Thursday, August 14, 2008

Mathematical Statistics Preliminary Examination

Statistics Group, Department of Mathematics and Statistics, Auburn University

Name: ____

1. It is a closed-book and in-class exam.
2. One page (letter size, 8.5-by-11in) cheat sheet is allowed.
3. Calculator is allowed. No laptop (or equivalent).
4. Show your work to receive full credits. *Highlight your final answer.*
5. Solve any **five** problems out of the seven problems.
6. Total points are **50**. Each question is worth **10** points.
7. If you work out more than five problems, your score is the sum of five highest points.
8. Time: **150** minutes. (9:00am–11:30am, Thursday, August 14, 2008)

1	2	3	4	5	6	7	Total

Notation:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \right).$$

means X and Y jointly follow a bivariate normal distribution such that

$$E(X) = \mu_x,\ E(Y) = \mu_y,\ var(X) = \sigma_x^2,\ var(Y) = \sigma_y^2,\ cov(X,Y) = \rho\sigma_x\sigma_y.$$

---

1. Let $Z = X_2|(X_1 > 0)$, where

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & \sigma^2 \end{pmatrix} \right).$$

   Find the pdf of Z.

2. The joint distribution of Y and X is given by the following hierarchical model

$$Y|X \sim \text{Possion}(X), \quad X \sim \text{Gamma}(\alpha, \beta).$$

   Calculate $E(X|Y)$.

3. (a) Let X be a Gamma(α, β) random variable with density function

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta}, \quad x > 0;\ \alpha, \beta > 0.$$

   Find the density of $Y = e^{-X}$.

   (b) Let $X_1, \ldots, X_n$ be an iid random sample from uniform(0, 1). Find the density of $Y = X_1X_2 \cdots X_n = \prod_{i=1}^n X_i$.

4. Let $X_1, \ldots, X_n$ be a random sample from the following distribution:

$$f(x|\theta) = \binom{2}{x}\theta^x(1-\theta)^{2-x}I_{\{0,1,2\}}(x),$$

   where the parameter space for the unknown θ is $\Theta = [0, 1]$.

   (a) Is there a one-dimensional sufficient statistic and if so, what is it? Does a complete sufficient statistic exist?

   (b) Find a maximum likelihood estimator of $\theta^2 = P(X_1 = 2)$. Is it unbiased?

   (c) Find a uniformly minimum variance unbiased estimator of θ^2 if such exists.

5. Let $\mathbf{X}$ be a random vector from an unknown distribution. According to the Neyman-Pearson Lemma, if H_0 is the simple null hypothesis that the joint density is $g(\mathbf{x})$ versus H_1 the simple alternative hypothesis that the joint density is $h(\mathbf{x})$, then $\mathcal{R}$ is the *best critical region of size* α, if, for $k > 0$: *(i)* $\frac{g(\mathbf{x})}{h(\mathbf{x})} \leq k$ for $\mathbf{x} \in \mathcal{R}$, *(ii)* $\frac{g(\mathbf{x})}{h(\mathbf{x})} \geq k$ for $\mathbf{x} \in \mathcal{R}^c$, and *(iii)* $\alpha = P_{H_0}(\mathbf{X} \in \mathcal{R})$.

   Now let $X_1, \ldots, X_n$ be a random sample from a distribution that has a probability mass function $f(x)$ that is positive only on the nonnegative integers. We wish to test the simple hypothesis

$$H_0 : f(x) = \begin{cases} \frac{e^{-1}}{x!} \ , & x = 0, 1, 2, \ldots \\ 0 \ , & \text{elsewhere,} \end{cases}$$

   against the alternative simple hypothesis

$$H_1 : f(x) = \begin{cases} \left(\frac{1}{2}\right)^{x+1} \ , & x = 0, 1, 2, \ldots \\ 0 \ , & \text{elsewhere.} \end{cases}$$

   (a) Determine the best critical region $\mathcal{R}$ for the case $n = 1$ and $k = 1$.

   (b) Compute the level of the test for the case $n = 1$ and $k = 1$.

   (c) Compute the power of the test (when H_1 is true) for the case $n = 1$ and $k = 1$.

6. Let $X_1, \ldots, X_n$ be a random sample from the Laplace distribution with density

$$f(x; \theta) = \frac{1}{2} e^{-|x-\theta|} \ .$$

   Suppose $n = 2k + 1$ is odd. Find the maximum likelihood estimator and show that it does not satisfy the likelihood equation $\partial \log L(\theta)/\partial\theta = 0$.

7. Suppose that $X_1, \ldots, X_n$ is an iid sample from a population with density $f(x; \theta)$, where θ is a unknown parameter. S_1 and S_2 are two different unbiased estimators of θ. It is known that the joint distribution of S_1 and S_2 is bivariate normal,

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \theta \\ \theta \end{pmatrix}, \frac{1}{n} \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right).$$

   We want to use $T(\lambda) = \lambda S_1 + (1 - \lambda) S_2$ for testing $H_0 : \theta = 0$ versus $H_a : \theta \neq 0$, where λ is a constant. Consider the rejection region of the form $|T(\lambda)| > c$.

   (a) Determine c at significance level α.

   (b) Expression the power function $\beta(\theta)$ in terms of Φ, where $\Phi(\cdot)$ is the cumulative distribution function of $N(0, 1)$.

   (c) Find λ that maximizes $\beta(\theta)$ for any given θ.

Solutions

1. The cdf of Z is

$$\begin{aligned}P(Z \le z) = P(X_2 \le z | X_1 > 0) &= \frac{P(X_2 \le z, X_1 > 0)}{P(X_1 > 0)} \\ &= 2P(X_2 \le z, Y > -(\rho/\sigma^2)X_2) \\ &= 2\int_{-\infty}^{z} f_{X_2}(x) \int_{-(\rho/\sigma^2)x}^{\infty} f_Y(y)dydx\end{aligned}$$

where $Y = X_1 - (\rho/\sigma^2)X_2 \sim N(0, 1-\rho^2/\sigma^2)$ and Y is independent of X_2. Thus,

$$f_Z(z) = \frac{d}{dz}P(Z \le z) = 2f_{X_2}(z)\int_{-(\rho/\sigma^2)z}^{\infty} f_Y(y)dy = \frac{2}{\sigma}\phi(z/\sigma)\Phi\left(\frac{-(\rho/\sigma^2)z}{\sqrt{1-\rho^2/\sigma^2}}\right),$$

where $\phi(\cdot)$ is a standard normal density and $\Phi(\cdot)$ is a cdf of a standard normal rv.

2. The joint density of (Y, X) is

$$\begin{aligned}f(y,x) = f(y|x)f(x) &= \frac{1}{y!}e^{-x}x^y \, \frac{1}{\Gamma(\alpha)\beta^\alpha} \, x^{\alpha-1}e^{x/\beta} \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha y!} \, e^{y+\alpha-1}e^{-(1+1/\beta)x}, \quad x > 0, y = 0, 1, 2, \ldots.\end{aligned}$$

The marginal density of Y is

$$\begin{aligned}f(y) = \int_0^\infty f(y,x)dx &= \frac{\Gamma(y+\alpha)(1+1/\beta)^{-(y+\alpha)}}{\Gamma(\alpha)\beta^\alpha y!} \\ &= \frac{\Gamma(y+\alpha)}{\Gamma(\alpha)y!}\frac{\beta^y}{(1+\beta)^{y+\alpha}}, \quad y = 0, 1, 2, \ldots.\end{aligned}$$

Therefore, the conditional density of $X|Y$ is

$$f(x|y) = \frac{(1+1/\beta)^{y+\alpha}}{\Gamma(y+\alpha)} \, x^{y+\alpha-1}e^{-(1+1/\beta)x}, \; x > 0.$$

It is a Gamma($y+\alpha$, $(1+1/\beta)^{-1}$) distribution and

$$E(X|Y) = \frac{Y+\alpha}{1+1/\beta}.$$

3. (a). It is easy to see that $X = -\log(Y)$. The density of Y is

$$f(y) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \, (-\log(Y))^{\alpha-1}y^{1/\beta-1}, \quad 0 < y < 1.$$

When $\alpha = \beta = 1$, Y follows uniform(0, 1).

(b). Let $T = -\log(Y) = \sum_{i=1}^n \{-\log(X_i)\}$. Because X_i follows uniform(0, 1), we know that $-\log(X_i)$ follows Gamma(1, 1) and T follows Gamma(n, 1). Applying the results in (a), the density of $Y = e^{-T}$ is

$$f(y) = \frac{1}{\Gamma(n)} (-\log(y))^{n-1} = \frac{(-\log(y))^{n-1}}{(n-1)!}, \quad 0 < y < 1.$$

4. (a). Note that

$$\begin{aligned} f(x_1, \cdots, x_n|\theta) &= \prod_{i=1}^n \binom{2}{x_i} \theta^{x_i} (1-\theta)^{2-x_i} I_{\{0,1,2\}}(x_i) \\ &= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{2n - \sum_{i=1}^n x_i} \cdot \prod_{i=1}^n \binom{2}{x_i} I_{\{0,1,2\}}(x_i). \end{aligned}$$

Thus, by the factorization theorem, $\sum_{i=1}^n X_i$ is a one-dimensional sufficient statistic for θ. Since

$$f(x_1, \cdots, x_n|\theta) = \prod_{i=1}^n I_{\{0,1,2\}}(x_i) \exp\left\{\sum_{i=1}^n x_i \log\left(\frac{\theta}{1-\theta}\right) + 2n\log(1-\theta) + \sum_{i=1}^n \log\binom{2}{x_i}\right\},$$

is one parameter exponential family with $\theta \in [0, 1]$, $\eta = \log(\theta/(1-\theta)) \in \bar{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{\infty\} = \mathcal{N}_0$. Since $\mathcal{N}_0$ contains a one-dimensional open rectangle, $\sum_{i=1}^n X_i$ is the CSS for θ.

(b). Note that

$$\begin{aligned} l(\theta) &= \sum_{i=1}^n x_i \log\theta + \left(2n - \sum_{i=1}^n x_i\right) \log(1-\theta) + \sum_{i=1}^n \log\binom{2}{x_i} \\ \frac{\partial l(\theta)}{\partial\theta} &= \frac{\sum_{i=1}^n x_i}{\theta} - \frac{2n - \sum_{i=1}^n x_i}{1-\theta} = 0. \end{aligned}$$

Thus, $\hat{\theta} = \bar{X}/2$ and so $\widehat{\theta^2} = (\bar{X}/2)^2$.

Also, since $E(X_1) = 2\theta$ and $E(X_1^2) = 2\theta(1+\theta)$,

$$\begin{aligned} E\left(\widehat{\theta^2}\right) &= \frac{1}{4n^2} E\left(\sum_{i=1}^n X_i^2 + 2\sum_{i<j} X_i X_j\right) \\ &= \frac{1}{4n} E(X_1^2) + \frac{n-1}{4n} E(X_1)^2 \\ &= \theta^2 + \frac{\theta(1-\theta)}{2n}. \end{aligned}$$

So, if $\theta \neq 0$ or 2, $\widehat{\theta^2}$ is not an unbiased estimator of θ^2.

(c). If we let $T(X) = I(X_1 = 2)$ then $E_\theta\big(T(X)\big) = P_\theta\big(X_1 = 2\big) = \theta^2$. Thus, $I(X_1 = 2)$ is an unbiased estimator of θ^2. And we know that $\sum_{i=1}^n X_i \sim \mathrm{B}(2n, \theta)$. Thus,

$$\begin{aligned}
E_\theta\left(I(X_1 = 2) \middle| \sum_{i=1}^n X_i = t \right) &= P_\theta\left(X_1 = 2 \middle| \sum_{i=1}^n X_i = t \right) \\
&= \frac{P_\theta(X_1 = 2) \cdot P_\theta(\sum_{i=2}^n X_i = t-2)}{P_\theta(\sum_{i=1}^n X_i = t)} \\
&= \frac{\frac{2!}{2!0!}\theta^2(1-\theta)^0 \cdot \frac{(2n-2)!}{(t-2)!(2n-t)!}\theta^{t-2}(1-\theta)^{2n-t}}{\frac{(2n)!}{t!(2n-t)!}\theta^t(1-\theta)^{2n-t}} \\
&= \frac{t(t-1)}{2n(2n-1)}.
\end{aligned}$$

Therefore, by Rao-Blackwell-Lehmann-Scheffé theorem, $\dfrac{\sum_{i=1}^n X_i\big(\sum_{i=1}^n X_i - 1\big)}{2n(2n-1)}$ is the UMVUE of θ^2.

5. (a). Here

$$\frac{g(\mathbf{x})}{h(\mathbf{x})} = \frac{e^{-n}/(x_1!x_2!\cdots x_n!)}{(1/2)^n(1/2)^{x_1+x_2+\cdots+x_n}} = \frac{(2e^{-1})^n 2^{\sum x_i}}{\prod x_i!} \ .$$

Thus the best critical region is

$$\mathcal{R} = \left\{ \mathbf{x} = (x_1, \ldots, x_n) : \frac{(2e^{-1})^n 2^{\sum x_i}}{\prod x_i!} \le k \right\} \ .$$

For $k = 1, n = 1$, $\mathcal{R} = \{x_1 : 2^{x_1}/x_1! \le e/2\}$. It is easy to see that this inequality is satisfied by all nonnegative integers except 1 and 2. Thus, $\mathcal{R} = \{0, 3, 4, 5, \ldots\}$.

(b). The level of the test is

$$\alpha = P_{H_0}(X_1 \in \mathcal{R}) = 1 - P_{H_0}(X_1 = 1, 2) = 1 - 1/e - 1/2e = 0.448 \ .$$

(c). The power of the test is given by

$$P_{H_1}(X_1 \in \mathcal{R}) = 1 - P_{H_1}(X_1 = 1, 2) = 1 - 1/4 - 1/8 = 0.625 \ .$$

6. The log likelihood function is given by

$$\log L(\mathbf{x}; \theta) = -n \log 2 - \sum_{i=1}^n |x_i - \theta| \ .$$

Note that maximizing $\log L$ is the same as minimizing $g(\theta) = \sum_{i=1}^n |x_i - \theta|$. We can write $g(\theta)$ using the order statistics $x_{(1)} \le \cdots \le x_{(n)}$ as $g(\theta) = \sum_{i=1}^n |x_{(i)} - \theta|$. To find the minimizer of $g(\theta)$, suppose $x_{(j)} \le \theta \le x_{(j+1)}$. Then

$$g(\theta) = \sum_{i=1}^{j-1} (\theta - x_{(i)}) + x_{(j+1)} - x_{(j)} + \sum_{i=j+2}^{n} (x_{(i)} - \theta) \ .$$

Now increasing θ by a small amount ϵ will increase the left-hand sum by $(j-1)\epsilon$ and decrease the right-hand sum by $(n-j-1)\epsilon$. Thus $g(\theta)$ will decrease iff $n-j-1 > j-1$ or $n > 2j$. Since $n = 2k+1$, the sum will drop if we increase j up to $j = k$. Moreover, if $x_{(k)} \leq \theta \leq x_{(k+1)}$, then increasing θ by ϵ will decrease $g(\theta)$ until $\theta = x_{(k+1)}$ since $n > 2k$. For $j \geq k+1$, we have $n < 2j$ and thus $g(\theta)$ increases in θ if $\theta > x_{(k+1)}$. Thus, the MLE is $\hat{\theta} = x_{(k+1)}$.

Now $\hat{\theta} = x_{(k+1)}$ does not satisfy the likelihood equation because the likelihood equation is not differentiable at any data point.

7. The distribution of $T(\lambda)$ is normal with mean θ and variance

$$\begin{aligned}\tau^2 = var[T(\lambda)] &= \lambda^2 var(S_1) + (1-\lambda)^2 var(S_2) + 2\lambda(1-\lambda)cov(S_1, S_2) \\ &= \frac{1}{n}(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)\lambda^2 - \frac{1}{n}2(\sigma_2^2 - \rho\sigma_1\sigma_2)\lambda + \frac{1}{n}\sigma_2^2.\end{aligned}$$

The power function is

$$\begin{aligned}\beta(\theta) &= P_\theta(|T(\lambda)| > c) = P_\theta(T(\lambda) > c) + P_\theta(T(\lambda) < -c) \\ &= 1 - \Phi\Big(\frac{c-\theta}{\tau}\Big) + \Phi\Big(\frac{-c-\theta}{\tau}\Big).\end{aligned}$$

Because the significance level is α, we have

$$\alpha = \beta(0) = 1 - \Phi\Big(\frac{c}{\tau}\Big) + \Phi\Big(\frac{-c}{\tau}\Big),$$

which yields $c = \tau z_{\alpha/2}$. So the power function can be written as

$$\beta(\theta) = 1 - \Phi\Big(z_{\alpha/2} - \frac{\theta}{\tau}\Big) + \Phi\Big(-z_{\alpha/2} - \frac{\theta}{\tau}\Big)$$

For any given θ,

$$\frac{\partial\beta(\theta)}{\partial\tau} = -\phi\Big(z_{\alpha/2} - \frac{\theta}{\tau}\Big)\frac{\theta}{\tau^2} + \phi\Big(-z_{\alpha/2} - \frac{\theta}{\tau}\Big)\frac{\theta}{\tau^2} < 0,$$

where ϕ is the density of $N(0,1)$. Notice that the above inequality holds because when $\theta > 0$, $\phi(-z_{\alpha/2} - \theta/\tau) < \phi(z_{\alpha/2} - \theta/\tau)$ and when $\theta < 0$, $\phi(-z_{\alpha/2} - \theta/\tau) > \phi(z_{\alpha/2} - \theta/\tau)$. Therefore, in order to maximize $\beta(\theta)$, we only need to minimize τ. Because τ is a quadratic function of λ and $(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) > 0$, τ^2 is minimized at

$$\lambda = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$