Probability and Statistics Preliminary Examination: August 2015

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); iid (independent and identically distributed).
  - MSE (mean squared error), MOME (method of moments estimator); MLE (maximum likelihood estimator); PBE (posterior Bayes estimator); UMVUE (uniform minimum variance unbiased estimator); UMP (uniformly most powerful); LRT (likelihood ratio test).
- 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 and mgf given respectively by

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

Geometric(p)**:** $\mathbb{E}(X) = 1/p$, $\mathbb{V}(X) = (1-p)/p^2$, and pmf and mgf given respectively by

$$f(x) = p(1-p)^{x-1} I(x \in \{1, 2, \cdots\}), \qquad M(t) = \frac{pe^t}{1-(1-p)e^t}, \quad t < -\log(1-p)$$

Negative-Binomial(r, p)**:** $\mathbb{E}(X) = r(1-p)/p$, $\mathbb{V}(X) = r(1-p)/p^2$, and pmf and mgf, respectively:

$$f(x) = \binom{r+x-1}{x} p^r (1-p)^x I(x \in \{0, 1, \cdots\}), \qquad M(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r, \quad t < -\log(1-p)$$

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

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} I(x > 0), \qquad M(t) = \left(\frac{1}{1-\beta t}\right)^\alpha, \quad t < 1/\beta$$

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. A monkey sitting at a keyboard types 8 digits at random. The possibilities for each keystroke are $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Each keystroke is independent of the others with all 10 possibilities being equally likely. Let X_i denote the number of digits that the monkey types exactly i times. For example, if the monkey types 51022888, then $X_1 = 3, X_2 = 1, X_3 = 1$, and $X_i = 0$ for all $i \geq 4$. Let Y denote the number of times that the monkey types a digit greater than 6.

   (a) Find the probability that $Y \leq 6$ given that $Y > 2$.

   (b) Derive the probability distribution of X_4.

   (c) Calculate $P(X_4 = 1|Y = 0)$.

2. Let (X_1, Y_1) and (X_2, Y_2) be random points on the plane such that X_1, X_2, Y_1, and Y_2 are independent $N(\mu, \sigma^2)$. Let D^2 denote the squared distance between the two points.

   (a) Show that $D^2 \sim 2\sigma^2\chi^2_2$.

   (b) Derive the pdf of D.

   (c) Determine whether or not D^2 and $T = \frac{1}{4}(X_1 + X_2 + Y_1 + Y_2)$ are independent.

3. Let $Y_1, \ldots, Y_n$ be iid random variables from $f(y) = \theta y^{\theta-1} I(0 < y < 1)$. Let $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ be the sample mean.

   (a) Derive the asymptotic distribution of $\sqrt{n}\left(\frac{\overline{Y}}{1-\overline{Y}} - \theta\right)$.

   (b) Find the lower bound for the MSE of an unbiased estimator of $\frac{\theta}{\theta+1}$.

   (c) Find the UMVUE for $\frac{\theta}{\theta+1}$. Does it achieve the bound you found in part (b)?

   (d) Derive an asymptotic level $100(1-\alpha)\%$ confidence interval for $\frac{\theta}{\theta+1}$.

4. Let $X_1, \ldots, X_n$ be iid $N(\theta, \theta)$ random variables, where $\theta > 0$.

   (a) Find a minimal sufficient statistic for θ.

   (b) Show that both $\overline{X}$ and S^2, the sample mean and sample variance, respectively, are unbiased estimators of θ.

   (c) Consider estimators for θ of the form $a\overline{X} + (1-a)S^2$ for $0 \leq a \leq 1$. Find the estimator of this form with the smallest variance.

   (d) Determine whether your estimator from part (c) is consistent.

5. Let $X_1, \ldots, X_n$ be iid random variables with cdf $F(x) = \left(1 - (\theta^3/x^3)\right) I(x \geq \theta > 0)$.

   (a) Calculate the expected value and variance of the MOME of θ.

   (b) Calculate the expected value and variance of the MLE of θ.

   (c) If $\theta \sim$Exponential(1), calculate the PBE of θ. Your answer should be expressed in terms of the observations $x_1, \ldots, x_n$ and a **single** probability for a Gamma(α, β) random variable Y, where you must determine the values α and β.

6. Let $X_1, \ldots, X_n$ be a random sample from the Uniform$(\theta, \theta + 1)$ distribution.

   (a) If the MLE is unique, then find it. If not, state what the maximizers of the likelihood are.

   (b) Construct the LRT for $H_0 : \theta = \theta_0$ vs. $H_0 : \theta \neq \theta_0$.

   (c) Derive the power function for the test in part (b).