May 2012

Ph.D. COMPREHENSIVE EXAMINATIONS

DEPARTMENT OF STATISTICS
UNIVERSITY OF TORONTO
THEORETICAL STATISTICS
COMPREHENSIVE EXAMINATION

May 18, 2012, 12:30 p.m. - 4:30 p.m.

Sidney Smith Hall

(# questions = 7); (#pages = 8 including cover page)

1. **ATTEMPT ALL QUESTIONS.**

   It is not necessary to completely solve every problem to achieve a good performance. Emphasize what you do know.

2. Please work neatly and legibly.

3. Start each question in a new book, with your name and the number of the question on the front cover. If there is more than one book for a question, then also indicate which is the first book and which second, e.g., Jane Smith, Question 5, Book 1 of 2.

4. The questions are not in any special order, nor are they all of equal difficulty.

5. The problems may be improperly phrased or may contain a misprint. Should this happen, reflect it in your discussion. Faculty members are not available to answer questions during the exam.

6. You are NOT permitted any aids (e.g., books, notes, etc.) aside from a single non programmable calculator.

7. Good luck!

1. Let $(X_i, Y_i)^\top$ take values $(1,0)^\top$, $(0,1)^\top$, $(1,1)^\top$, $(0,0)^\top$ with probabilities $p_{10}, p_{01}, p_{11}, p_{00}$, respectively, with $\sum_{i=0}^{1}\sum_{j=0}^{1} p_{ij} = 1$. Let $S_{xy} = \sum_i (X_i - \bar{X})(Y_i - \bar{Y})$, where $\bar{X} = n^{-1}\sum_i X_i$ and $\bar{Y} = n^{-1}\sum_i Y_i$. Assume that $p_{10} + p_{11} = p_{01} + p_{11} \equiv p$, derive the asymptotic distribution of the properly scaled S_{xy}.

2. Suppose that $X_1, \ldots, X_n$ is a sample from a Bernoulli(θ) distribution where $\theta \in [0, 1]$ is unknown and we want to estimate $\psi = \Psi(\theta) = \theta^2$.

   (a) What is the U-statistic estimator of ψ?

   (b) Is the U-statistic UMVU? If yes, justify your answer and if not, indicate how you would obtain the UMVU estimator for ψ.

   (c) Does the UMVU estimator attain the information lower bound? Justify your answer.

3. Suppose that we have k Poisson populations and independent samples of size n are drawn from each population. Denote the ith sample by $\{X_{i1}, \ldots, X_{in}\}$ that are i.i.d. Poisson(θ_i), $i = 1, \ldots, k$. We are interested in testing the null hypothesis $H_0 : \theta_1 = \ldots = \theta_k$ against all possible departures from this null.

   (a) Consider a test statistic $T_n = n \sum_{i=1}^{k} (\hat{\theta}_i - \bar{\theta})^2 / \bar{\theta}$, derive rigorously the asymptotic distribution of T_n.

   (b) Denote the likelihood ratio test statistic for testing H_0 by Λ_n, find the explicit expression of Λ_n and show that $2 \log \Lambda_n + T_n \xrightarrow{p} 0$ under H_0.

4. Suppose $(X, Y)^t \in R^{k+l}$ and

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim ?\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \right)$$

(meaning the random vector $(X, Y)^t$ has some unknown distribution with mean vector and variance matrix as specified).

(a) Let $G = \{g : R^k \to R^l : E(g(X)g^t(X)) \text{ exists finitely}\}$. Prove that G is a linear space and that $g(X) = E(Y \mid X)$ minimizes

$$E\left((Y - g(X))(Y - g(X))^t\right)$$

over $g \in G$ using the ordering on psd (positive semidefinite) matrices ($A \leq B$ whenever $y^t A y \leq y^t B y$ for every $y \in R^l$).

(b) Using (a) prove that $E(Y \mid X)$ is the orthogonal projection of Y onto G.

(c) Assume that $Var(X) = \Sigma_{XX}$ is invertible. Let

$$H = \{h : R^k \to R^l : h(X) = a + BX, \text{ for } a \in R^l, B \in R^{l \times k}\}.$$

Prove that H is a linear space and that

$$h(X) = \mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(X - \mu_X)$$

minimizes $E\left((Y - h(X))(Y - h(X))^t\right)$ using the ordering on psd matrices and is the orthogonal projection of Y onto H.

(d) Provide conditions under which $g = h$ where g and h are as specified in (a) and (c), respectively.

5. For a univariate c.d.f. $F(\cdot)$, let

$$\begin{aligned} T_1(F) &= \int [y - \mu(F)]^2 dF(y), \\ T_2(F) &= \int [y - \mu(F)]^3 dF(y), \\ T_3(F) &= T_2(F)/[T_1(F)]^{3/2}, \end{aligned}$$

where $\mu(F) = \int y dF(y)$. Let $F_n(\cdot)$ denote the empirical c.d.f. from a random sample $\{X_1, \ldots, X_n\}$ drawn from F, and assume that all necessary moments exist.

(a) Find the influence function of $T_3(F)$, denoted by $T_3'(x; F)$.

(b) Show that the remaining term

$$R_n = \sqrt{n}[T_3(F_n) - T_3(F)] - n^{-1/2} \sum_{i=1}^{n} T_3'(X_i; F) = o_p(1),$$

and derive the asymptotic distribution of the properly scaled $T_3(F_n)$.

6. Suppose we have a decision problem as specified by the statistical model $(\mathcal{X}, \mathcal{B}, \{P_\theta : \theta \in \Theta\})$, action space $\mathcal{A}$, correct action function $A : \Theta \to \mathcal{A}$ and loss function $L : \Theta \times \mathcal{A} \to [0, \infty)$.

   (a) Prove that, if a decision function δ is optimal with respect to minimizing expected loss for each $\theta \in \Theta$, then it it is degenerate at $A(\theta)$ a.s. P_θ.

   (b) Give an example of a decision problem for which there is no optimal δ.

   (c) Suppose that T is a sufficient statistic for $\{P_\theta : \theta \in \Theta\}$. If δ is a decision function, show how to find a decision function that depends on the data only through T and prove it has the same risk function as δ.

   (d) Prove that a Bayes rule depends on the data only through a sufficient statistic T.

7. Let F be a known absolutely continuous cumulative distribution function (c.d.f.) and assume that $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) with the c.d.f. $F(x|\theta) = [F(x)]^\theta$, $\theta > 0$.

   (a) Find the maximum likelihood estimator (MLE) $\hat{\theta}_n$ for θ, and its asymptotic distribution.

   (b) Find the asymptotic distribution of $\tilde{\theta}_n = \bar{y}/(1-\bar{y})$, where $\bar{y} = \sum_{i=1}^n F(X_i)/n$. Moreover, obtain an asymptotically efficient estimator based on $\tilde{\theta}_n$.

   (c) Let $U(x,\theta) = [F(x)]^\theta$, $0 < a \le \theta \le b < \infty$ for some $a, b > 0$. Let $S_n(\theta) = \sum_{i=1}^n U(X_i;\theta)$, rigorously justify and characterize the limiting behavior of $S_n(\hat{\theta}_n)/n$, where $\hat{\theta}_n$ is the MLE obtained in (a).