UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Tuesday 4 June 2002 1.30 to 4.30

PAPER 38

STATISTICAL THEORY

*Attempt **FOUR** questions*

*There are **six** questions in total*

The questions carry equal weight

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Describe in detail the p^* approximation for the density of the maximum likelihood estimator.

Consider two independent samples of independent exponential random variables, each of size n, and with means $1/\lambda$ and $1/(\psi\lambda)$, respectively.

Find the p^* approximation to the density of $(\hat{\psi}, \hat{\lambda})$, and hence find an approximation to the marginal density of $\hat{\psi}$. Comment on the exactness of the approximation.

[*You may assume that the exact distribution of* $\hat{\psi}/\psi$ *is an* F*-distribution with degrees of freedom* $(2n, 2n)$ *so that the exact density of* $\hat{\psi}$ *is given by*

$$\frac{\Gamma(2n)}{\Gamma(n)} \frac{1}{\psi} \left(\frac{\hat{\psi}}{\psi}\right)^{n-1} \left(\frac{\hat{\psi}}{\psi} + 1\right)^{-2n} \quad . \quad]$$

2 Write a brief account of the concept and properties of *profile likelihood.*

Define what is meant by *modified profile likelihood.*

Let $Y_1, \ldots, Y_n$ denote independent exponential random variables, such that Y_j has mean $\lambda \exp(\psi x_j)$, where $x_1, \ldots, x_n$ are scalar constants and ψ and λ are unknown parameters.

You may assume that in this model the maximum likelihood estimators are not sufficient and an ancillary statistic is needed. Let

$$a_j = \log Y_j - \log \hat{\lambda} - \hat{\psi} x_j, \quad j = 1, \ldots, n,$$

and take $a = (a_1, \ldots, a_n)$ as the ancillary.

Find the form of the profile log-likelihood function and of the modified profile log-likelihood function for ψ.

[*You are* not *required to show that* a *is ancillary.*]

3 Explain in detail what is meant by a *transformation model.*

What is meant by (i) a *maximal invariant*, (ii) an *equivariant estimator*, in the context of a transformation model?

Describe in detail how an equivariant estimator can be used to construct a maximal invariant. Illustrate the construction for the case of a *location-scale-model.*

4 Explain the terms (i) *functional statistic*, (ii) *influence function* of a functional T at a distribution F, as used in robustness theory.

Derive the relationship between the influence function and the asymptotic variance of a functional statistic.

Give a brief account of robustness measures derived from the influence function.

The p quantile q_p of a distribution F with density f is defined as the solution to the equation $F\{q_p(F)\} = p$. Find the form of the influence function of $q_p(F)$.

5 Let $X_1, \ldots, X_n$ be a random sample from a continuous distribution that is symmetric about the unknown median θ, $-\infty < \theta < \infty$.

Explain carefully how to test $H_0 : \theta = 0$ against $H_1 : \theta > 0$ using the Wilcoxon signed rank test.

Show that the null mean and variance of the Wilcoxon signed rank statistic are $n(n+1)/4$ and $n(n+1)(2n+1)/24$ respectively. State, without proof, the asymptotic null distribution of this statistic.

Each of the $n(n+1)/2$ averages $(X_i + X_j)/2$, $i \leqslant j = 1, \ldots, n$, is called a *Walsh average.*

It is proposed to test H_0 against H_1, using as statistic the total number of Walsh averages greater than 0. Show that this test is equivalent to the Wilcoxon signed rank test.

6 Write an account of *one* of the following:

(i) saddlepoint approximation methods;

(ii) the importance of parameter orthogonality in parametric inference;

(iii) ancillary statistics and conditional inference.