UNIVERSITY OF CAMBRIDGE

MAMA/210, NST3AS/210, MAAS/210

MAT3

MATHEMATICAL TRIPOS Part III

Friday 13 June 2025 1:30 pm to 3:30 pm

PAPER 210

TOPICS IN STATISTICAL THEORY

Before you begin please read these instructions carefully

Candidates have TWO HOURS to complete the written examination.

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
Rough 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 *(a)* Let $X, X_1, \ldots, X_n$ be independent and identically distributed on a measurable space $(\mathcal{X}, \mathcal{A})$ and let $\mathcal{G}$ denote a collection of integrable, real-valued functions on $\mathcal{X}$. What does it mean for $\mathcal{G}$ to satisfy a *Uniform Law of Large Numbers (ULLN)*? [You may assume that a relevant supremum is measurable.]

Given $\epsilon > 0$, let $N(\epsilon, \mathcal{G}) \in \mathbb{N} \cup \{\infty\}$ denote the minimal $N \in \mathbb{N} \cup \{\infty\}$ for which there exist integrable functions $g_1^{\mathrm{L}}, g_1^{\mathrm{U}}, \ldots, g_N^{\mathrm{L}}, g_N^{\mathrm{U}} : \mathcal{X} \to \mathbb{R}$ with the properties that $\mathbb{E}\big|g_j^{\mathrm{U}}(X) - g_j^{\mathrm{L}}(X)\big| \leqslant \epsilon$ for every $j \in [N]$, and for every $g \in \mathcal{G}$, there exists $j^* \equiv j^*(g) \in [N]$ such that $g_{j^*}^{\mathrm{L}}(x) \leqslant g(x) \leqslant g_{j^*}^{\mathrm{U}}(x)$ for all $x \in \mathcal{X}$. Prove that if $N(\epsilon, \mathcal{G}) < \infty$ for every $\epsilon > 0$, then $\mathcal{G}$ satisfies a ULLN.

(b) State the Glivenko–Cantelli theorem, and deduce it from *(a)*.

(c) Now suppose that the distribution function of X is continuous on $\mathbb{R}$. Let $\mathcal{B}(\mathbb{R})$ denote the Borel σ-algebra on $\mathbb{R}$ and for $A \in \mathcal{B}(\mathbb{R})$, define $g_A : \mathbb{R} \to \mathbb{R}$ by $g_A(x) := \mathbb{1}_{\{x \in A\}}$. Does $\mathcal{G} := \{g_A : A \in \mathcal{B}(\mathbb{R})\}$ satisfy a ULLN? Justify your answer.

2 In the context of kernel density estimation, define what is meant by a *kernel* K and the *scaled kernel* K_h, where $h > 0$. Given a bounded, Borel measurable function $g_1 : \mathbb{R} \to \mathbb{R}$ and an integrable function $g_2 : \mathbb{R} \to \mathbb{R}$, define their *convolution* $g_1 * g_2$.

Let f be a density on $\mathbb{R}$ with $f(x) = 0$ for $x \in (-\infty, 0]$, and suppose that $\mu := \int_0^\infty x f(x)\,dx < \infty$ and $\bar{\mu} := \int_0^\infty f(x)/x\,dx < \infty$. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} g$, where g is the density given by $g(x) := x f(x)/\mu$ for $x \in \mathbb{R}$, and suppose that our goal is to estimate f. Suppose further that μ is known, and consider the estimator $\hat{f}_n$ given by

$$\hat{f}_n(x) := \frac{\mu}{nh} \sum_{i=1}^n \frac{1}{X_i} K\biggl(\frac{x - X_i}{h}\biggr),$$

where K is a bounded kernel and $h > 0$. Find an exact expression, involving a convolution, for the bias of $\hat{f}_n(x)$. Compare this expression with the bias of a kernel density estimator based on observations having density f.

Show that

$$\int_{-\infty}^{\infty} \operatorname{Var} \hat{f}_n(x)\,dx = \frac{\mu\bar{\mu}R(K)}{nh} - \frac{1}{n}\int_{-\infty}^{\infty} (K_h * f)^2(x)\,dx \tag{1}$$

for some $R(K)$ that you should define. Prove further that $\mu\bar{\mu} > 1$ and hence compare (1) with the corresponding integrated variance of a kernel density estimator based on independent observations having density f.

3 For $\beta, L > 0$, define the *Hölder class* $\mathcal{H}(\beta, L)$ on $\mathbb{R}$.

Consider a vector $Y = (Y_1, \ldots, Y_n)^\top$ of responses generated via

$$Y_i = m(x_i) + \epsilon_i,$$

where $x_i = i/n$ for $i \in [n]$, where $m \in \mathcal{H}(\beta, L)$ and where $\epsilon_1, \ldots, \epsilon_n$ are independent with $\mathbb{E}(\epsilon_i) = 0$ and $\mathrm{Var}(\epsilon_i) \leqslant \sigma^2$ for $i \in [n]$. Fix $x \in [0, 1]$, define $K : \mathbb{R} \to [0, 1/2]$ by $K(u) := \mathbb{1}_{\{|u| \leqslant 1\}}/2$, let $p \in \mathbb{N}_0$ and let $h > 0$. Give the definition of the *local polynomial estimator* $\hat{m}_n(x; p, h)$ *of* $m(x)$ *of degree* p*, bandwidth* h *and kernel* K. Show that, for suitable matrices $X \in \mathbb{R}^{n \times (p+1)}$ and $W \in \mathbb{R}^{n \times n}$, and subject to a positive definiteness condition that you should state, $\hat{m}_n(x; p, h)$ can be expressed as an appropriate component of $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)^\top$, where $\hat{\beta} := (X^\top W X)^{-1} X^\top W Y$.

Define the *effective kernel* $\{w_i(x) : i \in [n]\}$ of $\hat{m}_n(x; p, h)$. Under the same positive definiteness condition as above, prove that if R is a polynomial of degree at most p, then

$$\frac{1}{n} \sum_{i=1}^{n} w_i(x) R(x_i) = R(x).$$

Writing $Q(u) := (1, u, u^2/2, \ldots, u^p/p!)^\top \in \mathbb{R}^{p+1}$ for $u \in \mathbb{R}$, explain why the minimum eigenvalue Λ_p of the matrix $\int_0^1 Q(u) Q(u)^\top \, du \in \mathbb{R}^{(p+1) \times (p+1)}$ is positive. Using the fact that the minimum eigenvalue $\lambda_0 \equiv \lambda_{0,n,h,x}(p)$ of $n^{-1} X^\top W X$ satisfies

$$\inf_{x \in [0,1]} \lambda_{0,n,h,x}(p) \geqslant \frac{1}{2} \left(\Lambda_p - \frac{6e}{nh} \right),$$

when $n \geqslant 2$ and $h \leqslant 1/4$, prove that for $p \geqslant \lceil \beta \rceil - 1$ there exist $n_0 \equiv n_0(p) \in \mathbb{N}$, $a \equiv a(p) > 0$ and $C \equiv C(p) > 0$ such that

$$\sup_{x \in [0,1]} \sup_{m \in \mathcal{H}(\beta, L)} \mathbb{E}\big[\big\{\hat{m}_n(x; p, h) - m(x)\big\}^2\big] \leqslant C \left(\frac{\sigma^2}{nh} + L^2 h^{2\beta} \right)$$

for $n \geqslant n_0$ and $h \in [a/n, 1/4]$.

4 Let P and Q be probability measures on a σ-finite measure space $(\mathcal{X}, \mathcal{A}, \mu)$, having Radon–Nikodym derivatives p and q respectively with respect to μ. Write down expressions for the *total variation distance* $\mathrm{TV}(P, Q)$ and *Hellinger distance* $\mathrm{H}(P, Q)$ as integrals with respect to μ. Prove that

$$\mathrm{TV}^2(P, Q) \leqslant \mathrm{H}^2(P, Q) \leqslant 2\mathrm{TV}(P, Q).$$

Writing P^n and Q^n for the n-fold product distributions of P and Q respectively, and using the fact that these probability measures have Radon–Nikodym derivatives p^n and q^n respectively with respect to the σ-finite product measure μ^n, prove that

$$\mathrm{H}^2(P^n, Q^n) = 2 - 2\left(1 - \frac{1}{2}\mathrm{H}^2(P, Q)\right)^n.$$

State Le Cam's two-point lemma.

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} U[\theta, \theta + 1]$ for some $\theta \in \mathbb{R}$. Prove that there exists a universal constant $c > 0$ such that

$$\inf_{\hat{\theta} \in \hat{\Theta}} \sup_{\theta \in \mathbb{R}} \mathbb{E}_\theta \bigl\{ \bigl(\hat{\theta}(X_1, \ldots, X_n) - \theta\bigr)^2 \bigr\} \geqslant \frac{c}{n^2}$$

for all $n \in \mathbb{N}$, where $\hat{\Theta}$ denotes the set of Borel measurable functions from $\mathbb{R}^n$ to $\mathbb{R}$. [You may use the fact that $(1 - x)^a \geqslant 1 - ax$ for $x \in [0, 1]$ and $a \geqslant 1$.]

END OF PAPER