Probability and Statistics Preliminary Examination: May 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 $X_{(1)} \leq \cdots \leq X_{(n)}$ 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:

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

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

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

1. Recall that an ordinary deck of 52 cards has 13 denominations (ace, 2, 3,..., 10, jack, queen, king) and 4 cards of each denomination (one for each suit: clubs, diamonds, hearts, spades). If a hand of 26 cards is randomly selected without replacement from an ordinary deck of 52 cards, what is the probability of:

   (a) There being either no aces or no kings in the hand?

   (b) There being no cards from k specified denominations in the hand? What are the possible values of k? [Note that (b) is not a generalization of (a).]

   (c) At least one card from each denomination?

2. Consider a sample of size 2 drawn without replacement from an urn containing three balls, numbered 1, 2 and 3. Let X be the number on the first ball drawn, and Y be the larger of the two numbers drawn.

   (a) Find the joint pmf of X and Y, and hence compute the marginal pmf's of X and Y. Are X and Y independent?

   (b) Find $P(X = 1|Y = 3)$.

   (c) Find $\mathbb{E}(X)$ and $\mathbb{E}(Y)$. Show that $\mathbb{V}(X) = 2/3$ and $\mathbb{V}(Y) = 2/9$.

   (d) Compute the correlation between X and Y.

3. Let $X_1, \ldots, X_n$ be a random sample from an Exponential(1) distribution. Let $\overline{X}_n = \sum_{i=1}^n X_i/n$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$ denote respectively the usual sample mean and the maximum order statistic.

   (a) Show that $2(\overline{X}_n - 1) \xrightarrow{p} 0$.

   (b) Find the limiting (asymptotic) distribution of: $2(\overline{X}_n - 1) - 2\sqrt{n}\log(\overline{X}_n)$.

   (c) Find the limiting (asymptotic) distribution of: $n[2(\overline{X}_n - 1) - 2\log(\overline{X}_n)]$.

   (d) Find centering and scaling sequences of constants $\{a_n\}$ and $\{b_n > 0\}$ such that

$$\frac{X_{(n)} - a_n}{b_n} \xrightarrow{d} W,$$

   where W is a non-degenerate limiting random variable, and identify the distribution of W.

4. Let $X_1, \ldots, X_n$ be a random sample from a $\mathrm{N}(\theta, \sigma^2)$, where σ^2 is known. Let $\overline{X} = \sum_{i=1}^n X_i/n$ denote the usual sample mean. Recall that the mgf of a $\mathrm{N}(\mu, \sigma^2)$ is given by $M(t) = \exp\{\mu t + \sigma^2 t^2/2\}$. We are interested in estimating $\tau(\theta) = e^\theta$.

   (a) Find the MLE of $\tau(\theta)$.

   (b) Find the UMVUE of $\tau(\theta)$.

   (c) Find the *mean squared error* (MSE) of $e^{\overline{X}}$ as an estimator of $\tau(\theta)$.

   (d) By using results about the asymptotic distribution of the MLE, show that for large n the MSE of $e^{\overline{X}}$ is approximately

$$\mathrm{MSE}(e^{\overline{X}}) \approx e^{2\theta}\frac{\sigma^2}{n}.$$

5. Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution on the interval $(0, \theta)$, with $\theta > 0$ unknown. Let $Y = X_{(n)}$ be the maximum order statistic.

   (a) Show that the pdf of Y is given by

$$f(y|\theta) = \frac{n}{\theta^n} y^{n-1} I(0 < y < \theta).$$

   (b) Determine if $(Y, 1.4Y)$ an interval estimator of θ, and, if so, find its confidence level. Justify your answer.

   (c) To test $H_0 : \theta = 1$ vs. $H_1 : \theta \neq 1$, find the power function of the test that rejects H_0 if and only if $Y > 1$ or $1.4Y < 1$.

   (d) For a prior distribution on θ that has the form $\pi(\theta) = \beta\alpha^\beta\theta^{-(\beta+1)} I(\theta > \alpha)$, where $\alpha > 0$ and $\beta > 1$ are fixed and known, find the *posterior Bayes estimator* of θ (under the usual squared error loss).

   (e) With the prior as in part (d), find a 90% *Bayesian credible interval* for θ that has the shortest possible length.

6. Let $X_1, \ldots, X_n$ be a random sample from the location-scale Exponential with pdf

$$f(x|\theta, \beta) = \frac{1}{\beta}\exp\left\{-\frac{x - \theta}{\beta}\right\} I(x > \theta),$$

where $-\infty < \theta < \infty$ and $\beta > 0$ are <u>both</u> unknown.

   (a) Find the LRT of $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$.

   (b) For $\theta_0 = 0$ and $n = 2$, show that the LRT in (a) can be based on the statistic

$$T = \frac{2X_{(1)}}{X_{(1)} + X_{(2)}}.$$

   Show that in this $n = 2$ case T has a pivotal distribution when $\theta = 0$, and use it to construct the corresponding size α LRT.

   (c) Find the LRT of $H_0 : \beta \leq \beta_0$ vs. $H_1 : \beta > \beta_0$.