Probability and Statistics Preliminary Examination: May 2013

<u>**Instructions:**</u>

- Work all 6 Problems. Neither calculators nor electronic devices of any kind are allowed. Clearly state every 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);
  - MSE (mean squared error), MOME (method of moments estimator); MLE (maximum likelihood estimator); UMVUE (uniform minimum variance unbiased estimator); UMP (uniformly most powerful); 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 and mgf given respectively by

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

  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 and mgf given respectively by

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

  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. Four (4) balls are to be distributed randomly into 4 urns. Let X denote the random variable that counts the number of urns containing exactly 1 ball. One way to determine the distribution of X is to count *patterns*. For example, the pattern $n_1 n_2 n_3 n_4$ denotes any allocation that results in n_1 balls in one urn, n_2 balls in another, n_3 balls in another, and n_4 balls in the remaining urn, in any order.

   (a) Reason that there are 256 equally likely sample points in this experiment.

   (b) Find $P(X = 4)$.

   (c) Find $P(X = 1)$.

   (d) Compute the remaining values of the pmf of X.

2. Let the discrete random variables (X, Y) have a joint pmf $f(x, y)$ given by the following table.

	X		
Y	$x = 1$	$x = 2$	$x = 3$
$y = 0$	0.15	0.10	0.30
$y = 1$	0.15	0.30	0.00

   Define $BUP(Y|X) = \mathbb{E}(Y|X)$ and $BLUP(Y|X) = a + bX$ to be, respectively, the *best unbiased predictor* and *best linear unbiased predictor* of Y based on X. If $\hat{Y}$ denotes any one of these predictors, then the "unbiased" condition ensures $\mathbb{E}(\hat{Y}) = \mathbb{E}(Y)$, and the "best" condition ensures $\hat{Y}$ is such that it minimizes the (prediction) MSE, $\mathbb{E}(Y - \hat{Y})^2$, in its respective class: either all functions of X for the BUP, or all linear functions of X for the $BLUP$.

   (a) For $BLUP(Y|X)$, show that the unbiased condition requires that $a = \mathbb{E}(Y) - b\mathbb{E}(X)$. Thus deduce that the value of b that minimizes the MSE is: $b = \text{Cov}(X, Y)/\mathbb{V}(X)$.

   (b) Compute $BUP(Y|X)$ and its MSE for (X, Y) with joint pmf as in the table.

   (c) Compute $BLUP(Y|X)$ and its MSE for (X, Y) with joint pmf as in the table.

3. Let $X_1, \ldots, X_n$ be a random sample from a Uniform distribution on $(\alpha - \beta, \alpha + \beta)$, where $\beta > 0$. Denote by $\overline{X} = \sum_{i=1}^{n} X_i/n$ the usual sample mean, and $X_{(1)} \leq \cdots \leq X_{(n)}$ the usual order statistics.

   (a) Show that $\sqrt{n}(\overline{X} - \alpha) \xrightarrow{d} \text{N}(0, \beta^2/3)$.

   (b) Find the limiting (asymptotic) distribution of

$$\frac{\sqrt{3n}(\overline{X}^2 - \alpha^2)}{2\alpha\beta}.$$

   (c) Does $X_{(n)} \xrightarrow{p} \alpha + \beta$? Justify your answer.

   (d) Deduce that the limiting (asymptotic) distribution of W_n is a pivotal quantity, where

$$W_n = \frac{\sqrt{n}(\overline{X} - \alpha)}{X_{(n)} - X_{(1)}},$$

   and thus show how it can be used to construct a $(1 - \gamma)100\%$ confidence interval for α, where $0 < \gamma < 1$.

4. Let $X_1, \ldots, X_n$ be a random sample from a Poisson(λ) distribution. We seek estimators of $\theta = e^{-\lambda} = P(X_1 = 0)$. To this end, define the statistics:

$$T = \sum_{i=1}^{n} X_i, \qquad \text{and} \qquad W = \{\text{number of } X_i\text{'s that equal } 0,\ i = 1, \ldots, n\}.$$

Several good estimators of θ are of the form $\hat{\theta} = a^{bT}$, for some constants $a > 0$ and $b \in \mathbb{R}$, possibly depending on n.

(a) Show that for estimators the form $\hat{\theta} = a^{bT}$, we have: $\mathbb{E}(\hat{\theta}) = \exp\{n\lambda(a^b - 1)\}$. Use this result to compute an expression for the MSE of $\hat{\theta}$.

(b) Find the UMVUE of θ. Does its variance attain the Cramer Rao Lower Bound (CRLB)?

(c) Find the MLE of θ. Is the MLE unbiased?

(d) Find the posterior Bayes estimator (PBE) of θ resulting from the (improper) prior, $\pi(\lambda) = \lambda^{-1} I(\lambda > 0)$. Is the PBE unbiased?

(e) Formulate an unbiased estimator of θ based on W, and compute its variance.

5. Let X_1, X_2 be a random sample of size 2 from a uniform distribution on the interval $(0, \theta)$, $\theta > 0$. Define $U = X_1 + X_2$ and $V = X_{(2)}$. In a *randomized* hypothesis test, the rejection rule is specified via a *critical function*, $\psi(\mathbf{x})$, which specifies the probability of rejecting H_0 given that the sample $\mathbf{x} = (x_1, \ldots, x_n)$ is observed. The *power* of such a test is defined to be $\beta(\theta) = \mathbb{E}_\theta[\psi(\mathbf{X})]$.

(a) Show that $U|V$ is uniformly distributed on the interval $(V, 2V)$.

(b) For testing $H_0 : \theta \leq 1$ vs. $H_1 : \theta > 1$, find the size and power function of the randomized test with critical function $\psi(v) = \mathbb{E}[U \geq 1 \mid V = v]$.

(c) Is the test in (b) UMP of its size? Justify.

6. For parameters $0 < \lambda_1 < \lambda_2 < \infty$, let the bivariate random vector (Y, Z) have joint pmf given by

$$f(y, z|\lambda_1, \lambda_2) = \frac{e^{-\lambda_2}}{(z-y)!y!} \lambda_1^y (\lambda_2 - \lambda_1)^{z-y} I_{\{0,1,\ldots,z\}}(y) I_{\{0,1,\ldots\}}(z).$$

Note that $Y \sim \text{Poisson}(\lambda_1)$, $Z \sim \text{Poisson}(\lambda_2)$, and $Y|Z \sim \text{Binomial}(Z, \lambda_1/\lambda_2)$. Assume that a random sample of size n from this joint pmf, $(Y_1, Z_1), \ldots, (Y_n, Z_n)$, is available to make inference on (λ_1, λ_2).

(a) Show that $(s, t) = (\sum_{i=1}^n y_i, \sum_{i=1}^n z_i)$, is a *complete* and *sufficient* statistic.

(b) Find the MLE of $\theta = \lambda_1/\lambda_2$.

(c) For $\theta = \lambda_1/\lambda_2$, construct the (asymptotic) size α LRT of $H_0 : \theta = 1/2$ vs. $H_1 : \theta \neq 1/2$.

(d) Find a $(1-\alpha)100\%$ confidence interval for $\lambda_2 - \lambda_1$ using a convenient asymptotic pivotal quantity.