Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Part I
August 25, 2015
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. Consider the linear regression model for independent data points $\{(\boldsymbol{X}_{k,i}, Y_{k,i}) : k = 1,\ldots,K;\ i = 1,\ldots,n\}$,

$$Y_{k,i} = \boldsymbol{X}_{k,i}^T \boldsymbol{\beta} + \epsilon_{k,i},$$

which is re-written in the matrix form,

$$\mathbf{y}_k = \mathbf{X}_k \boldsymbol{\beta} + \boldsymbol{\epsilon}_k,$$

where

$$\mathbf{y}_k = \begin{pmatrix} Y_{k,1} \\ \vdots \\ Y_{k,n} \end{pmatrix}, \quad \mathbf{X}_k = \begin{pmatrix} \boldsymbol{X}_{k,1}^T \\ \vdots \\ \boldsymbol{X}_{k,n}^T \end{pmatrix}, \quad \boldsymbol{\epsilon}_k = \begin{pmatrix} \epsilon_{k,1} \\ \vdots \\ \epsilon_{k,n} \end{pmatrix}, \quad \text{for } k = 1,\ldots,K,$$

with a finite integer $K \geq 2$. We assume that $E(\epsilon_{k,i} \mid \boldsymbol{X}_{k,i}) = 0$ and $\mathrm{var}(\epsilon_{k,i} \mid \boldsymbol{X}_{k,i}) = \sigma^2 \in (0,\infty)$, for $k = 1,\ldots,K$ and $i = 1,\ldots,n$. Let's compare two estimators of $\boldsymbol{\beta}$ formed by

$$\widehat{\boldsymbol{\beta}}_{\mathrm{I}} = \Big(\sum_{k=1}^{K} \mathbf{X}_k^T \mathbf{X}_k\Big)^{-1} \Big(\sum_{k=1}^{K} \mathbf{X}_k^T \mathbf{y}_k\Big),$$

and

$$\widehat{\boldsymbol{\beta}}_{\mathrm{II}} = \frac{1}{K} \sum_{k=1}^{K} (\mathbf{X}_k^T \mathbf{X}_k)^{-1} (\mathbf{X}_k^T \mathbf{y}_k).$$

(a) Compute $E(\widehat{\boldsymbol{\beta}}_{\mathrm{I}} \mid \mathbf{X}_1,\ldots,\mathbf{X}_K)$ and $E(\widehat{\boldsymbol{\beta}}_{\mathrm{II}} \mid \mathbf{X}_1,\ldots,\mathbf{X}_K)$.

(b) Compute $\mathrm{var}(\widehat{\boldsymbol{\beta}}_{\mathrm{I}} \mid \mathbf{X}_1,\ldots,\mathbf{X}_K)$ and $\mathrm{var}(\widehat{\boldsymbol{\beta}}_{\mathrm{II}} \mid \mathbf{X}_1,\ldots,\mathbf{X}_K)$.

(c) Compare the two covariance matrices in part (b) with derivations/justifications. Here for two symmetric non-negative definite matrices C_1 and C_2, $C_1 \geq C_2$ means that $C_1 - C_2$ is non-negative definite. (*Hint: you may use without proving the inequality* $E(AA^T) \geq E(AB^T)\{E(BB^T)\}^{-1}E(BA^T)$ *for appropriately dimensioned matrices* A *and* B.)

(d) From the computational perspective, which estimator will be more stable? Please provide your thoughts.

2. Let (X_i, Y_i), $i = 1, 2, \ldots, n$, be independent identically distributed random vectors with joint density function

$$\frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma^2} - \frac{2\rho xy}{\sigma\tau} + \frac{y^2}{\tau^2}\right)\right\}.$$

(a) Show that:

$$\begin{aligned} E(X_i^4) &= 3\sigma^4 \\ E(X_i^3 Y_i) &= 3\rho\sigma^3\tau \\ E(X_i^2 Y_i^2) &= (1+2\rho^2)\sigma^2\tau^2 \end{aligned}$$

(b) Let

$$S_x^2 = n^{-1}\sum_{i=1}^n X_i^2, \quad S_y^2 = n^{-1}\sum_{i=1}^n Y_i^2, \quad S_{xy} = n^{-1}\sum_{i=1}^n X_i Y_i$$

and

$$R = S_{xy}/(S_x S_y).$$

Show that $n^{1/2}(R-\rho) \to N(0, \delta^2)$ in distribution as $n \to \infty$ and find the value of δ^2.

3. With parameters $\beta, \gamma > 0$, random variables A, B, and C are mutually independent and Poisson distributed, with means $\beta\gamma$, β, and γ, respectively.

   (a) Define $X = A + B$ and $Y = A + C$ and put $D = (X, Y)^T$.

      i. Find the joint probability mass function of D.

      ii. Find the mean vector and covariance matrix of D.

      iii. Compute the Fisher information matrix for $(\beta, \gamma)^T$.

   (b) Data $D_1, D_2, \cdots, D_n$ are treated as mutually independent copies of D. Identify the maximum likelihood estimates of β and γ and confirm they are consistent.

   (c) Using data as in (b), consider testing the null hypothesis $H_0 : \beta = \gamma$ against the alternative $H_1 : \beta \neq \gamma$. Derive the generalized likelihood ratio statistic and describe a test that has a 5% level of significance.

   (d) Extending the model beyond Poisson components, how would you test the null hypothesis that the distribution of B equals the distribution of C, using data as in (b)?

4. Consider the *Gaussian sequence model* $y_i = \theta_i + \frac{w_i}{\sqrt{n}}$, for $1 \leq i \leq n$, and $w_i \sim \mathcal{N}(0,1)$ are independent and identically distributed. The goal is to estimate $\{\theta_i\}_{i=1}^n$ based on $\{y_i\}_{i=1}^n$.

   (a) Find the maximum likelihood estimator $\widehat{\theta}^{ML}$ for $\theta = (\theta_1, \theta_2, ..., \theta_n)$. Compute the squared bias, variance, and mean-squared error for $\widehat{\theta}^{ML}$.

   (b) Now consider the *shrinkage estimator* $\widehat{\theta}_i^{Shr,\alpha} = (1-\alpha)y_i$ for some $0 \leq \alpha \leq 1$. Assuming that $\|\theta\|_2^2 = \sum_{i=1}^n \theta_i^2 = 1$, find the squared bias, variance, and mean-squared error for $\widehat{\theta}^{Shr,\alpha}$ in terms of α. Find the value of α that minimizes the mean-squared error and the corresponding mean-squared error.

   (c) Show that $\widehat{\theta}^{Shr,\alpha}$ is a maximum a posteriori (MAP) estimator and find its corresponding prior distribution.

   (d) Now assume there exists an $S \subset \{1, 2, ..., n\}$ such that $|S| = s$ and $\theta_i \in \{-1, +1\}$ for $i \in S$ and $\theta_i = 0$ for $i \in S^c$. Find the mean-squared error for both $\widehat{\theta}^{ML}$ and $\widehat{\theta}^{Shr,\alpha}$ with the optimal choice of α (here, optimal refers to the value of α that minimizes mean-squared error). Can you suggest an estimator that might have lower mean-squared error in the regime where $1 \ll s \ll n$? Discuss.