Probability and Statistics Preliminary Examination: May 2017

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.
  - $E(X)$: expectation of random variable X.
  - $Var(X)$: 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

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

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. An urn contains N balls numbered $1, 2, ..., N$. From these balls n balls are drawn at random without replacement. Let X_i denote the number that appears on the ith sampled ball and $T = \sum_{i=1}^{n} X_i$. Find $E(T)$ and $Var(T)$.

2. Let the random variables $X_1, X_2 \overset{iid}{\sim} Exponential(1)$. Define $U = X_1 + X_2$ and $V = X_1 - X_2$.
   (a) Find the marginal distributions of U and V.
   (b) Obtain $P(U > 1, V > 1)$
   (c) Find the correlation between U and V.
   (d) Are the random variables U and V indpendent? Provide your reasoning.

3. Suppose X is a non-negative random variable for which, for some positive integer, r, $E(X^r)$ does not exist. Obtain the necessary and sufficient condition for the moment generating function of X to exist.

4. Consider the model $Y_i = x_i\beta + \epsilon_i \;\; i = 1, 2, ..., n$, where β is an unknown parameter to be estimated. ϵ_i are iid random variables from a distribution whose moments all exist, in particular $E(\epsilon_i) = 0, \;\; \forall i$ and $Var(\epsilon_i) = \sigma^2 \;\; \forall i$. Then an estimate of β is given by

$$\hat{\beta}_n = \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i y_i}$$

   If $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i^2 = S_{xx}$ (a positive constant), then obtain the limiting distribution of $\hat{\beta}_n$.

5. Consider the following Poisson-Gamma model: $X_1, X_2, ..., X_n|\lambda \overset{iid}{\sim} Poisson(\lambda)$. The prior, $\pi(\lambda) \propto \lambda^{a-1}e^{-b\lambda}$, where a and b are known positive constants.
   (a) Obtain the mode[1] of the posterior distribution of λ.
   (b) Using the posterior mode obtained in part (a), show that, as $n \to \infty$, the posterior distribution of λ can be approximated by a $Normal$ distribution. Clearly state the assumptions you need to make.
   (c) Obtain the mean and variance of the $Normal$ distribution you obtained in part (b).

6. Let $X_1, X_2, ...X_n \overset{iid}{\sim} Uniform(0, \theta), \theta > 0$. We wish to test $H_0 : \theta \leq \theta_0$ Vs. $H_1 : \theta > \theta_0$.
   (a) Find an UMP level α test and derive the power function for this.
   (b) Consider another decision rule given by

$$\begin{array}{lll} \text{Reject } H_0 & \text{with probability 1 if} & x_{(n)} > \theta_0 \\ \text{Reject } H_0 & \text{with probability } \alpha \text{ if} & x_{(n)} < \theta_0 \end{array}$$

   Would you prefer this test to the UMP test obtained in part (a)? Why, why not?

[1] for a continuous r.v, X, with pdf $f_X(x)$, mode is defined as $\underset{x}{argmax}\ f_X(x)$