1/II/27J **Principles of Statistics**

(a) What is a *loss function*? What is a *decision rule*? What is the *risk function* of a decision rule? What is the *Bayes risk* of a decision rule with respect to a prior π?

(b) Let $\theta \mapsto R(\theta, d)$ denote the risk function of decision rule d, and let $r(\pi, d)$ denote the Bayes risk of decision rule d with respect to prior π. Suppose that d^* is a decision rule and π_0 is a prior over the parameter space Θ with the two properties

(i) $r(\pi_0, d^*) = \min_d r(\pi_0, d)$

(ii) $\sup_\theta R(\theta, d^*) = r(\pi_0, d^*)$.

Prove that d^* is minimax.

(c) Suppose now that $\Theta = \mathcal{A} = \mathbb{R}$, where $\mathcal{A}$ is the space of possible actions, and that the loss function is

$$L(\theta, a) = \exp(-\lambda a \theta),$$

where λ is a positive constant. If the law of the observation X given parameter θ is $N(\theta, \sigma^2)$, where $\sigma > 0$ is known, show (using (b) or otherwise) that the rule

$$d^*(x) = x/\sigma^2\lambda$$

is minimax.

2/II/27J **Principles of Statistics**

Let $\{f(\cdot|\theta) : \theta \in \Theta\}$ be a parametric family of densities for observation X. What does it mean to say that the statistic $T \equiv T(X)$ is *sufficient* for θ? What does it mean to say that T is *minimal sufficient*?

State the Rao–Blackwell theorem. State the Cramér–Rao lower bound for the variance of an unbiased estimator of a (scalar) parameter, taking care to specify any assumptions needed.

Let $X_1, \ldots, X_n$ be a sample from a $U(0, \theta)$ distribution, where the positive parameter θ is unknown. Find a minimal sufficient statistic T for θ. If $h(T)$ is an unbiased estimator for θ, find the form of h, and deduce that this estimator is minimum-variance unbiased. Would it be possible to reach this conclusion using the Cramér–Rao lower bound?

3/II/26J **Principles of Statistics**

Write an essay on the rôle of the Metropolis–Hastings algorithm in computational Bayesian inference on a parametric model. You may for simplicity assume that the parameter space is finite. Your essay should:

(a) explain what problem in Bayesian inference the Metropolis–Hastings algorithm is used to tackle;

(b) fully justify that the algorithm does indeed deliver the required information about the model;

(c) discuss any implementational issues that need care.

4/II/27J **Principles of Statistics**

(a) State the strong law of large numbers. State the central limit theorem.

(b) Assuming whatever regularity conditions you require, show that if $\hat{\theta}_n \equiv \hat{\theta}_n(X_1, \ldots, X_n)$ is the maximum-likelihood estimator of the unknown parameter θ based on an independent identically distributed sample of size n, then under P_θ

$$\sqrt{n}(\hat{\theta}_n - \theta) \to N(0, J(\theta)^{-1}) \qquad \text{in distribution}$$

as $n \to \infty$, where $J(\theta)$ is a matrix which you should identify. A rigorous derivation is not required.

(c) Suppose that $X_1, X_2, \ldots$ are independent binomial $\mathrm{Bin}(1, \theta)$ random variables. It is required to test $H_0 : \theta = \frac{1}{2}$ against the alternative $H_1 : \theta \in (0, 1)$. Show that the construction of a likelihood-ratio test leads us to the statistic

$$T_n = 2n\{\hat{\theta}_n \log \hat{\theta}_n + (1 - \hat{\theta}_n) \log(1 - \hat{\theta}_n) + \log 2\},$$

where $\hat{\theta}_n \equiv n^{-1} \sum_{k=1}^{n} X_k$. Stating clearly any result to which you appeal, for large n, what approximately is the distribution of T_n under H_0? Writing $\hat{\theta}_n = \frac{1}{2} + Z_n$, and assuming that Z_n is small, show that

$$T_n \simeq 4nZ_n^2.$$

Using this and the central limit theorem, briefly justify the approximate distribution of T_n given by asymptotic maximum-likelihood theory. What could you say if the assumption that Z_n is small failed?