UNIVERSITY OF CAMBRIDGE

M. PHIL. IN STATISTICAL SCIENCE

Thursday 5 June 2003 9 to 12

PAPER 41

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.

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

Section A

1 (i) Let $Y_1, \ldots, Y_n$ be independent, identically distributed exponential random variables with common density $f(y\,;\,\lambda) = \lambda e^{-\lambda y}, y > 0$, and suppose that inference is required for $\theta = E(Y_1)$.

Find the maximum likelihood estimator of θ, and explain carefully why, with $\bar{Y} = n^{-1}\sum_{i=1}^{n} Y_i$ and Φ the distribution function of $N(0,1)$,

$$\left(\frac{\bar{Y}}{1 + n^{-1/2}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \,,\; \frac{\bar{Y}}{1 - n^{-1/2}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)} \right) \qquad (*)$$

is a confidence interval for θ of asymptotic coverage $1 - \alpha$.

(ii) Define the r^{th} *degree Hermite polynomial* $H_r(x)$.

Let $X_1, \ldots, X_n$ be independent, identically distributed random variables, with common mean μ and common variance σ^2, and let

$$T = \left(\sum_{i=1}^{n} X_i - n\mu \right) / \sqrt{n}\sigma.$$

An *Edgeworth expansion* of the distribution function of T is

$$P(T \leqslant t) = \Phi(t) - \phi(t)\left\{ \frac{\rho_3}{6\sqrt{n}} H_2(t) + \frac{\rho_4}{24n} H_3(t) + \frac{\rho_3^2}{72n} H_5(t) \right\} + O(n^{-3/2}).$$

in terms of standardised cumulants ρ_r.

Use an appropriate Edgeworth expansion to show that the confidence interval $(*)$ in (i) above has coverage error of order $O(n^{-1})$.

2 Explain briefly the concept of a *conditional likelihood*.

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

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

Find the cumulant generating function of $\tau_2(Y_i)$, and a saddlepoint approximation to the density of $S = n^{-1}\sum_{i=1}^{n} \tau_2(Y_i)$.

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

$$l(\psi, \hat{\lambda}_\psi) + \tfrac{1}{2}\log|d_{\lambda\lambda}(\psi, \hat{\lambda}_\psi)|,$$

in terms of quantities $\hat{\lambda}_\psi$, $d_{\lambda\lambda}$ which you should define carefully.

3 Describe in detail the p^* formula for the density of the maximum likelihood estimator $\hat{\theta}$ of a scalar parameter θ, and describe briefly its properties.

How would you estimate: (i) the density of the score function, (ii) the distribution function of $\hat{\theta}$, in the presence of an ancillary statistic?

4 Explain what is meant by a *functional statistic*, and the notion of *Fisher consistency*.

Define what is meant by the *influence function* of a functional T at a distribution F, and explain what is meant by an M-*estimator* of a parameter θ, based on a given ψ function. Find the form of the associated influence function and derive an expression for the asymptotic variance of the M-estimator at a distribution F.

Describe briefly, with reference to the location model on $\mathbb{R}$, the principles behind choice of an appropriate ψ function.

5 (i) Let $X_{(1)}, \ldots, X_{(n)}$ be an ordered, random sample from a continuous distribution function F.

Derive the distribution of $R = F(X_{(n)}) - F(X_{(1)})$.

Hence show that, for small ϵ, δ, the sample size n required to ensure that

$$P(R \geqslant 1 - \epsilon) \geqslant 1 - \delta$$

is approximately θ/ϵ, where θ is the unique positive solution to $1 + \theta - \delta e^{\theta} = 0$.

(ii) Describe in detail the Wilcoxon signed rank test, used to test whether a continuous distribution is symmetric about a point θ_0.

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

(i) Bartlett correction;

(ii) Modified profile likelihood;

(iii) Laplace approximation;

(iv) maximal invariants and equivariant estimators;

(v) parameter orthogonality;

(vi) one-sample U-statistics;

(vii) finite-sample versions of influence measures;

(viii) curved exponential families.

Section B

7 Suppose (r_{ij}) are independent observations, with

$$r_{ij} \sim Bi(n_{ij}, p_{ij}) \quad , \quad 1 \leqslant i, j \leqslant 2,$$

where $n_{11}, n_{12}, n_{21}, n_{22}$ are given totals. Consider the model

$$\omega : \operatorname{logit} p_{ij} = \mu + \alpha_i + \beta_j \quad , \quad 1 \leqslant i, j \leqslant 2$$

where $\alpha_1 = \beta_1 = 0$.

(i) Write down the log-likelihood under ω, and discuss carefully how $\hat{\alpha}_2, \hat{\beta}_2$ and their corresponding standard errors may be derived. [Do not attempt to find analytical expressions for $\hat{\alpha}_2, \hat{\beta}_2$ and their se's.]

(ii) With (r_{ij}/n_{ij}) as

498/796	878/1625
54/142	197/660

and fitting ω by glm( ), the deviance was found to be .00451, with

$$\hat{\alpha}_2 = -1.013(\text{se} = .0872)$$
$$\hat{\beta}_2 = -0.3544(\text{se} = .0804).$$

What do you conclude from these figures?

8 Let $Y_1, \ldots, Y_n$ be independent variables, such that $Y = X\beta + \epsilon$, where X is a given $n \times p$ matrix, of rank p, β is an unknown vector of dimension p, and $\epsilon_1, \ldots, \epsilon_n$ are independent normal variables, each with mean 0 and unknown variance σ^2.

(i) Derive an expression for $\hat{\beta}$, the least squares estimator of β, and derive the distribution of $\hat{\beta}$.

(ii) How would you estimate σ^2?

(iii) In fitting the model

$$Y_i = \mu + \alpha x_i + \beta z_i + \gamma t_i + \epsilon_i \quad , \quad 1 \leqslant i \leqslant n,$$

where $(x_i), (z_i), (t_i)$ are given vectors, and $\epsilon_1, \ldots, \epsilon_n$ has the distribution given above, explain carefully how you would test

$$H_0 : \beta = \gamma = 0.$$

(You may quote any standard theorems needed.)

9 Write an account, with appropriate examples, of the decision theory approach to inference. Your account should include discussion of *all* of the following:

(i) the main elements of a decision theory problem;

(ii) the Bayes and minimax principles;

(iii) admissibility;

(iv) finite decision problems;

(v) decision theory approaches to point estimation and hypothesis testing.

10 Suppose that Y_1 and Y_2 are independent Poisson random variables with means $\psi\mu$ and μ respectively. We are interested in testing the null hypothesis $H_0 : \psi \leqslant \psi_0$ against the alternative hypothesis $H_1 : \psi > \psi_0$, where ψ_0 is given and μ is unknown.

Explain in detail why the appropriate test is a conditional test, based on the conditional distribution of Y_1 given $Y_1 + Y_2$, and find its form.

Let $S = Y_1 + Y_2$. Show that the significance probability for observed (Y_1, Y_2) is approximately

$$1 - \Phi\left[\frac{Y_1 - S\psi_0/(1+\psi_0)}{\{S\psi_0/(1+\psi_0)^2\}^{1/2}}\right].$$