May 2015

Ph.D COMPREHENSIVE EXAMINATIONS
DEPARTMENT OF STATISTICAL SCIENCES
UNIVERSITY OF TORONTO

THEORETICAL STATISTICS COMPREHENSIVE EXAMINATION

May 19, 2015, 12:30 p.m. – 4:30 p.m.

Sidney Smith Hall

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

1. ATTEMPT ALL OF QUESTIONS 1-7

2. Please work neatly and legibly, and start each question in a new book With your name and the number of the question, then also indicate which is the first book for a question, then also indicate which is the first book and which is the second.

3. It is not necessary to completely solve every problem to achieve a good Performance: *emphasize what you know.*

4. The questions are not in any special order, and they are not of equal difficulty.

5. Like typical real problems, the problems here may be improperly phrased or may contain a misprint. Should this happen, reflect it in your discussion.

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

7. Good Luck!

1. Suppose we know that a response variable is generated by randomly selecting an individual π from a population Π and obtaining the measurement $x = X(\pi) \in \{1,2,3\}$. Suppose it is also known that the proportions of elements of Π taking these values is given by one of three possibilities, that we label by a, b and c, provided in the following table.

	$X=1$	$X=2$	$X=3$
a	1/3	1/3	1/3
b	1/6	1/6	2/3
c	1/2	1/6	1/3

Suppose we observe a single value x.

a) (5 marks) What is the statistical model, the parameter space Θ and the parameter θ of the model? Discuss what the "true value" of the parameter means.

b) (3 marks) Determine the MLE of θ as a function of x.

c) (3 marks) Determine the sampling distribution of the MLE in (b) for each value of θ.

d) (3 marks) Determine a minimal sufficient statistic and justify your answer.

e) (3 marks) Suppose we put a prior on Ω given by $\Pi(\theta = a) = 1/4$, $\Pi(\theta = b) = 1/4$, and $\Pi(\theta = c) = 1/2$. Determine the posterior distributions of θ given each possible data value.

f) (3 marks) If our goal is to estimate θ and the loss function is taken to be $L(\theta, z) = 1 - I_{\{\theta\}}(z)$, then determine a Bayes rule.

2. Given a parameter θ in $(0, 1/2)$, the random variables $X_1, X_2, \ldots, X_n$ with values in $\{1, 2, 3\}$ are i.i.d. with distribution given by

$$P(X_i = k \mid \theta) = \begin{cases} 1/2 & \text{if } k = 1 \\ \theta & \text{if } k = 2 \\ 1/2 - \theta & \text{if } k = 3 \end{cases}$$

a) (4 marks) Find a simple form for the minimal sufficient statistic for this model, and prove that it is minimal sufficient.

b) (4 marks) Find a non-constant ancillary statistic that is a function of the minimal sufficient statistic.

c) (1 mark) Prove that the minimal sufficient statistic is not complete.

d) (4 marks) Find the maximum likelihood estimator (MLE) for this model.

e) (4 marks) Find a simple expression for the observed information, and its value when θ is replaced by the MLE.

f) (4 marks) Find a simple expression for the Fisher information, and its value when θ is replaced by the MLE.

g) (4 marks) Discuss how one could best quantify how accurate the MLE for this model is as an estimate for θ, for the particular data set observed. Do *not* assume that n is very large.

3. Suppose $(X,Y)^T \in R^{k+l}$, with $X^T \in R^k, Y^T \in R^l$, 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) (5 marks) Let $G = \left\{ g : R^k \to R^l : E\left(g(X)g^T(X)\right) \text{ exists finitely} \right\}$. Prove that G is a linear space and that $g(X) = E(Y \mid X)$ minimizes

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

using the ordering on psd matrices given by $A \leq B$ whenever $y^T A y \leq y^T B y$ for every $y \in R^l$.

b) (5 marks) Using (a) prove that $E(Y \mid X)$ is the orthogonal projection of Y onto G.

c) (5 marks) Assume that $Var(X) = \Sigma_{XX}$ is invertible. Let

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

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

4. Let X_0, X_1, X_2, ... be a sequence of non-negative integer random variables. Conditional on a positive integer parameter θ, the joint distribution for this sequence is defined by $P(X_0 = 0) = 1$ and

$$P(X_{i+1} = k \mid X_0 = x_0, \ldots, X_i = x_i) = \begin{cases} 1/3 & \text{if } x_i > 0 \text{ and } x_i < \theta \text{ and } |k - x_i| \leq 1 \\ 2/3 & \text{if } x_i = 0 \text{ and } k = 0 \\ 2/3 & \text{if } x_i = \theta \text{ and } k = \theta \\ 0 & \text{otherwise} \end{cases}$$

That is, the sequence is a random walk on the integers in the range 0 to θ, starting at 0, with probabilities of 1/3, 1/3, and 1/3 of moving up one, moving down one, or staying at the same place, except that the probability of staying at the same place is 2/3 if the previous position was at 0 or θ, and the probability of moving outside the range 0 to θ is zero.
We observe $x_1, \ldots, x_n$, and wish to infer θ.

a) (6 marks) Write down a simple expression for the likelihood function for this model given the data $x_1, \ldots, x_n$.

b) (6 marks) Write down a simple form of the minimal sufficient statistic for this model, and prove that it is minimal sufficient.

c) (5 marks) Suppose that we decide to use a prior distribution in which the prior probability of θ is $2^{-\theta}$. Write down a simple expression for the posterior probability mass function of θ given the data $x_1, \ldots, x_n$.

   Also, give the posterior probability distribution explicitly for the following data set:

$$x_1 = 1,\ x_2 = 2,\ x_3 = 2,\ x_4 = 1,\ x_5 = 2,\ x_6 = 1$$

d) (4 marks) Prove that, for any n, there is no unbiased estimator of θ based on $x_1, \ldots, x_n$.

e) (4 marks) Find an estimator for θ that converges to θ with probability 1 as n goes to infinity.

5. Suppose that f is a density on $\mathcal{X} = R^n$ with respect to Lebesgue measure and the statistical model is given by the set of density functions $\{\theta^{-n} f(x/\theta) : \theta > 0\}$.

a) (2 marks) Show that we can consider Θ as a group with product $\theta_1 \cdot \theta_2 = \theta_1 \theta_2$.

b) (2 marks) Prove that $T_\theta x = \theta x$ defines an action of G on $\mathcal{X}$.

c) (2 marks) Determine the set which must be deleted from the sample space so that G acts freely.

d) (2 marks) Suppose that we want to estimate $\Psi(\theta) = \theta$ and the loss function is given by $L(\theta, \psi) = \rho(\psi/\theta)$ for some convex function $\rho : (0, \infty) \to R^1$. Prove that G leaves this decision problem invariant.

e) (3 marks) Show that $l(x) = T_{\|x - \bar{x}\|}$ is equivariant and determine a maximal invariant $m(x)$.

f) (3 marks) Putting $s = \|x - \bar{x}\|$ and $u = m(x)$, show that $J(x \to s, u) = s^{n-1} h(u)$ for some function h.

g) (3 marks) Determine the form of the Pitman estimate.

h) (3 marks) Determine the Pitman estimate when f corresponds to the $N_n(0, I)$ distribution and $\rho(z) = z^2$.

6. (10 marks) Suppose that $\mathcal{X} = [0,2]$, $\theta \in [0,1]$ and

$$f_\theta(x) = \begin{cases} (2\theta - 1)\,x + (1-\theta) & x \in [0,1] \\ (1-2\theta)\,x + (3\theta - 1) & x \in [1,2]\,. \end{cases}$$

Show that this family has monotone likelihood ratio form and find the form of the UMP size α test for $H_0 = \{\theta \le \theta_0\}$ versus $H_a = \{\theta > \theta_0\}$.

7. Let X_1, X_2, ... be a sequence of random variables, all with the same range, G, which is a finite set.

a) (6 marks) Suppose that given a parameter θ, the X_i are independent, and all have the same distribution, specified by θ. Suppose that for each n, the statistic $S_n(x_1, \ldots, x_n)$ is sufficient for this model. Suppose that we give some prior distribution to θ. Prove that for each n, the predictive distribution for X_{n+1} given $x_1, \ldots, x_n$ (*not* conditional on θ) can be written as a function of only $S_n(x_1, \ldots, x_n)$. You may use the definition of a sufficient statistic and the usual factorization theorem in your proof.

b) (14 marks) Suppose that the infinite sequence X_1, X_2, ... is exchangeable, and that there is a set A (a subset of G) such that for each n,

$$P(X_1 = x_1, \ldots, X_n = x_n) = \begin{cases} a_n & \text{if } x_1, \ldots, x_n \text{ are all in } A \\ b_n & \text{otherwise} \end{cases}$$

where a_n and b_n are constants, which may be different for each n, and which must of course result in these probabilities summing to 1, and with the marginal distributions obtained from the above with a certain n matching the results with smaller n.

By De Finetti's Representation Theorem, there must be some parameter, θ, with some prior distribution, $\pi(\theta)$, such that these probabilities can be written as

$$P(X_1 = x_1, \ldots, X_n = x_n) = \int \prod_{i=1}^{n} f(x_i|\theta)\, d\pi(\theta)$$

for some probability function $f(x|\theta)$.

Give a non-trivial example of model of this sort — that is, specify G, A, a_n, and b_n, or alternatively, G, π, and f. Discuss in general what form f can have for a model of this sort.