UNIVERSITY OF CAMBRIDGE

M. PHIL. IN STATISTICAL SCIENCE

Thursday 2 June 2005 1.30 to 4.30

STATISTICAL THEORY

*Attempt **FOUR** questions, not more than **TWO** of which should be from Section B.*

*There are **TEN** questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script paper

SPECIAL REQUIREMENTS
None

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

Section A

1 Explain briefly the concepts of *profile likelihood* and *conditional likelihood*, for inference about a parameter of interest ψ, in the presence of a nuisance parameter λ.

Suppose $Y_1, \ldots, Y_n$ are independent, identically distributed from the exponential family density

$$f(y; \psi, \lambda) = \exp\{\psi \tau_1(y) + \lambda \tau_2(y) - d(\psi, \lambda) - Q(y)\},$$

where ψ, λ are both scalar.

Obtain a saddlepoint approximation to the density of $S = n^{-1} \sum_{i=1}^{n} \tau_2(Y_i)$.

Show that use of the saddlepoint approximation leads to an approximate conditional log-likelihood function for ψ of the form

$$l_p(\psi) + B(\psi),$$

where $l_p(\psi)$ is the profile log-likelihood, and $B(\psi)$ is an adjustment which you should specify carefully.

2 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.

3 Let $Y_1, \ldots, Y_n$ be independent, identically distributed $N(\mu, \mu^2), \mu > 0$.

Show that this model is an example of a *curved exponential family*, and find a minimal sufficient statistic.

Show that

$$a = \sqrt{n} \frac{(\sum Y_i^2)^{1/2}}{\sum Y_i}$$

is an ancillary statistic.

Assume that $a > 0$. Show that the maximum likelihood estimator of μ is

$$\hat{\mu} = \frac{(\sum Y_i^2)^{1/2}}{q\sqrt{n}},$$

where $q = \{(1 + 4a^2)^{1/2} + 1\}/(2a)$.

Show further that (apart from a constant) the log-likelihood may be written

$$l(\mu; \hat{\mu}, a) = -\frac{n}{2\mu^2}\left(q^2\hat{\mu}^2 - \frac{2q\mu\hat{\mu}}{a}\right) - n\log\mu,$$

and obtain the p^* approximation to the (conditional) density of $\hat{\mu}$.

How would you approximate $\text{Prob}(\hat{\mu} \leqslant t|a)$, for given t?

4 Explain what is meant by an *M-estimator* of a parameter θ, based on a given ψ function. Show that under appropriate conditions allowing the interchange of order of integration and differentiation, the influence function is proportional to ψ and derive an expression for the asymptotic variance $V(\psi, F)$ of the M-estimator at a distribution F.

A location model on $\mathbb{R}$, with parameter space $\mathbb{R}$, is specified by $F_\theta(x) = F(x - \theta)$, and an M-estimator is constructed using a ψ function of the form $\psi(x, \theta) = \psi(x - \theta)$.

For the particular choice

$$\psi(x) = \min\{b, \max\{x, -b\}\}, \quad b < \infty \ :$$

(i) Find the asymptotic variance $V(\psi, \Phi)$, where Φ is the standard normal distribution;

(ii) Verify that the estimator is B-robust, by determining an explicit bound on the influence function.

5 Let $Y_1, \ldots, Y_n$ be independent, identically distributed from a distribution F, with density f symmetric about an unknown point θ. Suppose we wish to test $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$.

Explain how to test H_0 against H_1 using (i) the sign test, *and* (ii) the Wilcoxon signed rank test.

Show that the null mean and variance of the Wilcoxon signed rank statistic are $\frac{1}{4}n(n+1)$ and $\frac{1}{24}n(n+1)(2n+1)$ respectively.

What is meant by a one-sample U-statistic?

State, without proof, a result concerning the asymptotic distribution of a one-sample U-statistic, and use it to deduce asymptotic normality of the Wilcoxon signed rank statistic.

6 Write brief notes on *four* of the following:

(i) Edgeworth expansion;

(ii) parameter orthogonality;

(iii) Laplace approximation;

(iv) Bartlett correction;

(v) the invariance principle;

(vi) finite-sample versions of robustness measures;

(vii) tests based on the empirical distribution function;

(viii) large-sample likelihood theory.

Section B

7 Assume that the n-dimensional observation vector Y may be written

$$Y = X\beta + \epsilon,$$

where X is a given $n \times p$ matrix of rank p, β is an unknown vector, and

$$\epsilon \sim N_n(0, \sigma^2 I).$$

Let $Q(\beta) = (Y - X\beta)^T(Y - X\beta)$. Find $\hat{\beta}$, the least-squares estimator of β, and show that

$$Q(\hat{\beta}) = Y^T(I - H)Y$$

where H is a matrix that you should define.

If now $X\beta$ is written as $X\beta = X_1\beta_1 + X_2\beta_2$, where $X = (X_1 : X_2)$, $\beta^T = (\beta_1^T : \beta_2^T)$, and β_2 is of dimension p_2, state without proof the form of the F-test for testing $H_0 : \beta_2 = 0$.

What is meant by saying that β_1 is orthogonal to β_2? What is the practical relevance of orthogonality?

8 Suppose that $Y_1, \ldots, Y_n$ are independent binomial observations, with

$$Y_i \sim B(t_i, \pi_i) \text{ and } \log(\pi_i/(1 - \pi_i)) = \beta^T x_i, \quad \text{for } 1 \leqslant i \leqslant n,$$

where $t_1, \ldots, t_n$ and $x_1, \ldots, x_n$ are given. Discuss carefully the estimation of β.

Your solution should include

(i) the method of checking the fit of the above logistic model,

and

(ii) the method for finding an approximate 95% confidence interval for β_2, the second component of the vector β.

9 Write a brief account of Bayesian decision theory. Derive the general form of a Bayes rule, in terms of the posterior distribution and loss function.

Let $X_1, \ldots, X_n$, conditional on θ, be a random sample from a $N(\theta, \sigma^2)$ population and let the prior distribution on θ be $N(\mu, \tau^2)$, where the values of σ^2, μ and τ^2 are known. Consider testing $H_0 : \theta \geqslant \theta_0$ against $H_1 : \theta < \theta_0$, with loss function,

$$L(\theta, \text{accept } H_0) = \begin{cases} 0 & \text{if } \theta \geqslant \theta_0 \\ 1 & \text{if } \theta < \theta_0 \end{cases}$$

and

$$L(\theta, \text{accept } H_1) = \begin{cases} 1 & \text{if } \theta \geqslant \theta_0 \\ 0 & \text{if } \theta < \theta_0 \end{cases} .$$

Show that the Bayes test rejects H_0 if

$$\bar{x} < \theta_0 - \frac{\eta(\mu - \theta_0)}{1 - \eta},$$

where $\eta = \sigma^2/(n\tau^2 + \sigma^2)$ and $\bar{x} = n^{-1} \sum_{i=1}^{n} x_i$.

10 Let $X_1, \ldots, X_n$ be independent identically distributed random variables, with joint density $f(x; \theta)$, where θ is a scalar parameter.

What is meant by the *Fisher information* $I_n(\theta)$ about θ contained in the sample $X_1, \ldots, X_n$?

State the Cramer–Rao lower bound for the variance of an arbitrary unbiased estimator of θ. Give the condition for this lower bound to be attained.

Let $X_1, \ldots, X_n$ be an independent, identically distributed sample of size $n \geqslant 3$ from the exponential density $\theta e^{-\theta x}$, $x > 0$, where $\theta > 0$ is unknown.

Obtain the maximum likelihood estimator $\hat{\theta}$ of θ. Show that it is biased, but a multiple of it is not. Determine whether the Cramer–Rao lower bound is attained by the variance of this unbiased estimator.

State carefully, but without proof, the asymptotic distribution of $\hat{\theta}$.

Explain in detail the rationale for the Wald test of the hypothesis $H_0 : \theta = \theta_0$, and find the form of the test statistic for the above exponential example.

[*Hint: You may assume that the density of* $X_1 + \ldots + X_n$ *is* $\dfrac{\theta^n t^{n-1} e^{-\theta t}}{(n-1)!}, \quad t > 0$.]

END OF PAPER