UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Wednesday, 8 June, 2022 1:30 pm to 3:30 pm

PAPER 224

INFORMATION 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*) State Stein's lemma and give its proof.

(*b*) Define the total variation distance between two PMFs P, Q on a discrete alphabet A, and show that, alternatively, it can be expressed as:

$$\|P - Q\|_{\mathrm{TV}} = 2 \sup_{B \subset A} |P(B) - Q(B)|.$$

(*c*) Consider a hypothesis test between two arbitrary PMFs P_n and Q_n on A^n, for some discrete alphabet A. Denote the probabilities of error associated with an arbitrary decision region $B_n \subset A^n$ by $e_1^{(n)}(B_n)$ and $e_2^{(n)}(B_n)$. Let $P_e^{(n)}(B_n) = e_1^{(n)}(B_n) + e_2^{(n)}(B_n)$ denote the "total" error probability. Show that the smallest possible total error probability that can be achieved by any hypothesis test is equal to:

$$1 - \frac{1}{2}\|P_n - Q_n\|_{\mathrm{TV}}.$$

2

(*a*) State the theorem on error exponents in fixed-rate data compression and prove its direct part.

(*b*) Consider a simple-versus-simple hypothesis test between two distributions P and $Q \neq P$ having full support on a finite alphabet A. For $n \geqslant 1$ and any $\delta > 0$ define the decision regions $B_n^* = \{x_1^n : \text{the type } \hat{P}_n \text{ of } x_1^n \text{ satisfies } D(\hat{P}_n \| P) \leqslant \delta\}$. Show that the associated probabilities of error $e_1^{(n)}$ and $e_2^{(n)}$ satisfy, for $i = 1, 2$,

$$\limsup_{n \to \infty} \frac{1}{n} \log e_i \leqslant -D_i(\delta),$$

for appropriate exponents $D_1(\delta), D_2(\delta)$. Identify these exponents and find an interval of values of δ where they are both strictly positive.

3

(a) Define the *type* of a string x_1^n with values in the finite alphabet A. Given a probability mass function (PMF) Q on A, derive an expression for the probability $Q^n(x_1^n)$ in terms of its type.

(b) Let P be an n-type on a finite alphabet A of size $m = |A|$. Define the *type class* $T(P)$ and show that:

$$|T(P)| \geqslant (n+1)^{-m} 2^{nH(P)}.$$

If you give the proof that was given in class, you should also carefully state and prove any intermediate lemmas used in the proof.

(c) Let B be an arbitrary nonempty subset of A^n, write $\hat{P}_{x_1^n}$ for the type of a string $x_1^n \in A^n$, and define the PMF:

$$P_B(a) = \frac{1}{|B|} \sum_{x_1^n \in B} \hat{P}_{x_1^n}(a), \qquad a \in A.$$

Show that:

$$|B| \leqslant 2^{nH(P_B)}.$$

Justify all the steps in your argument carefully. *Hint.* Consider X_1^n to be a random string drawn uniformly among all elements of B, and let J be a uniformly drawn index from $\{1, 2, \ldots, n\}$, independent of X_1^n.

4

(a) State Kraft's inequality and prove its converse part.

(b) Suppose that the PMF P of a RV X values in $A = \{1, 2, \ldots\}$ has nonincreasing probabilities, i.e., $P(k+1) \leqslant P(k)$ for all k. Show that, if $H(X) < \infty$, then $\mathbb{E}[\log X] < \infty$.

(c) Show that the geometric distribution with parameter $p = \frac{1}{\mu}$ has maximal entropy among all distributions on $A = \{1, 2, \ldots\}$ with mean $\mu > 0$.

(d) Let A be a finite alphabet, let $f : A \to \mathbb{R}$ be a given function, and let v be a constant such that $\min_{x \in A} f(x) < v < \max_{x \in A} f(x)$. Identify the PMF on A that has maximal entropy among all PMFs P on A with $\sum_{x \in A} P(x) f(x) = v$.

[*Hint.* Recall the form of the minimising P^* in Sanov's theorem in the context of the Chernoff bound.]

END OF PAPER