Preliminary Exam: Probability and Statistics
2004

Work all 9 problems. Begin each problem on a new page, using one side of the sheet. Tables of the normal and chi-square distributions are attached. Throughout “p.d.f.” means “probability density function,” \( \mathbb{R} \) is the real line, and \( 1_A \) the indicator function of the set \( A \) (i.e. \( 1_A(x) = 1 \) if \( x \in A, 0 \) otherwise).

1. Let \( X \) and \( Y \) be nonnegative random variables of continuous type with p.d.f.’s \( f \) and \( g \) respectively. These densities are supposed to be continuous on \( \mathbb{R} \) with respective cumulative distribution functions \( F \) and \( G \) satisfying \( \lim_{x \to \infty} x \{1 - F(x)\} = 0 \) and \( \lim_{x \to \infty} x \{1 - G(x)\} = 0 \).
   a. Prove that \( E(X) = \int_0^\infty \{1 - F(x)\} \, dx \).
   b. If \( P\{X > x\} \geq P\{Y > x\} \) for all \( x \in \mathbb{R} \), prove that \( E(X) \geq E(Y) \).

2. Let \( X_1, \ldots, X_n \) be independent and identically distributed random variables with common density
   \[
   f_\theta(x) = \theta x^{-\theta-1} 1_{(1,\infty)}(x), \quad x \in \mathbb{R}, \quad \theta > 0.
   \]
   Define \( S_n = (\prod_{i=1}^n X_i)^{1/n} \).
   a. Compute \( E_\theta(\log S_n) \) and \( \text{Var}_\theta(\log S_n) \).
   b. Find the limiting distribution of \( \log S_n \).
   c. Choose \( \theta = 10, n = 100 \), and approximate \( P\{S_n > e^{0.11}\} \).

3. Let the random variables \( X \) and \( Y \) have the joint p.d.f.
   \[
   f(x, y) = \begin{cases} 
   8xy, & \text{for } 0 < x < y < 1, \\
   0, & \text{elsewhere}.
   \end{cases}
   \]
   a. Determine the marginal p.d.f.’s.
   b. Compute \( E(Y), \text{Var}(Y) \).
   c. Find the conditional p.d.f. of \( Y \) given \( X = x \).
   d. Find the conditional expectation of \( Y \) given \( X = x \).

continued on page 2
4. Let $X$ be a random sample of size 1 from the p.d.f.

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

$\theta > 0$.

a. Compute $E_\theta(\sin X)$.

b. Is $X$ complete for $\theta > 0$? (Explain your answer.)

c. Let $X_1, \ldots, X_n$ be a random sample of size $n$ from the above p.d.f. Find a real valued random variable that is a sufficient statistic for $\theta > 0$.

5. Let $X_1, \ldots, X_n$ be a random sample of size $n$ from the Poisson $(\theta)$ - distribution.

a. Explain why $T = \sum_{i=1}^{n} X_i$ is a sufficient and complete statistic for $\theta > 0$.

b. Find the exact probability distribution of $T$ using the moment generating function technique.

c. Find the uniform minimum variance unbiased estimator of $h(\theta) = \theta^2, \theta > 0$.

(Hint: you may use that $\sum_{k=2}^{\infty} k(k-1)e^{-n\theta}(n\theta)^k/k! = (n\theta)^2$.)

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

$$f_\theta(x) = \theta(1 - \theta)^{x-1}, x = 1, 2, \ldots, 0 < \theta < 1.$$

a. Explain why $\sum_{i=1}^{n} X_i$ is a sufficient and complete statistic for $0 < \theta < 1$.

Henceforth let $n = 1$ and simply write $X_1 = X$. Define

$$T(X) = \begin{cases} 1, & \text{if } X = 1, \\ 0, & \text{if } X \geq 2. \end{cases}$$

b. Show that $T$ is the uniform minimum variance unbiased estimator of $\theta$.

c. Compute $E_\theta(X)$. (Hint: you may use the relation $\sum_{x=1}^{\infty} x(1 - \theta)^{x-1} = \frac{d}{d\theta} \left( \sum_{x=1}^{\infty} (1 - \theta)^x \right), 0 < \theta < 1$.)

d. Compute the Cramér-Rao lower bound for unbiased estimators of $\theta$.

e. Is $T$ an efficient estimator of $\theta$? (Explain your answer.)

continued on page 3
7. Let $X$ be a random sample of size 1 from the p.d.f.

$$f_\theta(x) = \left\{ 1 + \theta^2 \left( \frac{1}{2} - x \right) \right\} \mathbf{1}_{(0,1)}(x), x \in \mathbb{R},$$

$\theta \in [-1, 1]$. On the basis of this sample, consider the following.

a. Derive the most powerful test for testing $H_0 : \theta = 0$ against $H_1 : \theta = \theta_1, \theta_1 \in [-1, 1], \theta_1 \neq 0$, at the level of significance $\alpha = 0.05$.

b. Investigate whether or not the test derived in part a is uniformly most powerful at that level for testing $H_0 : \theta = 0$ against the alternative $\bar{H}_1 : \theta \neq 0$.

c. Determine the power function of the test in part a.

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

$$f_\theta(x) = \theta \cdot e^{-\theta x} \mathbf{1}_{(0,\infty)}(x), x \in \mathbb{R},$$

$\theta > 0$. Derive the family of likelihood ratio tests for testing $H_0 : \theta = \theta_0$, for fixed $\theta_0 > 0$, against the alternative $H_1 : \theta \neq \theta_0$, and prove that these tests are equivalent with rejecting $H_0$ when either $\bar{X} \leq c_1$ or $\bar{X} \geq c_2$ for suitable numbers $c_1 \leq c_2$.

9. Each person in a sample of size $n = 100$ from a population of males (M) and females (F) is classified according to whether he or she is a smoker (S) or not (NS). The data are summarized in the table below:

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>F</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>65</td>
</tr>
</tbody>
</table>

Test for independence of gender and the habit of smoking or not smoking at the approximate level of significance $\alpha = 0.05$. 

Page 3