UNIVERSITY OF LONDON

FOR EXTERNAL STUDENTS

B. Sc. Examination 2003

STATISTICS

ST305 Statistical Inference

Duration: 2 hours 15 minutes

Date and time:

Full marks may be obtained for complete answers to FOUR questions.

A formula sheet is attached to the end of this examination paper.

Electronic calculators may be used. The make and model should be specified on the script. The calculator must not be programmed prior to the examination. Calculators which display graphics, text or algebraic equations are not allowed.

NOTE: Full details of all calculations are to be shown; pre-programmed statistical tests and procedures on a calculator, apart from mean and standard deviation, must not be used.

WHITE, YEATS & SKIPWORTH: Tables for Statisticians to be provided.

THIS EXAMINATION PAPER MUST NOT BE REMOVED FROM THE EXAMINATION ROOM

Question 1 Suppose that $X_1, \ldots, X_n$ is a random sample of n independent observations from a distribution with p.d.f.

$$f(x \mid \theta) = \frac{x}{\theta} e^{-x^2/2\theta}, \qquad x \geq 0.$$

(a) Determine the Cramer-Rao lower bound (CRLB) for the variance of unbiased estimators of θ. [10]

(b) Does there exist an unbiased estimator which attains the CRLB? If so, find it. [8]

(c) Now consider the unbiased estimation of $\sqrt{\theta}$. Determine the CRLB, and say whether or not it can be attained. [7]

Question 2 Let $X_1, \ldots, X_n$ be a random sample from one of the distributions below.

(a) In which case is the only sufficient statistic the set of order statistics $X_{(1)}, \ldots, X_{(n)}$? [3]

(b) In each of the other cases find a suitable sufficient statistic for θ:

(i) $N(\theta, \theta)$ $(\theta > 0)$

(ii) $N(\theta, \theta^2)$

(iii) $f(x \mid \theta) = \frac{1}{2} \exp(- \mid x - \theta \mid)$

(iv) $U(0, \theta)$

(v) $f(x \mid \theta) = (1-p)p^{x-\alpha}$, $x = \alpha, \alpha+1, \ldots,$ $\qquad \theta = (p, \alpha)$

(vi) $f(x \mid \theta) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log x - \mu)^2\right)$, $x > 0$, $\theta = (\mu, \sigma^2)$.

[22]

Question 3 The number of beta-particles given off from a piece of radioactive material in a one minute interval is Poisson distributed with known mean λ, and the number given off in n disjoint one minute intervals are independent. A machine used for counting these beta-particles records the value zero with probability θ $(0 < \theta < 1)$ whatever the number of particles, and accurately records the true number (which may be zero) with probability $(1 - \theta)$, independently of other time periods.

Suppose that the recorded values are $X_1, \ldots, X_n$ and θ is unknown.

(a) Show that the joint mass function of these data is given by

$$f_{\mathbf{X}}(\mathbf{x} \mid \theta) = (\theta + (1-\theta)e^{-\lambda})^m (1-\theta)^{n-m} e^{-\lambda(n-m)} \lambda^{\sum_{i=1}^n x_i} / \prod_{i=1}^n x_i!,$$

where $M = m$ is the number of zeroes amongst the data $\mathbf{x} = (x_1, \ldots, x_n)$. [6]

(b) Hence show that M is a complete and sufficient statistic for θ. [7]

(c) Hence find a UMVUE of θ, quoting any theorems you may use. [12]

Question 4 (a) State clearly but do not prove the Neyman-Pearson Lemma. [9]

(b) Under what conditions can the test procedure from the above lemma be uniformly most powerful (UMP)? [4]

(c) If $f(x \mid \theta) = x^2 e^{-x/\theta}/2\theta^3$, show that the UMP test (amongst tests of size α) of

$$H_0 : \theta = \theta_1 \quad \text{versus} \quad H_1 : \theta > \theta_1$$

is to reject H_0 if

$$\sum_{i=1}^{n} x_i \geq \frac{\theta_1}{2} \chi^2_{6n}(\alpha).$$ [12]

Question 5 Let X_{ij} be independent Poisson random variables with means θ_i, $j = 1, 2, \ldots, n$, $i = 1, 2, \ldots, p$.

(a) Derive the generalized likelihood ratio test statistic for testing

$H_0 : \theta_1 = \theta_2 = \cdots = \theta_p$ versus H_1 : at least two of the θ_i, $i = 1, \ldots, p$ are different. [17]

(b) Hence show that a test of approximate size α is to reject H_0 if

$$\sum_{i=1}^{p} T_i \log T_i - T \log(T/p) > \frac{1}{2} \chi^2_{p-1}(\alpha),$$

where $T_i = \sum_{j=1}^{n} x_{ij}$, $i = 1, \ldots, p$ and $T = \sum_{i=1}^{p} T_i$. [8]