UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Friday, 7 June, 2013 9:00 am to 11:00 am

PAPER 33

NONPARAMETRIC 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

Let $X, X_1, \dots, X_n$ be independent and identically distributed random variables, taking values in a measurable space T. Let $\mathcal{H}$ be a class of measurable functions $T \to \mathbb{R}$, which satisfies $\mathbb{E}|h(X)| < \infty$ for all $h \in \mathcal{H}$.

State and prove a uniform law of large numbers for the random variables $\frac{1}{n}\sum_{i=1}^n h(X_i)$, $h \in \mathcal{H}$, assuming a condition on the existence of brackets which you should specify.

Now let $T = [0, \infty)$, $m(s) = \mathbb{E}[\exp(sX)]$, and suppose that $m(s) < \infty$ for $s \in [0, t]$. Define an estimator $\hat{m}_n(s)$ of $m(s)$ which depends only on $X_1, \dots, X_n$, and prove it satisfies

$$\sup_{s \in [0,t]} |\hat{m}_n(s) - m(s)| \overset{a.s.}{\to} 0.$$

2

Let $X_1, \dots, X_n$ be independent and identically-distributed real-valued random variables with density f, and $K = 1_{[-1/2,1/2]}$. For $K_h(x) = h^{-1}K(x/h)$, $h > 0$, define the convolution $K_h * f$, and the kernel density estimator $\hat{f}^K_{n,h}$. Show that

$$\mathbb{E}\|\hat{f}^K_{n,h} - K_h * f\|_2 \leqslant \kappa\sqrt{1/nh},$$

where $\|\cdot\|_2$ denotes the L^2 norm on $\mathbb{R}$, and κ is a constant you should specify.

Now let

$$x_0 < x_1 < \cdots < x_k,$$

and suppose f is constant on the intervals $(-\infty, x_0), [x_0, x_1), \dots, [x_k, \infty)$. Show that, for a suitable choice of h,

$$\mathbb{E}\|\hat{f}^K_{n,h} - f\|_2 = O(n^{-1/4}).$$

3

Define the Haar scaling functions $\{\varphi_{j,k} : k \in \mathbb{Z}\}$, for $j \in \{0, 1, \dots\}$, and the Haar basis $\{\varphi_{0,k}, \psi_{l,k} : l \in \{0, 1, \dots\}, k \in \mathbb{Z}\}$. Prove that these sets are orthonormal sets in $L^2(\mathbb{R})$.

Given a locally-integrable function $f : \mathbb{R} \to \mathbb{R}$, define the Haar approximation $H_j(f)$ to f, both in terms of Haar scaling functions, and in terms of Haar basis functions. State why these two definitions are equivalent.

Now let $f : \mathbb{R} \to \mathbb{R}$ be symmetric, decreasing on $[0, \infty)$, and satisfy $f(x) \to 0$ as $x \to \infty$. Show that

$$\|H_j(f) - f\|_1 \leqslant f(0)2^{1-j}.$$

4

Let m and V be functions $[0,1] \to \mathbb{R}$, and $x_i = i/n$. Also let $Y_1, \ldots, Y_n$ be independent real-valued random variables, with Y_i having mean $m(x_i)$, and variance $V(x_i)$.

Let the kernel $K = 1_{[-1/2,1/2]}$, and define the corresponding local polynomial estimator $\hat{m}^P_{n,h}$ of order zero. Show that if $n \geqslant 2h^{-1}$, this estimator equals the Nadaraya-Watson estimator $\hat{m}^K_{n,h}$.

Now suppose that m is differentiable, with $m' \in L^\infty$, and V is bounded by $\sigma^2 > 0$. Prove that if $n \geqslant 2h^{-1}$, and $x \in [0,1]$,

$$\mathbb{E}|\hat{m}^P_{n,h}(x) - m(x)| \leqslant \kappa \left(\frac{\sigma}{\sqrt{nh}} + \|m'\|_\infty h \right),$$

for a constant $\kappa > 0$.

END OF PAPER