Probability and Statistics Preliminary Examination: May 2022

Instructions:

- Work all 5 Problems. Neither calculators nor electronic devices of any kind are allowed. Clearly state any theorem or fact that you use. Each of the 5 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.

  Bin(n, p)**:** $\mathbb{E}(X) = np$, $\mathbb{V}(X) = np(1-p)$, and pmf and mgf given respectively by

  $$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} I(x \in \{0, 1, \cdots, n\}), \qquad M(t) = (1 - p + pe^t)^n$$

  Unif(a, b)**:** $\mathbb{E}(X) = (a+b)/2$, $\mathbb{V}(X) = (b-a)^2/12$, and pdf given by

  $$f(x) = (b-a)^{-1} I(a < x < b)$$

  Pois(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

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

Exp(β)**:** Gamma$(1, \beta)$.

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

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

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

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

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

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

Problems

1. Suppose the joint pdf of X and Y is given by

$$f(x,y) = \begin{cases} kxy & \text{if } y < x < y+1, 0 < y < 1, \\ 0 & \text{elsewhere.} \end{cases}$$

   (a) Find the value of k.

   (b) Find the marginal pdf of Y.

   (c) Find the marginal pdf of X.

2. Consider three boxes and n marbles. For each marble, I will randomly put it into box i with probability p_i for $i = 1, 2, 3$. Suppose that there are Y_1, Y_3, Y_3 marbles in the three boxes.

   (a) Find a LRT statistic for the hypotheses $H_0 : p_1 + p_2 = 2p_3$ V.S. $H_1 : H_0$ is false.

   (b) Find the asymptotic size α LRT test.

3. Let $X_1, \dots, X_n$ be iid observations with pdf $f(x) = 3x^2 I(0 < x < 1)$. Let $Y_n = \max\{X_1, \dots, X_n\}$.

   (a) Find the cdf of Y_n.

   (b) Find the constant c such that $Y_n \xrightarrow{P} c$.

   (c) Let c be the constant in (b). Find the limiting distribution of $n(c - Y_n) \xrightarrow{D}$?

4. Let $X_1, \dots, X_n$ be independent random variables from Pois(λ). Let $T = \sum_{i=1}^n X_i$.
   (a) Find the conditional pmf of X_1 given $T = t$.
   (b) Find the UMVUE of $\lambda e^{-\lambda}$.

5. Let X be an observations with pdf $f(x|\theta) = \theta x^{\theta-1} I(0 < x < 1)$ with $\theta > 0$.
   (a) Find the UMP size α test for $H_0 : \theta = \theta_0$ V.S. $H_1 : \theta = \theta_1$. Here $0 < \theta_0 < \theta_1$ are two known constants.
   (b) Find the UMP size α test for $H_0 : \theta = \theta_0$ V.S. $H_1 : \theta > \theta_0$. Here $\theta_0 > 0$ is a known constant.
   (c) Find the power function of the test in (b).