Preliminary Exam: Probability and Statistics
August 2005

Work all 8 problems. Begin each problem on a new page, using one side of the sheet. Throughout "p.d.f." means "probability density function", $\mathbb{R} = (-\infty, \infty)$, and $\mathbf{1}_A$ denotes the indicator of the set A (i.e. $\mathbf{1}_A(x) = 1$ if $x \in A, 0$ otherwise). A table of the standard normal distribution is attached.

1. Let $X_1, \ldots, X_n$ be independent random variables and assume that X_i has a normal distribution with mean $c_i\mu, \mu \in \mathbb{R}$, and variance $\sigma^2 > 0$, where c_i is a known number $(i = 1, \ldots, n)$. Assume that $c_i \neq 0$ for at least one index i. The parameters μ and σ^2 are unknown.

- **a.** Find the maximum likelihood estimators of μ and σ^2.
- **b.** Find the family of likelihood ratio tests for testing the null hypothesis $H_0 : \mu = 0$ against the alternative $H_1 : \mu \neq 0$.
- **c.** For $c_1 = \cdots = c_n = 1$, show that the LR test statistic can be reduced to a statistic with a student distribution.

2. Suppose that p and q are p.d.f.'s that are strictly positive and continuous on the interval $[0, 1]$. Let $X_1, \ldots, X_n$ be a random sample of size n from the p.d.f.

$$f_\theta(x) = c(\theta)\{p(x)\}^\theta\{q(x)\}^{1-\theta}\mathbf{1}_{[0,1]}(x), x \in \mathbb{R},$$

where the parameter $\theta \in (0, 1)$ is unknown and $c(\theta) = [\int_0^1 \{p(x)\}^\theta\{q(x)\}^{1-\theta}dx]^{-1}$.

- **a.** Show that the family of p.d.f.'s $f_\theta, 0 < \theta < 1$, is an exponential family.
- **b.** Determine the family of uniformly most powerful critical regions for testing the null hypothesis $H_0 : \theta = \theta_0$ against the alternative $H_1 : 0 < \theta < \theta_0$, for some given $\theta_0 \in (0, 1)$.

Henceforth let $p(x) = 2x\mathbf{1}_{[0,1]}(x), q(x) = \mathbf{1}_{[0,1]}(x), x \in \mathbb{R}, n = 2$, and $\theta_0 = \frac{1}{2}$.

- **c.** Determine the density $f_{\frac{1}{2}}(x), x \in \mathbb{R}$, explicitly.
- **d.** Find, in the present situation, the most powerful test of size $\alpha, 0 < \alpha < 1$.

3. Let X be a random variable with a double exponential p.d.f.

$$f_\theta(x) = \frac{1}{2\theta}\, e^{-|x|/\theta}, x \in \mathbb{R},$$

for some unknown $\theta \in (0, \infty)$. Suppose $X_1, \ldots, X_n$ is a random sample of size n from this p.d.f.

continued on page 2

a. Compute $E_\theta|X|$ and $E_\theta X^2$.

b. Show that the statistic $T = \frac{1}{n}\sum_{i=1}^{n}|X_i|$ is an unbiased estimator of θ, and calculate its variance.

c. Calculate the Fisher information $I(\theta)$ at θ for the family of p.d.f.'s given above.

d. Is T in part **b** the uniform minimum variance unbiased estimator of θ? Why?

4. Suppose that X and Y are jointly continuous random variables with

$$f_{Y|X}(y|x) = \mathbf{1}_{(x,x+1)}(y), f_X(x) = \mathbf{1}_{(0,1)}(x), x, y \in \mathbb{R}.$$

a. Compute $\text{Cov}(X, Y)$.

b. Compute $P(X + Y < 1)$.

c. Find $f_{X|Y}(x|y)$ for $0 < y < 2$.

d. Find the conditional expectation $E(X|Y)$.

5. Let $p : \mathbb{R} \to \mathbb{R}$ be a given function that is strictly positive and continuous on $[0, \infty)$. Let $P(x) = \int_0^x p(y)dy, x \in \mathbb{R}$.

a. For $\theta > 0$ determine the number $c(\theta)$ in such a way that the function

$$f_\theta(x) = c(\theta)p(x)\mathbf{1}_{[0,\theta]}(x), x \in \mathbb{R},$$

is a p.d.f.

A random sample $X_1, \ldots, X_n$ of size n from the density f_θ is given, where the parameter $\theta > 0$ is unknown.

b. Compute the maximum likelihood estimator of θ.

c. Find a complete sufficient statistic for θ. Explain why it is complete.

d. Find the minimum variance unbiased estimator of $P(\theta)$.

e. Determine the conditional expectation $E(P(X_1)|S)$, if S is a complete sufficient statistic for θ.

6. Suppose that the random variables $X_1, \ldots, X_n$ are independent and identically distributed with p.d.f.

$$f(x) = 2x\mathbf{1}_{[0,1]}(x), x \in \mathbb{R},$$

under the null hypothesis H_0, and p.d.f.

$$g(x) = 4x^3\mathbf{1}_{[0,1]}(x), x \in \mathbb{R},$$

under the alternative H_1.

continued on page 3

a. Determine the family of most powerful tests for testing H_0 against H_1.

b. Determine the most powerful test of approximate size $\alpha, 0 < \alpha < 1$, assuming that the sample size n is large enough for application of the central limit theorem.

7. Let the discrete random variables X and Y have the joint p.d.f.

$$f_{X,Y}(x, y) = \frac{2}{n(n+1)}, y = 1, \dots, x, \; x = 1, \dots, n.$$

Compute:

a. The marginal p.d.f f_Y.

b. The conditional p.d.f. $f_{X|Y}(x|y)$.

c. The conditional expectation $E(X|Y)$.

d. Of course we have either $\mathrm{Var}(X) \leq \mathrm{Var}(E(X|Y))$ or $\mathrm{Var}(X) \geq \mathrm{Var}(E(X|Y))$. Which of these two inequalities should hold according to theory?

8. Let $X_1, \dots, X_n$ be independent and identically distributed random variables with common density

$$f(x) = 3\, x^2 \mathbf{1}_{[0,1]}(x), x \in \mathbb{R},$$

and let $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Find the limiting normal distribution of $\sin(\overline{X}_n)$, as $n \to \infty$. (Hint: you may use the "delta-method".)