Imperial College London

UNIVERSITY OF LONDON

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2006

This paper is also taken for the relevant examination for the Associateship.

M3S3/M4S3

STATISTICAL THEORY II

Date: Tuesday, 23rd May 2006 Time: 10 am – 12 noon

Credit will be given for all questions attempted but extra credit will be given for complete or nearly complete answers.

Calculators may not be used.

1. (a) Consider the statement

A sequence of random variables $\{X_n\}$ converges almost surely to random variable X as $n \longrightarrow \infty$, denoted

$$X_n \xrightarrow{a.s.} X.$$

(i) Give two equivalent definitions of this statement.

(ii) Show that

$$X_n \xrightarrow{a.s.} X \quad \Longrightarrow \quad X_n \xrightarrow{p} X,$$

as $n \longrightarrow \infty$, that is if $\{X_n\}$ converges almost surely to X, then $\{X_n\}$ *converges in probability* to X also.

(b) Let $Z \sim Uniform(0,1)$, and define a sequence of random variables $\{X_n\}$ by

$$X_n = nI_{[1-n^{-1},1)}(Z) \qquad n = 1, 2, \ldots$$

where, for set A

$$I_A(Z) = \begin{cases} 1 & Z \in A \\ 0 & Z \notin A \end{cases}$$

that is, I_A is the indicator random variable associated with the set A.

Does the sequence $\{X_n\}$ converge in any mode (that is *almost sure*, *in probability* or *in law*) to any limit random variable ? Justify your answer.

(c) Suppose, for $n = 1, 2, \ldots$, $X_n \sim Bernoulli(p_n)$ are a sequence of independent random variables where

$$P[X_n = 1] = p_n = \frac{1}{\sqrt{n}} .$$

Does $P[\ X_n = 1$ infinitely often $] = 1$? Justify your answer.

2. (a) State the Central Limit Theorem for a sequence of independent and identically distributed random $k \times 1$ vectors $\{\boldsymbol{X}_n\}$ with

$$E_{f_{\boldsymbol{X}_n}}[\boldsymbol{X}_n] = \boldsymbol{\mu} \qquad Var_{f_{\boldsymbol{X}_n}}[\boldsymbol{X}_n] = \Sigma$$

with $\boldsymbol{\mu}$ and Σ finite, with Σ positive definite. Suppose that $\boldsymbol{Y_n} = \boldsymbol{g}(\boldsymbol{X_n})$, where $\boldsymbol{g}$ is a mapping

$$\boldsymbol{g} : \mathbb{R}^k \longrightarrow \mathbb{R}^d$$

for $d \leq k$, with continuous first partial derivative matrix $\dot{\boldsymbol{g}}$ of dimension $(d \times k)$.

Derive asymptotic normal approximations to the distribution of

$$\boldsymbol{M}_n = \overline{\boldsymbol{Y}}_n = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Y}_i$$

and, if $k = d = 1$, the distribution of

$$g(M_n) = M_n(M_n + 1).$$

(b) Suppose that $X_1, X_2, \ldots$ are independent $Exponential(\phi)$ random variables, with probability density function

$$f_{X|\phi}(x|\phi) = \phi e^{-\phi x} \qquad x > 0$$

and zero otherwise, for parameter $\phi > 0$. Let β_L and β_U be the lower and upper quartiles of the distribution, that is, the real values satisfying

$$F_X(\beta_L) = \frac{1}{4} \qquad F_X(\beta_U) = \frac{3}{4}$$

where F_X is the cumulative distribution function corresponding to f_X.

Find the asymptotic variance of the sample interquartile range, R_n

$$R_n = U_n - L_n$$

where L_n and U_n are the sample lower and upper quartiles respectively derived from a sample of size n.

Hint: find the asymptotic distribution of the sample quartiles in this case using the general result, then use a suitable transformation.

3. (a) Suppose $X_1, \ldots, X_n$ are independent Gamma random variables, $X_i \sim Gamma(\alpha, \beta)$, with probability density function $f_{X|\boldsymbol{\theta}}(x|\boldsymbol{\theta})$ given by

$$f_{X|\boldsymbol{\theta}}(x|\boldsymbol{\theta}) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\} \qquad x > 0$$

and zero otherwise, where $\boldsymbol{\theta} = (\alpha, \beta)^\mathsf{T}$, Γ is the Gamma function

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

and parameter space $\Theta \equiv \mathbb{R}^+ \times \mathbb{R}^+$.

(i) Find the (unit) Fisher information for $\boldsymbol{\theta}$; leave your answer in terms of derivatives of the Gamma function.

(ii) Show that, if $\beta = 1$, then

$$E_{f_{X|\boldsymbol{\theta}}}[\log X] = \frac{\dot{\Gamma}(\alpha)}{\Gamma(\alpha)}$$

where $\dot{\Gamma}$ is the first derivative of the Gamma function. Hence verify that the expected score is zero.

Note: here, exchanging the order of differentiation and integration is permissible.

(iii) Are parameters α and β orthogonal ? Justify your answer.

(b) Suppose that $X_1, \ldots, X_n$ are independent and identically distributed with probability density function $f_{X|\boldsymbol{\theta}}(x|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a $d \times 1$ vector.

(i) State the theorem describing the existence, weak consistency and asymptotic normality of solutions to the likelihood equations. Outline the key regularity assumptions under which this theorem holds.

(ii) Find the asymptotic normal distribution of the consistent estimator found as the solution to the likelihood equation when $d = 1$, and $f_{X|\theta}(x|\theta)$ is given by

$$f_{X|\theta}(x|\theta) = \theta x^{\theta-1} \qquad 0 < x < 1$$

for parameter $\theta > 0$.

4. (a) (i) Describe what is meant by the concept of *efficiency* in the estimation of a $d \times 1$ parameter $\boldsymbol{\theta}$.

(ii) For $d = 1$, derive the information inequality, and hence derive a lower bound on the variance of unbiased estimators of θ when data $x_1, \ldots, x_n$ are available in a model satisfying standard regularity conditions.

(iii) State the equivalent lower bound for $d > 1$. Give careful definition of the relevant concepts in this vector parameter case.

(b) (i) Explain how consistent but inefficient estimators may be modified and improved using *scoring* or *one-step* procedures. State, without proof, any relevant theorems concerning the efficiency of the modified estimators.

(ii) Discuss, without computing derivatives explicitly, the use of the one-step procedure for the model with $d = 1$ for the probability density function $f_{X|\theta}(x|\theta)$ given by

$$f_{X|\theta}(x|\theta) = \frac{\exp\{-(x-\theta)\}}{(1+\exp\{-(x-\theta)\})^2} \qquad -\infty < x < \infty$$

Note: this density is symmetric about θ.

5. (a) Suppose $X_1, \ldots, X_n$ are independent Poisson random variables, $X_i \sim Poisson(\lambda)$, with probability density function $f_{X|\lambda}(x|\lambda)$ given by

$$f_{X|\lambda}(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, \ldots$$

for parameter $\lambda > 0$.

Find the posterior distribution for λ, in light of data $x_1, \ldots, x_n$ under the following prior specifications:

(i) The *proper* $Gamma(1, 1)$ prior

$$p_\lambda(\lambda) = e^{-\lambda} \qquad \lambda > 0$$

(ii) The *non-informative* improper prior

$$p_\lambda(\lambda) = 1 \qquad \lambda > 0$$

(iii) *Jeffreys' prior*

$$p_\lambda(\lambda) \propto |I(\lambda)|^{1/2} \qquad \lambda > 0$$

where I is the (unit) Fisher information for the model.

Find a prior such that the posterior expectation is equal to the ML estimator, provided that at least one $x_i > 0$.

(b) The Bayes estimator, $\widehat{\lambda}_n^{(B)}$ under a quadratic loss function is the posterior expectation. Show that for each prior in part (a), $\widehat{\lambda}_n^{(B)}$ is a consistent estimator (in the frequentist sense) of λ, that is

$$\widehat{\lambda}_n^{(B)} \xrightarrow{a.s.} \lambda$$

as $n \longrightarrow \infty$.

(c) Find the *posterior predictive* distribution for new data random variable Y, independent from and identically distributed to $X_1, \ldots, X_n$, in light of the data $\boldsymbol{x} = (x_1, \ldots, x_n)$, denoted $f_{Y|\boldsymbol{X}}(y|\boldsymbol{x})$ and defined by

$$f_{Y|\boldsymbol{X}}(y|\boldsymbol{x}) = \int f_{Y|\lambda}(y|\lambda) p_{\lambda|\boldsymbol{X}}(\lambda|\boldsymbol{x}) \, d\lambda$$

under Jeffreys' Prior.