STT 871-872 Fall Preliminary Examination
Wednesday, August 28, 2013
12:30 - 5:30 pm

1. Consider the set up in which our data are $(x_i, Y_i, w_i), 1 \leq i \leq n$, obeying the model

$$Y_i = \beta_1 + w_i\beta_2 + x_i\beta_3 + \varepsilon_i, \qquad i = 1, 2, \cdots, n,$$

where $w_1, w_2. \cdots, w_n$ and $x_1, x_2, \cdots, x_n$ are known constants; β_1, β_2, and β_3 are real parameters; and $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ are i.i.d. $N(0; \sigma^2)$ random errors. Assume that $\sum_1^n w_i = 0 = \sum_1^n x_i$. For notation, let $S_{ww} = \sum w_i^2, S_{xx} = \sum x_i^2, S_{wx} = \sum w_i x_i$ and so on.

a. Write the model in matrix form as $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ describing entries in the matrix $\mathbf{X}$. [5]

b. If $n > 3$, show that $\mathbf{X}$ will be full rank iff $D = S_{xx}S_{ww} - S_{wx}^2 \neq 0$. [5]

c. Assuming X is of full rank, give an explicit formula for the least squares estimator $\hat{\beta}$ of $\beta = (\beta_\mathbf{1}, \beta_\mathbf{2}, \beta_\mathbf{3})$ (It will involve terms such as S_{xx}, S_{xY} , etc.). You may use the following fact. [5]

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} \frac{1}{n} & 0 & 0 \\ 0 & S_{xx}/D & -S_{wx}/D \\ 0 & -S_{wx}/D & S_{ww/D} \end{pmatrix}.$$

2. Let $X, X_1, \ldots, X_n$ be i.i.d. r.v.'s such that for $\theta > 0$, they have common density (with respect to Lebesgue measure),

$$\begin{aligned} f_\theta(x) &= x\theta^2 e^{-\theta x}, && x > 0; \\ &= 0, && x \leq 0. \end{aligned}$$

Let $p = g(\theta) = (1+\theta)e^{-\theta} = P_\theta(X > 1)$. The two natural estimators of p are

$$\tilde{p}_n = n^{-1}\sum_{i=1}^n I(X_i > 1), \qquad \text{and} \qquad \hat{p}_n = g(\hat{\theta}_n),$$

where $\hat{\theta}_n$ is the maximum likelihood estimator of θ.

a. Find the limiting distribution of $\sqrt{n}(\tilde{p}_n - p)$. [5]

b. Find the limiting distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$, [5]

c. Derive the asymptotic relative efficiency of $\tilde{p}_n$ with respect to $\hat{p}_n$. [5]

3. Let $\theta > 0$ and $X_1, X_2, \ldots,$ be i.i.d. having uniform distribution on $(0, \theta)$. Let P_n and Q_n denote the joint distributions of $X_1, X_2, \cdots, X_n$, when $\theta = 1$, and when $\theta = 1 - 1/n^p$, respectively, where p is a fixed positive constant.

a. For which values of p are $\{P_n\}$ and $\{Q_n\}$ mutually contiguous? [5]

b. When $\{P_n\}$ and $\{Q_n\}$ mutually contiguous, identify the limit points of the distribution of dQ_n/dP_n, under P_n. [5]

4. Let X be a $N(\theta, 1)$ r.v., with θ in the set of integers $\mathbf{N} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$. Consider the problem of estimating of θ with the loss function $L(\theta, a)$ as the 0-1 loss.

a. Suppose an estimator T is equivariant, i.e., satisfies $T(x+k) = T(x) + k$, for all $x \in \mathbb{R}$ and all $k \in \mathbf{N}$. Show that the risk function of T is constant in θ. [5]

b. Let $S(X) = X - [X]$, where $[X]$ is the integer nearest to X. Show that every equivariant estimate is of the form $[X] - v(S(X))$, for some measurable function v of $S(X)$. [5]

c. Find the minimum risk equivariant estimate of θ. [5]

d. Which of the three estimates X, $[X]$, and S, are (i) sufficient for θ, (ii) complete sufficient for θ. [5]

5. Suppose that $X_1, X_2, \cdots, X_n$ are independent r.v.'s, with X_i having $N(\mu_i, 1)$ distribution. Consider the following hypotheses:

$$(1) \qquad H_0 : \mu_i = 0 \quad \text{for all } i = 1, \cdots, n, \qquad \text{vs.}$$
$$H_1 : \mu_i = ab_i \text{ with } b_i \text{ i.i.d. Bernoulli}(p), \text{ independent of all } X_j, 1 \leq j \leq n,$$

where $a \in \mathbb{R}$ and $0 < p < 1$ are known constants. Let $\mu = (\mu_1, \cdots, \mu_n)^T$. Note that under the null hypothesis, $\mu = \mathbf{0}_{n\times 1}$.

a. Find the marginal distributions of $X_1, X_2, \cdots, X_n$ under H_1. [5]

b. Show that the likelihood ratio statistic for testing (1) is [5]

$$W = \prod_{i=1}^{n} \Big\{ 1 + p\Big(\exp(-\frac{a^2}{2}) \exp(aX_i) - 1 \Big) \Big\}.$$

c. For a given $c > 0$, let $\varphi_c = I\{W > c\}$ be a test function corresponding to the hypotheses (1). Namely, the test φ_c rejects H_0 if $W > c$ and accept H_0, if $W \leq c$. Define the risk of φ_c to be

$$\text{Risk}_\pi(\varphi_c) = P_0(W > c) + E_\pi \left[P_\mu(W \leq c|\mu) \right].$$

where P_0 is the probability measure under the null hypothesis and P_μ is the probability measure under the alternative conditional on μ, and the expectation is taken with respect to the distribution π of $\mu = (ab_1, ab_2, \ldots, ab_n)$. Show that the test $\varphi_1 = I\{W > 1\}$ minimizes the risk $\text{Risk}_\pi(\varphi_c)$ w.r.t. $c > 0$. [5]

d. Show that the risk of φ_1 has a lower bound

$$\text{Risk}_\pi(\varphi_1) \geq 1 - \frac{1}{2}\sqrt{E_0(W^2) - 1},$$

where the expectation E_0 is taken with respect to P_0. [5]

6. Let $\mathcal{X} = \{1, 2, \cdots, k\}$, with $k < \infty$ and $\{P_\theta, \theta \in \mathbb{R}\}$ be a family of probabilities on $\mathcal{X}$ such that $P_\theta(x) > 0$, for all $\theta \in \mathbb{R}$ and $x \in \mathcal{X}$.

a. Suppose that T_n is a sequence of estimates such that $\sup_n E_{\theta_0} T_n^2 < \infty$. Show that there is a subsequence T_{n_i} and T such that, for all θ, $E_\theta T_{n_i} \longrightarrow E_\theta T$. [5]

b. Suppose that there is no unbiased estimate of the function $g(\theta)$. Let T_n is a sequence of estimates which are asymptotically unbiased, i.e. for all θ, $E_\theta T_n \longrightarrow g(\theta)$. Show that for all θ, $\text{Var}_\theta(T_n) \longrightarrow \infty$. [5]

7. Suppose $\theta = (\theta_1, \theta_2)$ is a bivariate parameter and the parameter space is $\Theta = \Theta_1 \times \Theta_2$. Suppose that $\{f(x|\theta) : \theta \in \Theta\}$ is family of densities such that $f(x|\theta) > 0$ for all x, θ. Suppose T_1 is sufficient for θ_1, whenever θ_2 is fixed and known and T_2 is sufficient for θ_2, whenever θ_1 is fixed and known. Show that (T_1, T_2) is sufficient for (θ_1, θ_2). [5]

8. Let $X_1, X_2, \ldots, X_n$ be i.i.d observations from $U(\theta - 1, \theta + 1)$ with θ in the set of integers $\mathbf{N} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$.

a. Find a MLE $\hat{\theta}_n$ such that under 0-1 loss $\hat{\theta}_n$ has constant risk. [5]

b. Is it consistent ? [5]

c. Show that $\hat{\theta}_n$ is minimax. [5]

d. Show that $\hat{\theta}_n$ is not admissible by constructing an estimate that has 0 risk at $\theta = 0$. [5]