UNIVERSITY OF CAMBRIDGE

# MATHEMATICAL TRIPOS Part III

Monday, 2 June, 2014 9:00 am to 11:00 am

## PAPER 30

## STATISTICAL THEORY

*Attempt no more than **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**

In a parametric statistical model $\{f(\cdot,\theta) : \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^p$, define the maximum likelihood estimator and the Fisher information.

Formulate and prove a result on the asymptotic distribution of the maximum likelihood estimator. In the proof you may assume consistency of the maximum likelihood estimator, as well as standard regularity conditions of the model without specifying them.

**2**

In a standard linear model $Y = X\theta + \varepsilon$ with $X$ a nonstochastic $n \times p$ design matrix of full column rank, noise vector $\varepsilon \sim N(0, \sigma^2 I)$ where $\sigma^2 > 0$ is unknown, and with parameter vector $\theta = (\theta_1, \ldots, \theta_p)^T \in \Theta = \mathbb{R}^p$, $p \leqslant n$, consider a $k$-dimensional sub-model $M$ of $\mathbb{R}^p$, $k \leqslant p$, consisting of any $k$ components of $\theta$. Define the least squares estimator $\hat{\theta}^M$ of the submodel $M$ and construct an unbiased estimator for its prediction risk

$$R(M) \equiv E\|X\hat{\theta}^M - X\theta\|^2.$$

That is, find a data-driven statistic $\widehat{R(M)}$ that satisfies

$$E[\widehat{R(M)}] = E\|X\hat{\theta}^M - X\theta\|^2, \quad \forall\theta \in \mathbb{R}^p.$$

[Here $\|\cdot\|$ denotes the standard Euclidean norm.]

Discuss heuristically how you would use this estimator to construct a statistical model selection procedure.

**3**

Let $X$ be a $n \times p$ matrix. Define the concepts of a *Gram matrix* $\hat{\Sigma}$ and of the *restricted isometry property.*

Let $g_i$, $i = 1, \ldots, n$, be i.i.d. $N(0,1)$ and set $Z = \sum_{i=1}^{n}(g_i^2 - 1)$. Show that for all $t \geqslant 0$ and every $n \in \mathbb{N}$,

$$\Pr(Z \geqslant t) \leqslant 2\exp\left\{-\frac{t^2}{4(n+t)}\right\}; \tag{1}$$

and moreover that for every $z \geqslant 0$ and every $n \in \mathbb{N}$,

$$\Pr(Z \geqslant 4(\sqrt{nz} + z)) \leqslant 2e^{-z}. \tag{2}$$

Now let the matrix $X$ be formed of i.i.d. entries $X_{ij} \sim N(0,1)$, and let $\hat{\Sigma}$ be the associated Gram matrix. Formulate and prove a result about the concentration of $\theta^T\hat{\Sigma}\theta$ around $\|\theta\|_2^2$ for every *fixed* $\theta \in \mathbb{R}^p$ satisfying $\|\theta\|_2 \leqslant 1$.

**4**

Suppose $\Theta \subset \mathbb{R}^p$ is compact and assume that $Q : \Theta \to \mathbb{R}$ is a non-random function that is continuous on $\Theta$, and that $\theta_0$ is the unique minimizer of $Q$. If

$$\sup_{\theta \in \Theta} |Q_n(\theta; Y_1, \ldots, Y_n) - Q(\theta)| \to^P 0 \tag{1}$$

as $n \to \infty$, show that any solution $\hat{\theta}_n$ of

$$\min_{\theta \in \Theta} Q_n(\theta, Y_1, \ldots, Y_n)$$

converges to $\theta_0$ in probability as $n \to \infty$.

Let now $\Theta \subseteq \mathbb{R}$ and let $S_n$ be a sequence of random real-valued continuous functions defined on $\Theta$ such that, as $n \to \infty$, $S_n(\theta)$ converges to $S(\theta)$ in probability $\forall \theta \in \Theta$, where $S : \Theta \to \mathbb{R}$ is nonrandom. Suppose for some $\theta_0 \in \Theta$ and every $\varepsilon > 0$ we have $S(\theta_0 - \varepsilon) < 0 < S(\theta_0 + \varepsilon)$, and that $S_n$ has *exactly one* zero $\hat{\theta}_n$ for every $n \in \mathbb{N}$. Deduce that $\hat{\theta}_n \to^P \theta_0$ as $n \to \infty$.

## END OF PAPER