_i > 1)\}$ be independent random samples from normal distributions with means 0 and variance σ_i^2 respectively. Let $\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{i,j}, i = 1, 2$ and $S_i^2 = \sum_{j=1}^{n_i}(X_{i,j} - \bar{X}_i)^2, i = 1, 2$. Put $Y_1 = \frac{\sqrt{n_1}\bar{X}_1}{\sqrt{\hat{\sigma}_1^2}}, Y_2 = (n_1 - 1)S_2^2/((n_2 - 1)S_1^2)$, where $\hat{\sigma}_1^2 = S_1^2/(n_1 - 1)$.

(a) Obtain the joint pdf (probability density function) of $\{Y_1, Y_2\}$.

(b) What is the marginal pdf of Y_1?

(c) What is the marginal distribution of Y_2?

10. Let $\{(X_{i,1}, \ldots, X_{i,n_i}), i = 1,2,3\}$ be independent random samples from the density $f_i(x;\theta_i)$ $(i = 1,2,3)$ respectively, where $f_i(x;\theta_i)$ is given by

$$f_i(x, \theta_i) = \theta_i x^{\theta_i - 1}, 0 < x < 1, \theta_i > 0.$$

Put $Y_i = -\sum_{j=1}^{n_i} \log X_{i,j}$ $(i = 1,2,3)$.

(a) Derive the joint probability density function of $\{Z_1 = Y_2/Y_1, Z_2 = Y_3/Y_1\}$.

(b) Assuming that $\theta_2 = \theta_3$, show that the generalized likelihood ratio test for testing $H_0 : \theta_1 = \theta_2$ versus $H_1 : \theta_1 \neq \theta_2$ is based on the statistic $Z = Z_1 + Z_2$.

11. Let $\{X_1, \ldots, X_n\}(n > 20)$ be a random sample from a population with density $f(x, \Theta) = \sum_{i=1}^{3} \omega_i f_i(x; \mu_i, \sigma^2)$, where $\{\omega_1 = \theta^2, \omega_2 = 2\theta(1-\theta), \omega_3 = (1-\theta)^2 \quad (0 < \theta < 1)\}$ and $f_i(x; \mu_i, \sigma^2)$ is the density of a normal distribution with mean μ_i and variance σ^2. Let the prior distribution of $\Theta = (\theta, \mu_i, i = 1, 2, 3, \sigma^2)$ be given by the non-informative prior

$$P(\Theta) \propto (\sigma^2)^{-1}.$$

Illustrate how you will use the Gibbs sampling procedure to derive estimates of the parameters.

12. Let $\{X_1, \ldots, X_m\}$ be a random sample from the population with density $f(x; \theta_1, \sigma_1^2) = \frac{1}{\sigma_1^2} exp\{-\frac{1}{\sigma_1^2}(x - \theta_1)\}, x > \theta_1, \sigma_1^2 > 0$. Let $\{Y_1, \ldots, Y_n\}$ be a random sample from the population with density $f(y; \theta_2, \sigma_2^2) = \frac{1}{\sigma_2^2} e^{-\frac{1}{\sigma_2^2}(y-\theta_2)}, y > \theta_2, \sigma_2^2 > 0$, independently of $\{X_1, \ldots, X_m\}$.

   (a) Derive a set of sufficient and complete statistics for the parameters $\{\theta_i, \sigma_i^2, i = 1, 2\}$.

   (b) Derive the UMVUE (Uniformly Minimum Varianced Unbiased Estimator) of $\phi = \sigma_1^2/\sigma_2^2$.

   (c) Let the prior distribution of $\Omega = \{\theta_i, \sigma_i^2, i = 1, 2\}$ be given by the non-informative prior $P(\Omega) = \prod_{j=1}^{2} (\sigma_j^2)^{-1}$, derive the Bayesian estimator of ϕ under squared loss function.