Statistics & Probability preliminary examination

May 1, 2013

Instructions: To pass this exam you need to pass both parts: probability and statistics.

Part I Statistics

1. (Stat) Let \( X_1, X_2, \ldots, X_n \) be a random sample from the Cauchy location family, i.e. the common density function is:

\[
f_\theta(x) = \frac{1}{\pi \{1 + (x - \theta)^2\}}
\]

where \( \theta \in \mathbb{R} \). Let \( M_n = \text{median}(X_1, X_2, \ldots, X_n) \) be the sample median. Compute the relative efficiency of \( \{M_n: n = 1, 2, \ldots\} \) as a sequence of estimators of the location parameter \( \theta \).

Hint: You may use the following well-known result about Order Statistics:

**Theorem 1** Suppose that \( X_1, X_2, \ldots, X_n \) is a random sample from a population with continuous distribution \( F \), density function \( f \). Let \( G \) be the inverse function of \( F \). Assume that \( \lambda \) is a positive number such that \( f(G(\lambda)) > 0 \), and \( \alpha_n \) is a sequence of positive integers such that \( \lim_{n \to \infty} \sqrt{n} \left( \frac{\alpha_n}{n} - \lambda \right) = 0 \).

Then

\[
\sqrt{n} \left( X_{[n;\alpha_n]} - G(\lambda) \right) \overset{D}{\to} N(0, \frac{\lambda(1 - \lambda)}{\{f(G(\lambda))\}^2}) \quad \text{as } n \to \infty.
\]

where \( X_{[n;\alpha_n]} \) is the \( \alpha_n \)th smallest among \( X_1, X_2, \ldots, X_n \).

2. (Stat). Let \( X_1, X_2, \ldots, X_n \) be a random sample from a population with density function

\[
f(x; \theta) = \theta c^\theta x^{-(\theta+1)}, \quad x > c,
\]

where \( c > 0 \) is known. Use a prior

\[
\lambda(\theta) = \frac{b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)}, \quad \theta > 0,
\]

where \( a > 0, b > 0 \) are both user-specified and known. Find the posterior distribution and the Bayes estimator (under the square error loss), i.e. the posterior mean, of \( \theta \).

3. (Stat) Let \( X_1, X_2, \ldots, X_n \) be a random sample from the Poisson distribution with mean \( \theta > 0 \).

(a) Find the (uniformly) minimum variance unbiased estimator of \( a\theta + be^{-\theta} + c \), where \( a, b, \) and \( c \) are given positive numbers.

(b) Does there exist an unbiased estimator of \( \frac{1}{\theta} \)? JUSTIFY your answer!

4. (Stat) Assume that \( X_1, X_2, \ldots, X_n \) is a random sample from the Rayleigh distribution with density function

\[
f(x; \theta) = \frac{2x}{\theta^2} \exp \left( -\frac{x^2}{\theta} \right), \quad x > 0,
\]

where \( \theta > 0 \) is an unknown parameter. Obtain the Likelihood Ratio Test (LRT) statistic for testing \( H_0: \theta = \theta_0 \) versus \( H_0: \theta \neq \theta_0 \) where \( \theta_0 > 0 \) is a given constant. Write down the \( \alpha \)-level critical region for the LRT as an inequality in terms of a sufficient statistic, and show how you would find the critical region that is valid for any sample size.
Part II Probability

5. (Prob) Suppose $X_{1,n}, X_{2,n}, \ldots$ are independent random variables centered at expectations (mean 0) and set $s_n^2 = \sum_{k=1}^{n} E(X_{n,k})^2$. Assume for all $k$ that $|X_{n,k}| \leq M_n$ with probability 1 and that $M_n/s_n \to 0$. Let $Y_{n,i} = 3X_{n,i} + X_{n,i+1}$. Show that $\frac{Y_{n,1} + Y_{n,2} + \ldots + Y_{n,n}}{s_n}$ converges in distribution and find the limiting distribution.

6. (Prob) Let $X_1, X_2, \ldots$ be independent and identically distributed random variables with distribution function

$$F(x) = \begin{cases} 0 & \text{if } x \leq 1 \\ 1 - x^{-a} & \text{if } x \geq 1 \end{cases}$$

where $a > 0$. Show that $Z_n = n^{-1/a} \max(X_1, X_2, ..., X_n)$ converges in distribution and find its limit.

7. (Prob) Let $X_1, X_2, \ldots$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$. Suppose $X_n$ has mean $\mu_n$ and variance $\sigma_n^2$. Assume that $\lim_{n \to \infty} \mu_n = \mu \in \mathbb{R}$ and $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$. Show that $X_n \to \mu$ almost surely.

8. (Prob) Prove Kolmogorov’s inequality: Let $X_1, \ldots, X_n$ be independent with mean zero and finite variances, and let $S_k = \sum_{j=1}^{k} X_j$. Then, for $t > 0$, prove that

$$\Pr(\max_{1 \leq k \leq n} |S_k| \geq t) \leq \frac{\text{Var}(S_n)}{t^2}$$