Preliminary Exam: Probability and Statistics
2003

Work any 8 problems and clearly indicate which problems you wish to be graded. Begin each problem on a new page, using one side of the sheet. A table of chi-square distributions is attached.
Throughout "p.d.f." means "probability density function", $\mathbb{R}$ is the real line $(-\infty, \infty)$, and $\mathbb{N}$ is the set of the natural numbers $\{1, 2, \dots\}$.

1. Let X and Y denote the portions of two different types of components in a sample from a mixture of chemicals used as an insecticide. Suppose that X and Y have the joint p.d.f. given by

$$f(x,y) = \begin{cases} 2, & 0 \le x \le 1, 0 \le y \le 1, 0 \le x+y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

a. Find $P\{X \ge \frac{1}{2}, Y \le \frac{1}{2}\}$.

b. Find the marginal p.d.f. of Y.

c. Find the conditional p.d.f. of X given $Y = y$.

2. Let $X_1, \dots, X_n$ be a random sample of size n from a distribution with mean 0, variance 1, and finite fourth moment. Find the limiting distribution of

$$\frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n X_i^2}}.$$

3. Suppose that p and q are p.d.f.'s that are strictly positive and continuous on the real line. Let $X_1, \dots, X_n$ be a random sample of size n from the p.d.f.

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

for some $0 < \theta < 1$, where $c(\theta) = 1/\int_{-\infty}^{\infty} \{p(x)\}^\theta \{q(x)\}^{1-\theta} dx$.

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 $H_0 : \theta = \theta_0$ versus the alternative $H_1 : \theta_0 < \theta < 1$, for some given $0 < \theta_0 < 1$.

continued on page 2

4. Let $X_1, \ldots, X_n$ be a random sample of size n from a distribution with p.d.f.

$$f_\theta(x) = \begin{cases} 1/\theta, & \text{for } x = 1, \ldots, \theta, \\ 0, & \text{elsewhere,} \end{cases}$$

for some $\theta \in \mathbb{N}$. The maximum of the $X_i (i = 1, \ldots, n)$, i.e. the largest order statistic, is denoted by $X_{(n)}$.

- **a.** Prove that $X_{(n)}$ is a sufficient statistic for $\theta \in \mathbb{N}$.
- **b.** Find the p.d.f. of $X_{(n)}$.
- **c.** Prove that $X_{(n)}$ is a complete statistic for $\theta \in \mathbb{N}$.
- **d.** Write $h(\theta) = E_\theta(X_{(n)})$. Is $X_{(n)}$ the uniform minimum variance unbiased estimator of $h(\theta), \theta \in \mathbb{N}$? State any theorem you apply.

5. Let $X_1, \ldots, X_n$ be a random sample of size n from the p.d.f.

$$f_\theta(x) = \begin{cases} (1/\theta)^2 x e^{-x/\theta}, & x > 0, \\ 0, & \text{elsewhere,} \end{cases}$$

for some $\theta > 0$. Denote the sample mean by $\overline{X}$.

- **a.** Compute $E_\theta(\frac{1}{2}\overline{X})$ and $\mathrm{Var}_\theta(\frac{1}{2}\overline{X})$. Show your calculations.
- **b.** Find the Cramér-Rao lower bound for unbiased estimators of $\theta > 0$, based on this sample.
- **c.** Is the estimator $\frac{1}{2}\overline{X}$ unbiased for $\theta > 0$?
  Is it efficient for $\theta > 0$? (Justify your answers.)

6. Let X_1, X_2 be a random sample of size 2 from the p.d.f.

$$f(x) = \begin{cases} 1, & 0 \le x \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Define $Y = \min\{X_1, X_2\}$ and $Z = \max\{X_1, X_2\}$.

- **a.** Find the joint p.d.f. of the pair (Y, Z).
- **b.** Find the conditional p.d.f. of Y given $Z = z$.
- **c.** Find the p.d.f. of the quotient Y/Z of Y and Z.

continued on page 3

7. On the basis of one single "observation" X one wants to test the null hypothesis H_0 that X has the standard normal density $f_0(x) = \left(1/\sqrt{2\pi}\right) e^{-x^2/2}, x \in \mathbb{R}$, versus the alternative H_1 that X has the standard double exponential density $f_1(x) = (1/2)e^{-|x|}, x \in \mathbb{R}$.

a. Determine the family of most powerful critical regions.

b. Compute the type II error probability of the most powerful critical region of size $\alpha = 2\{1 - \Phi(2)\}$.

8. Use the data shown in the following table to test at the 0.01 level of significance (approximate) whether a person's ability in mathematics is independent of his or her interest in statistics.

		ability in mathematics		
		low	average	high
interest in statistics	low	63	42	15
	average	58	61	31
	high	14	47	29

9. Let $B(\theta, \theta) = \{\Gamma(\theta)\}^2/\Gamma(2\theta), \theta > 0$, where Γ denotes the gamma function. For natural numbers $k = 1, 2, \ldots$ we have $\Gamma(k) = (k-1)!$. Suppose that $X_1, \ldots, X_n$ is a random sample of size n from the p.d.f.

$$f_\theta(x) = \begin{cases} \frac{1}{B(\theta,\theta)} x^{\theta-1}(1-x)^{\theta-1}, & 0 < x < 1, \\ 0, & \text{elsewhere}, \end{cases}$$

for some $\theta > 0$.

a. Prove that the statistic $S = \Pi_{i=1}^n X_i(1 - X_i)$ is a sufficient statistic for $\theta > 0$.

b. Compute $E_\theta S$.

c. Restricting θ to $\mathbb{N}$, find an unbiased estimator for $\{\theta/(2\theta+1)\}^n$ which is a function of S.

10. Consider the family of p.d.f.'s

$$f_\theta(x) = \begin{cases} (1/\theta)e^{-x/\theta}, & x \geq 0, \\ 0, & \text{elsewhere}, \end{cases}$$

$\theta > 0$. Suppose that $X_1, \ldots, X_m$ is a random sample of size m from the p.d.f. f_ξ, for some $\xi > 0$, and that $Y_1, \ldots, Y_n$ is a random sample of size n from the p.d.f. f_η, for some $\eta > 0$. The two samples are stochastically independent. The null hypothesis $H_0 : \xi = \eta$ is to be tested versus the alternative $H_1 : \xi \neq \eta$.

Find the likelihood ratio statistic λ for this testing problem. Simplify the expression for λ as much as you can.