## A1/12 B1/15 Principles of Statistics

(i) Explain in detail the minimax and Bayes principles of decision theory.

Show that if $d(X)$ is a Bayes decision rule for a prior density $\pi(\theta)$ and has constant risk function, then $d(X)$ is minimax.
(ii) Let $X_{1}, \ldots, X_{p}$ be independent random variables, with $X_{i} \sim N\left(\mu_{i}, 1\right), i=1, \ldots, p$. Consider estimating $\mu=\left(\mu_{1}, \ldots, \mu_{p}\right)^{T}$ by $d=\left(d_{1}, \ldots, d_{p}\right)^{T}$, with loss function

$$
L(\mu, d)=\sum_{i=1}^{p}\left(\mu_{i}-d_{i}\right)^{2}
$$

What is the risk function of $X=\left(X_{1}, \ldots, X_{p}\right)^{T}$ ?
Consider the class of estimators of $\mu$ of the form

$$
d^{a}(X)=\left(1-\frac{a}{X^{T} X}\right) X,
$$

indexed by $a \geqslant 0$. Find the risk function of $d^{a}(X)$ in terms of $E\left(1 / X^{T} X\right)$, which you should not attempt to evaluate, and deduce that $X$ is inadmissible. What is the optimal value of $a$ ?
[You may assume Stein's Lemma, that for suitably behaved real-valued functions $h$,

$$
E\left\{\left(X_{i}-\mu_{i}\right) h(X)\right\}=E\left\{\frac{\partial h(X)}{\partial X_{i}}\right\}
$$

## A2/11 $\quad$ B2/16 Principles of Statistics

(i) Let $X$ be a random variable with density function $f(x ; \theta)$. Consider testing the simple null hypothesis $H_{0}: \theta=\theta_{0}$ against the simple alternative hypothesis $H_{1}: \theta=\theta_{1}$.

What is the form of the optimal size $\alpha$ classical hypothesis test?
Compare the form of the test with the Bayesian test based on the Bayes factor, and with the Bayes decision rule under the 0-1 loss function, under which a loss of 1 is incurred for an incorrect decision and a loss of 0 is incurred for a correct decision.
(ii) What does it mean to say that a family of densities $\{f(x ; \theta), \theta \in \Theta\}$ with real scalar parameter $\theta$ is of monotone likelihood ratio?

Suppose $X$ has a distribution from a family which is of monotone likelihood ratio with respect to a statistic $t(X)$ and that it is required to test $H_{0}: \theta \leqslant \theta_{0}$ against $H_{1}: \theta>\theta_{0}$.

State, without proof, a theorem which establishes the existence of a uniformly most powerful test and describe in detail the form of the test.

Let $X_{1}, \ldots, X_{n}$ be independent, identically distributed $U(0, \theta), \theta>0$. Find a uniformly most powerful size $\alpha$ test of $H_{0}: \theta \leqslant \theta_{0}$ against $H_{1}: \theta>\theta_{0}$, and find its power function. Show that we may construct a different, randomised, size $\alpha$ test with the same power function for $\theta \geqslant \theta_{0}$.

## A3/12 B3/15 Principles of Statistics

(i) Describe in detail how to perform the Wald, score and likelihood ratio tests of a simple null hypothesis $H_{0}: \theta=\theta_{0}$ given a random sample $X_{1}, \ldots, X_{n}$ from a regular oneparameter density $f(x ; \theta)$. In each case you should specify the asymptotic null distribution of the test statistic.
(ii) Let $X_{1}, \ldots, X_{n}$ be an independent, identically distributed sample from a distribution $F$, and let $\hat{\theta}\left(X_{1}, \ldots, X_{n}\right)$ be an estimator of a parameter $\theta$ of $F$.

Explain what is meant by: (a) the empirical distribution function of the sample; (b) the bootstrap estimator of the bias of $\hat{\theta}$, based on the empirical distribution function. Explain how a bootstrap estimator of the distribution function of $\hat{\theta}-\theta$ may be used to construct an approximate $1-\alpha$ confidence interval for $\theta$.

Suppose the parameter of interest is $\theta=\mu^{2}$, where $\mu$ is the mean of $F$, and the estimator is $\hat{\theta}=\bar{X}^{2}$, where $\bar{X}=n^{-1} \sum_{i=1}^{n} X_{i}$ is the sample mean.

Derive an explicit expression for the bootstrap estimator of the bias of $\hat{\theta}$ and show that it is biased as an estimator of the true bias of $\hat{\theta}$.

Let $\hat{\theta}_{i}$ be the value of the estimator $\hat{\theta}\left(X_{1}, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{n}\right)$ computed from the sample of size $n-1$ obtained by deleting $X_{i}$ and let $\hat{\theta} .=n^{-1} \sum_{i=1}^{n} \hat{\theta}_{i}$. The jackknife estimator of the bias of $\hat{\theta}$ is

$$
b_{J}=(n-1)(\hat{\theta} .-\hat{\theta}) .
$$

Derive the jackknife estimator $b_{J}$ for the case $\hat{\theta}=\bar{X}^{2}$, and show that, as an estimator of the true bias of $\hat{\theta}$, it is unbiased.

## A4/13 B4/15 Principles of Statistics

(a) Let $X_{1}, \ldots, X_{n}$ be independent, identically distributed random variables from a one-parameter distribution with density function

$$
f(x ; \theta)=h(x) g(\theta) \exp \{\theta t(x)\}, x \in \mathbb{R} .
$$

Explain in detail how you would test

$$
H_{0}: \theta=\theta_{0} \text { against } H_{1}: \theta \neq \theta_{0}
$$

What is the general form of a conjugate prior density for $\theta$ in a Bayesian analysis of this distribution?
(b) Let $Y_{1}, Y_{2}$ be independent Poisson random variables, with means $(1-\psi) \lambda$ and $\psi \lambda$ respectively, with $\lambda$ known.

Explain why the Conditionality Principle leads to inference about $\psi$ being drawn from the conditional distribution of $Y_{2}$, given $Y_{1}+Y_{2}$. What is this conditional distribution?
(c) Suppose $Y_{1}, Y_{2}$ have distributions as in (b), but that $\lambda$ is now unknown.

Explain in detail how you would test $H_{0}: \psi=\psi_{0}$ against $H_{1}: \psi \neq \psi_{0}$, and describe the optimality properties of your test.
[Any general results you use should be stated clearly, but need not be proved.]

Part II 2002

