UNIVERSITY OF CAMBRIDGE

M. PHIL. IN STATISTICAL SCIENCE

Wednesday 6 June 2007 1.30 to 3.30

STATISTICAL THEORY

*Attempt **THREE** questions.*

*There are **FOUR** 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.

1 State and prove Cochran's theorem.

Explain how Cochran's theorem may be applied to the problem of testing hypotheses in a linear model.

[*You should define the model and the hypotheses carefully. You may assume that the likelihood ratio statistic is of the form*

$$w_{LR}(H_0) = n \log\Big(1 + \frac{A}{B}\Big),$$

but should state explicit expressions for A *and* B*. Formal verification of the conditions of Cochran's theorem is not required.*]

2 We write $Y \sim IG(\phi, \lambda)$ if the density of Y is

$$f(y; \phi, \lambda) = \frac{\lambda^{1/2}}{(2\pi)^{1/2} y^{3/2}} e^{(\lambda\phi)^{1/2}} \exp\Big\{-\frac{1}{2}\Big(\frac{\lambda}{y} + \phi y\Big)\Big\}, \quad y \in (0, \infty), \phi \in (0, \infty), \lambda \in (0, \infty).$$

Let $Y_1, \ldots, Y_n$ be independent $IG(\phi, \lambda)$ random variables. By first computing the cumulant generating function of $n^{-1}\sum_{i=1}^n Y_i$, find the density of $S_n = \sum_{i=1}^n Y_i$.

What is meant by a *saddlepoint approximation* to the density of a sum of independent and identically distributed random variables? [*An explicit expression for the* $O(n^{-1})$ *term is not required.*]

Compute the saddlepoint approximation to the density of S_n defined above. Comment on the accuracy of the approximation.

Now suppose $Y_1, \ldots, Y_n$ are independent with density

$$g(y) = \frac{5(\sqrt{5} - 1)}{4\pi(1 + y^{10})}, \quad y \in \mathbb{R}.$$

Without doing any calculations, explain briefly why it would not be appropriate to try to compute the saddlepoint approximation to the density of $\sum_{i=1}^n Y_i$ in this case.

3 Consider a model with two real-valued parameters, where one is of interest and the other is a nuisance parameter. What is meant by saying that the two parameters are *orthogonal*? What is meant by an *interest-respecting reparametrisation*? Give an informal derivation of a differential equation that may be used to find an orthogonal, interest-respecting reparametrisation.

Now suppose Y has density

$$f(y;\psi,\sigma)=\frac{1}{\sigma}\Big(1+\frac{\psi y}{\sigma}\Big)^{-(1+\frac{1}{\psi})},\quad y\in(0,\infty),\psi\in(0,\infty),\sigma\in(0,\infty),$$

where ψ is of interest and σ is a nuisance parameter. Show that ψ and σ are not orthogonal.

Find an orthogonal, interest-respecting reparametrisation.

[*You may assume without proof that if a is a non-negative integer, $b\in(0,\infty)$ satisfies $b-a>1$ and $c\in(0,\infty)$, then*

$$\int_0^\infty y^a(1+cy)^{-b}\,dy=\frac{a!}{c^{a+1}(b-1)(b-2)\ldots(b-a-1)}.$$
]

4 Let $Y_1,\ldots,Y_n$ be independent $N(\mu,\sigma^2)$ random variables, and suppose that we are interested in testing $H_0:\sigma^2=\sigma_0^2$ against $H_1:\sigma^2\neq\sigma_0^2$. Write down expressions for the maximum likelihood estimator $\hat{\sigma}^2$ of σ^2, and the constrained maximum likelihood estimator $\hat{\mu}_{\sigma^2}$ of μ for a fixed value of σ^2. Show that, under H_0, the likelihood ratio statistic may be written as

$$w_{LR}(\sigma_0^2)=n\{-\log(1+V)+V\},$$

where $V=(U-n)/n$ and $U\sim\chi^2_{n-1}$.

Define what is meant by the *Bartlett correction factor*, and the *Bartlett-corrected likelihood ratio statistic*. By integrating an asymptotic expansion term by term, which you may assume is valid, show that in the example above, the Bartlett correction factor is 11/6.

[*You may assume without proof that if $r\in\mathbb{N}$ and $U\sim\chi^2_{n-1}$, then*

$$\mathbb{E}(U^r)=\big(n-1+2(r-1)\big)\big(n-1+2(r-2)\big)\ldots(n-1).$$

You should bound the higher moments of V by using the fact that if $T_1,\ldots,T_{n-1}$ are independent and identically distributed with $\mathbb{E}(T_i)=0$ and $\mathbb{E}(|T_i|^r)<\infty$ for some $r\in\mathbb{N}$ with $r\geqslant 2$, then there exists a finite constant $C(r)$ such that

$$\left|\mathbb{E}\Big\{\Big(\sum_{i=1}^{n-1}T_i\Big)^r\Big\}\right|\leqslant\begin{cases}C(r)n^{r/2} & \text{if } r \text{ is even}\\ C(r)n^{(r-1)/2} & \text{if } r \text{ is odd.}\end{cases}$$
]

END OF PAPER