FINAL EXAMINATION
PROBABILITY AND STATISTICS
6 DEC 2012 – 2:00-5:00

Instructions: *The duration of the test is 3 hours. The maximum you can score is* 50. *The marks for each question is indicated in bold,* [**4**] *means that part carries* 4 *marks. Give all details but try to write succinctly.*

1. For each statement, state whether it is true or false and justify your answer (Justification means that if the statement is true you must give a short proof/argument and if the statement is false you must provide a counterexample).

(A) [**2**] Let X be a positive random variable. If $\mathbf{E}[X^3]$ is finite, then $\mathbf{E}[X^2]$ must be finite.

(B) [**2**] Let $X_1,\ldots,X_n$ be i.i.d. from $N(\mu,1)$ distribution where μ is unknown. Then, the estimate $\overline{X}_n$ has uniformly smaller mean-squared error than any other estimate.

(C) [**2**] Let X be a random variable with CDF $F(t)$ and pdf $f(t)$ (assume that the pdf is continuous). Then $F(1) = F(2)$ if and only if $f(s) = 0$ for all $s \in [1,2]$.

(D) [**2**] Let $X_1,\ldots,X_n$ be i.i.d. Geo(p) random variables. Then $\frac{1}{\overline{X}_n}$ is an unbiased estimate for p.

(E) [**2**] A hypothesis test rejects the null hypothesis at level α if and only if α is less than the p-value.

2. For the first three parts, choose the correct option. For the remaining just write the answer (one or two words or numbers). Justification is not needed.

(A) [**2**] Consider the problem of constructing a 95% confidence interval for the mean of a normal population whose variance is known. To reduce the length of the confidence interval by a factor of 4, the sample size

(a) must increase by a factor of 2.

(b) must increase by a factor of 4.

(c) must increase by a factor of 16.

(d) must increase, but the factor depends on the known value of the variance.

(B) [**2**] Let X,Y be two positive random variables with CDFs F and G, respectively. Assume that X and Y have finite expectations. Suppose $F(t) \le G(t)$ for all t. Then,

(a) $\mathbf{E}[X] \le \mathbf{E}[Y]$.

(b) $\mathbf{E}[X] \ge \mathbf{E}[Y]$.

(c) Neither of the above inequalities necessarily hold.

(d) There do not exist such CDFs F and G.

(C) [**2**] Let X,Y be two positive random variables with joint density $f(x,y)$. Let $U = XY$ and $V = \frac{X}{Y}$. Then, the joint density of (U,V) is

(a) $g(u,v) = \frac{1}{2} f\left(uv, \frac{u}{v}\right)$.

Date: December 6, 2012.

(b) $g(u,v) = 2\frac{u}{v} f\left(uv, \frac{u}{v}\right)$.

(c) $g(u,v) = \frac{1}{2v} f\left(\sqrt{uv}, \frac{\sqrt{u}}{\sqrt{v}}\right)$.

(d) $g(u,v) = f\left(\sqrt{uv}, \frac{\sqrt{u}}{\sqrt{v}}\right)$.

(D) **[2]** Suppose $Y \sim \text{Pois}(\lambda)$ and the conditional distribution $X \mid_{Y=n}$ if $\text{Bin}(n, \frac{1}{2})$ for every positive integer n, then the distribution of X is ______.

(E) **[2]** Let $X_1, \ldots, X_{48}$ be i.i.d. random variables with distribution uniform$[1,3]$. Then a reasonable approximation for $\mathbf{P}\{1.9 \le \overline{X}_{48} \le 2\}$ is ______.

3. Let $X_1, \ldots, X_n$ be i.i.d. samples from $N(\mu, \sigma^2)$ distribution. Both μ and σ^2 are unknown.

(1) **[3]** Find MLE for (μ, σ^2).

(2) **[2]** Find the mean squared error for the MLE of μ.

(3) **[5]** The summary data is given below. Construct a 95% confidence interval for σ^2.

$$n = 15, \quad \sum_{i=1}^{n} X_i = 30, \quad \sum_{i=1}^{n} X_i^2 = 240.$$

[**Note:** In each case, you must give the general formulas as well as the numerical answer].

4. A "psychic" predicts the order of a deck of n cards. Let X_n be the number of correct guesses.

(1) **[3]** Assuming random guessing, compute $\mathbf{E}[X_n]$ and $\text{Var}(X_n)$.

(2) **[3]** Consider the hypothesis testing problem with

$$H_0 : \textit{the predictions are made at random}.$$
$$H_1 : \textit{the predictions are better than random guessing}.$$

Describe (briefly) how you would carry out a test of significance level α for this problem.

(3) **[4]** If the observed value is $X_n = 5$, find the p-value (you may assume that n is large enough so that Poisson approximation holds, but state clearly what is meant by that).

5. Let X, Y be independent random variables with $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$.

(1) **[3]** Find $\mathbf{P}\{X \le Y\}$.

(2) **[4]** Find the density of $X + Y$.

(3) **[3]** Let a, b, c, d be real numbers and let $U = aX + bY$ and $V = cX + dY$. Find $\text{Cov}(U, V)$.

[**Note:** The answers are in terms of λ, μ and a, b, c, d, of course.]

6. The following two questions are unrelated.

(1) **[5]** A box contains four coupons labelled $1, 2, 3, 4$. Coupons are drawn at random with replacement n times. What is the chance that all coupons have been seen? Present the exact answer in terms of n.

(2) **[5]** Show that $f(t) = \frac{e^t}{(1+e^t)^2}$ is a density (it is known as the *logistic density*). Explain how you would simulate from the Logistic distribution using uniform random numbers.