Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Part I
August 26, 2014
12:30-4:30pm, Room 133 SMI

- There are a total of FOUR (4) problems in this exam. Please do a total of THREE (3) problems.
- Each problem must be done in a separate exam book.
- Please turn in THREE (3) exam books.
- Please write your code name and **NOT** your real name on each exam book.

(1) Let $X_1, \ldots, X_n$ be i.i.d. random variables following Beta distribution Beta$(\theta, 1)$ with the probability density function $f(x; \theta) = \theta x^{\theta-1}$, $\theta > 0$, $0 < x < 1$.

(a) Find the maximum likelihood estimator (MLE) of $1/\theta$. Calculate the expectation of the MLE and show whether or not it is unbiased.

(b) Calculate the Cramer-Rao lower bound for the variance of the unbiased estimators of $\theta/(\theta + 1)$. Find an unbiased estimator of $\theta/(\theta + 1)$. Calculate the variance of your unbiased estimator. Show whether or not your estimator achieves the lower bound.

(c) Let $X_1, \ldots, X_n$ be i.i.d. random variables following Beta distribution Beta(a, b) with the probability density function $f(x; a, b) = \frac{1}{\mathrm{B}(a,b)} x^{a-1}(1-x)^{b-1}$, $a > 0$, $0 < b < 1$, where $\mathrm{B}(a, b)$ is the Beta function. Let $Y_1, \ldots, Y_n$ be i.i.d. random variables following Beta distribution Beta$(a+b, 1-b)$. Assume that $X_i \perp Y_j$, $\forall i, j$. Suppose in practice we can only observe the product of these two Beta random variables, that is, observed data is $Z_1, \ldots, Z_n$ where $Z_i = X_i Y_i$, $i = 1, \ldots, n$. Derive an asymptotic 95% confidence interval for $1/a$ based on Z_i.

(2) Assume that $\{X_1, \ldots, X_n\} \overset{\text{i.i.d.}}{\sim}$ a cumulative distribution function (C.D.F.) G with support on $[0, 1]$. Fix $t_0 \in (0, 1)$. Define a constant

$$\pi = \frac{1 - G(t_0)}{1 - t_0}.$$

Define an estimator of π by

$$\widehat{\pi}_n = \min\{\widetilde{\pi}_n, 1\},$$

where

$$\widetilde{\pi}_n = \frac{1 - \mathbb{G}_n(t_0)}{1 - t_0}$$

and $\mathbb{G}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{I}(X_i \le t)$ denotes the empirical distribution function of $\{X_i\}_{i=1}^{n}$ with an indicator function $\mathrm{I}(\cdot)$. We wish to study the asymptotic properties of $\widehat{\pi}_n$ as $n \to \infty$, in the case of $G(t_0) \ge t_0$.

(a) If $G(t_0) > t_0$,

(i) Derive the asymptotic convergence (in probability) limit of $\widehat{\pi}_n$.

(ii) Derive the asymptotic distribution of $\widehat{\pi}_n$.

(iii) Give the explicit form of the C.D.F. of the asymptotic distribution in part (a)(ii).

(b) If $G(t_0) = t_0$,

(i) Derive the asymptotic convergence (in probability) limit of $\widehat{\pi}_n$.

(ii) Derive the asymptotic distribution of $\widehat{\pi}_n$.

(iii) Give the explicit form of the C.D.F. of the asymptotic distribution in part (b)(ii).

(3) Consider a linear regression problem where $Y \sim \mathcal{N}(X\beta, I_{n\times n})$, $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$, and $\beta \in \mathbb{R}^p$. Throughout we assume that $n \leq p$.

(a) For an estimator $\widehat{\beta}$, define the mean-squared prediction risk as $\mathcal{R}(\beta, \widehat{\beta}) = \mathbb{E}[\|X(\widehat{\beta}-\beta)\|_2^2]$. Prove that for any estimator, $\mathcal{R}(\beta, \widehat{\beta}) = \mathrm{Bias}^2(X\widehat{\beta}) + \mathrm{Var}(X\widehat{\beta})$ where $\mathrm{Bias}^2(X\widehat{\beta}) = \|\mathbb{E}[X\widehat{\beta}] - X\beta\|_2^2$ and $\mathrm{Var}(X\widehat{\beta}) = \mathbb{E}[\|X\widehat{\beta} - \mathbb{E}[X\widehat{\beta}]\|_2^2]$.

(b) Assume a Gaussian prior on β, with distribution $\beta \sim \mathcal{N}\left(0, \frac{1}{\lambda} I_{p\times p}\right)$ for a fixed $\lambda > 0$. Find the posterior density $\beta|Y$ and the corresponding *Maximum a Posteriori* (MAP) estimator. Note: The MAP estimator is defined to be the mode of the posterior density.

(c) Assume $\mathrm{rank}(X) = n$ and let $X = U\Sigma V^T$ be the singular value decomposition of X where $U \in \mathbb{R}^{n\times n}$, $\Sigma \in \mathbb{R}^{n\times p}$, and $V \in \mathbb{R}^{p\times p}$. For the MAP estimator in part (b), derive $\mathrm{Bias}^2(X\widehat{\beta})$ and $\mathrm{Var}(X\widehat{\beta})$ in terms of U, Σ, V, and λ.

(d) Find $\mathrm{Bias}^2(X\widehat{\beta})$ and $\mathrm{Var}(X\widehat{\beta})$ as $\lambda \to 0$ for the MAP estimator in part (b).

(4) With parameters α and β fixed, both in $(0, 1)$, consider a sequence of zero-one random variables $X_1, X_2, \ldots, X_N$ with a joint distribution defined as follows. First, $P(X_1 = 1) = \alpha/(\alpha + \beta)$. Next, for any $n = 1, 2, \ldots, N - 1$ and any $x_1, x_2, \ldots, x_n$ with all $x_i \in \{0, 1\}$,

$$P(X_{n+1} = 1 | X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = \begin{cases} \alpha & \text{if } x_n = 0, \\ 1 - \beta & \text{if } x_n = 1. \end{cases}$$

(a) Set $\theta = \alpha/(\alpha + \beta)$, and show that $P(X_n = 1) = \theta$ for all $n = 1, 2, \ldots, N$

(b) Consider $n = 1, 2, \ldots, N - 1$, and define $\theta_n = P(X_{n+1} = 1 | X_1 = 1)$. Treating $\theta_0 = 1$, show that $\theta_n = \alpha + \theta_{n-1}(1 - \alpha - \beta)$ for all n.

(c) Assume that $\alpha + \beta < 1$.

(i) Show that $\theta_n > \theta$ for all $n = 1, 2, \ldots N - 1$.

(ii) Show that $\text{var}(\sum_{n=1}^{N} X_n) > \sum_{n=1}^{N} \text{var}(X_n)$.

(d) Upon observing $X_n = x_n$ for all n, a likelihood function (not involving information from X_1) is defined: $L(\alpha, \beta) = P(X_2 = x_2, X_3 = x_3, \ldots, X_N = x_N | X_1 = x_1)$.

(i) Determine maximum likelihood estimators of $\widehat{\alpha}$ and $\widehat{\beta}$.

(ii) Derive the likelihood ratio test statistic for the null hypothesis $H_0 : \alpha + \beta = 1$. Evaluate this statistic for the following small ($N = 11$) data set: $0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0$.