Probability and Statistics Preliminary Examination: August 2012

Instructions:

- Work all 6 Problems. Neither calculators nor electronic devices of any kind are allowed. Clearly state any theorem or fact that you use. Each of the 6 Problems is equally weighted (but the parts within a Problem may not be).
- Abbreviations/Acronyms.
  - pmf (probability mass function); pdf (probability density function); cdf (cumulative distribution function); mgf (moment generating function);
  - MOME (method of moments estimator); MLE (maximum likelihood estimator); UMVUE (uniform minimum variance unbiased estimator); UMP (uniformly most powerful test); LRT (likelihood ratio test); MLR (monotone likelihood ratio).
- Notation.
  - $I(x \in A)$ or $I_A(x)$: indicator function for set A; takes on the value 1 if $x \in A$ and 0 otherwise.
  - $\mathbb{E}(X)$: expectation of random variable X.
  - $\mathbb{V}(X)$: variance of random variable X.
  - $X \sim \mathrm{N}(a, b)$: X has a normal distribution with mean a and variance b.
- Common distributions and other results.

Poisson(λ)**:** $\mathbb{E}(X) = \lambda$, $\mathbb{V}(X) = \lambda$, and pmf

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} I(x \in \{0, 1, \cdots\})$$

Exponential(λ)**:** $\mathbb{E}(X) = \lambda$, $\mathbb{V}(X) = \lambda^2$, and pdf

$$f(x) = \frac{1}{\lambda} e^{-x/\lambda} I(x > 0)$$

Beta(α, β)**:** $\mathbb{E}(X) = \alpha/(\alpha+\beta)$, $\mathbb{V}(X) = \alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$, and pdf

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} I(0 < x < 1)$$

Gamma(α, β)**:** $\mathbb{E}(X) = \alpha\beta$, $\mathbb{V}(X) = \alpha\beta^2$, and pdf

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} I(x > 0)$$

Order Statistics: Let $X_{(1)} \le \cdots \le X_{(n)}$ denote the order statistics from a random sample $X_1, \ldots, X_n$. If X_1 is continuous with pdf $f(x)$ and cdf $F(x)$, the pdf of $X_{(j)}$, $(X_{(i)}, X_{(j)})$, and $(X_{(1)}, \ldots, X_{(n)})$, are given by:

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!}[F(x)]^{j-1}[1-F(x)]^{n-j} f(x) I(-\infty < x < \infty)$$

$$f_{X_{(i)},X_{(j)}}(x_i, x_j) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f(x_i) f(x_j) [F(x_i)]^{i-1}[F(x_j)-F(x_i)]^{j-1-i}[1-F(x_j)]^{n-j} \times I(-\infty < x_i \le x_j < \infty)$$

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) = n! f(x_1) \cdots f(x_n) I(-\infty < x_1 \le \cdots \le x_n < \infty)$$

1. Toss n fair dice together. This constitutes the first toss. Now pick up only those not showing a 6, and toss them (together). This constitutes the second toss. Continue doing this until all dice show a 6. Let X denote the *total number of tosses* made. For example, if all dice show a 6 after the first toss, then $X = 1$. (Assume independence between tosses, and among dice within a toss.)

   (a) How many tosses do you *expect* to make if $n = 1$?

   (b) Compute the cdf of X for general $n \geq 1$.

   (c) Show that for a discrete random variable Y whose range is the non-negative integers,

   $$\mathbb{E}(Y) = \sum_{y=0}^{\infty} P(Y > y).$$

   (d) Using the above results, or otherwise, evaluate $\mathbb{E}(X)$ for the case $n = 2$.

2. Let $X_1, \ldots, X_5$ be a random sample of size 5 from a Uniform distribution on the interval $(0, 1)$. Denote the corresponding order statistics by $Y_1 \leq \cdots \leq Y_5$. Define $U = Y_4 - Y_2$ and $V = Y_2$.

   (a) Determine the joint distribution of (Y_2, Y_4).

   (b) Find the joint distribution of (U, V).

   (c) Show that the marginal distributions of U and V are each Beta$(2, 4)$.

   (d) Find $\rho_{U,V}$, the correlation coefficient between U and V.

3. Let the random vector $(X_1, \ldots, X_n)$ follow a multivariate normal distribution with $\mathbb{E}(X_i) = \mu_X$, $\mathbb{V}(X_i) = \sigma^2$, and $\text{Cov}(X_i, X_j) = \rho\sigma^2$, for $i = 1, \ldots, n$, $j = 1, \ldots, n$, $i \neq j$, where $0 < \rho < 1$. Additionally, let $Y_1, \ldots, Y_m$ be a random sample from a $\text{N}(\mu_Y, \sigma^2)$ distribution, and assume that X_i and Y_k are independent, for any $i = 1, \ldots, n$ and any $k = 1, \ldots, m$. Let $\overline{X} = \sum_{i=1}^{n} X_i/n$ and $\overline{Y} = \sum_{k=1}^{m} Y_k/m$ denote the sample means of the X's and Y's, respectively, and let $S_Y^2 = \sum_{k=1}^{m}(Y_k - \overline{Y})^2/(m-1)$ be the sample variance of the Y's.

   (a) Show that the variance of $\overline{X}$ is given by

   $$\mathbb{V}(\overline{X}) = \frac{\sigma^2}{n}\left[1 + (n-1)\rho\right].$$

   (b) Is $\overline{X}$ a *consistent* estimator of μ_X? Justify your answer.

   (c) Find the exact distribution of

   $$\frac{(\overline{X} - \mu_X)}{S_Y}\sqrt{\frac{n}{1 + (n-1)\rho}}.$$

   (d) Does $S_Y^2 \xrightarrow{P} \sigma^2$? Justify your answer.

4. One way to skew a symmetric distribution, is to introduce a skewness parameter $\beta \in (-\infty, \infty)$ into the definition of the pdf. If $\phi(z)$ and $\Phi(z)$ denote respectively the pdf and cdf of a standard normal random variable Z, then X is said to have a *skew-normal* distribution if its pdf is given by

$$f(x) = 2\phi(x)\Phi(\beta x)I(-\infty < x < \infty).$$

Assume that a random sample $X_1, \ldots, X_n$ is available from the distribution of X.

(a) Show that

$$\mathbb{E}(X) = \sqrt{\frac{2}{\pi}} \frac{\beta}{1+\beta^2}.$$

(b) Show that $X^2 \sim \chi_1^2$, a chi-square distribution with one degree of freedom. What does this fact imply about the relationship between the even moments of X and Z?

(c) Find a MOME for β.

(d) In the case $n = 2$ and $x_1 + x_2 = 0$ with $x_1 \neq 0$, find the MLE of β.

5. Let the pdf of the bivariate random vector (X, Y) be the product of the pdfs of two independent exponentials, $X \sim \text{Exp}(1/\theta)$ and $Y \sim \text{Exp}(\theta)$, such that the joint pdf is given by

$$f(x, y|\theta) = \exp\left\{-\theta x - \frac{y}{\theta}\right\} I(x > 0)I(y > 0)I(\theta > 0).$$

For a random sample of size n from this joint pdf, $(X_1, Y_1), \ldots, (X_n, Y_n)$, define the following statistics: $W = \sum_{i=1}^n X_i$, $V = \sum_{i=1}^n Y_i$, $T = \sqrt{V/W}$, and $U = \sqrt{VW}$.

(a) Show that (W, V) is a *minimal sufficient* statistic for θ, but is not a *complete* statistic.

(b) Show that U is *ancillary* for θ. [Hint: after obtaining the joint distribution of (T, U), use the fact that $\int_0^\infty \frac{1}{x} \exp\{-\theta y/x - yx/\theta\}dx = 2K_0(2y)$, where $K_0(\cdot)$, a modified Bessel function of the 2nd kind, is independent of θ.]

(c) Show that T is the MLE of θ.

(d) Compute the asymptotic (large sample) distribution of the MLE of θ. Use this result as the basis for constructing an asymptotic $(1 - \alpha)$ confidence interval for θ.

6. Let $X_1, \ldots, X_n$ be a random sample from the Weibull distribution with pdf

$$f(x|\theta) = \theta\gamma x^{\gamma-1} \exp\{-\theta x^\gamma\}I(x > 0),$$

where $\gamma > 0$ is known and $\theta > 0$ is unknown. Note that $\mathbb{E}(X_1^\gamma) = 1/\theta$, and let $T = \sum_{i=1}^n X_i^\gamma$. We are interested in testing $H_0 : \frac{1}{\theta} \leq \frac{1}{\theta_0}$ vs. $H_1 : \frac{1}{\theta} > \frac{1}{\theta_0}$.

(a) Show that a UMP test exists, and can be based on T. Find the form of the rejection region.

(b) Show that the critical value of the rejection region for the UMP test can be based on an appropriate quantile from a χ^2 distribution, and find the power function of the corresponding size α test.

(c) Obtain an *exact* (non-asymptotic) $(1 - \alpha)$ one-sided confidence interval for $1/\theta$ of the form $(\ell(\mathbf{x}), \infty)$, for some $\ell(\mathbf{x})$ which is a function of the data sample $\mathbf{x} = (x_1, \ldots, x_n)$.

(d) Suppose $1/\theta_0 = 12$. By using a normal approximation to the distribution of T, find a formula for the minimum sample size n that ensures the UMP size $\alpha = 0.01$ test of (a) has power at least 0.95 at the alternative value $1/\theta_1 = 15$. [Note the following standard normal quantiles: $z_{0.01} = 2.33$ and $z_{0.05} = 1.645$.]