UNIVERSITY OF CAMBRIDGE

MAMA/224, NST3AS/224, MAAS/224

MAT3

MATHEMATICAL TRIPOS Part III

Wednesday, 7 June, 2023 9:00 am to 11:00 am

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) Suppose X, Y are independent RVs with values in $A = \{0, 1, \ldots\}$. Show that $H(X+Y) \geqslant \max\{H(X), H(Y)\}$.

(b) Suppose X has PMF P_X and mean $\mu > 0$ on A and let Z be an independent geometric with mean μ, i.e., with parameter $p = \frac{1}{1+\mu}$ and PMF $P_Z(k) = p(1-p)^k$, $k \geqslant 0$. Using the result of part (a) show that,

$$D(P_X \| P_Z) \leqslant 2 d_R(X, Z),$$

where $d_R(X, Z)$ is the *Ruzsa distance* between X and Z:

$$d_R(X, Z) := H(X - Z) - \frac{1}{2} H(X) - \frac{1}{2} H(Z).$$

(c) Suppose X has PMF P_X on A with $P_X(0) = 0$, let $q := \sum_{k=0}^{\infty} \min\{P_X(k), P_X(k+1)\}$, and write P_{X-1} for the PMF of $X - 1$ on A. Show that,

$$\|P_X - P_{X-1}\|_{\text{TV}} = 2(1 - q).$$

(d) With X as in part (c), use the result of part (c) to show that:

$$2^{-2H(X)+1} \leqslant (\log_e 2) D(P_X \| P_{X-1}).$$

Hint. Find an upper bound for q and a lower bound for $H(X)$, both in terms of $\max_k P_X(k)$.

2 Let A be a finite alphabet and Q be a PMF on A.

(a) Let x_1^n be a string in A^n. Define its type $P = \hat{P}_{x_1^n}$, define the type class $T(P)$, and state an upper bound for the probability $Q^n(T(P))$.

Let B be an arbitrary nonempty subset of A^n. In the next five parts you will establish a corresponding upper bound for $Q^n(B)$.

(b) Suppose that $Y_1^n = (Y_1, Y_2, \ldots, Y_n)$ are independent and identically distributed with each $Y_i \sim Q$, and that X_1^n are distributed as Y_1^n conditional on $Y_1^n \in B$. Write down a formula for the joint PMF P_n of X_1^n.

(c) Let J be uniformly distributed on $\{1, 2, \ldots, n\}$, independent of Y_1^n. Derive a formula for the PMF $\bar{P}$ of X_J in terms of the types $\hat{P}_{x_1^n}$ of the strings $x_1^n \in B$.

(d) Show that $H(X_1^n) \leqslant nH(\bar{P})$.

(e) Show that:

$$\sum_{x_1^n \in B} \frac{Q^n(x_1^n)}{Q^n(B)} \log Q^n(x_1^n) = n \sum_{x \in A} \bar{P}(x) \log Q(x).$$

(f) Using parts (b), (d) and (e) show that: $Q^n(B) \leqslant 2^{-nD(\bar{P}\|Q)}$.

3

(a) State Kraft's inequality for prefix-free free codes (C_n, L_n) on A^n for a finite alphabet A.

(b) State and prove both the direct and converse parts of the codes-distributions correspondence.

(c) Let $X_1^n = (X_1, \ldots, X_n)$ be random variables with values in the finite alphabet A, and let $W(x_1^n)$ denote the "weight" of a string $x_1^n \in A^n$ for some fixed weight function $W : A^n \to (0, \infty)$. Find the smallest achievable value of the average weighted description length, $\mathbb{E}[W(X_1^n)L_n(X_1^n)]$, among all prefix-free codes, ignoring integer codelength constraints. Describe the length function L_n^* that achieves that minimum.

4

(*a*) Suppose $X_1, X_2, \ldots, X_n$ are (not necessarily independent) Bernoulli random variables (RVs), and let $S_n = X_1 + \cdots + X_n$. State and prove a Poisson approximation bound for the PMF P_{S_n} of S_n. You may assume, without proof, that $D_e(\text{Bern}(q)\|\text{Po}(q)) \leqslant q^2$.

(*b*) Let $\{X_i^{(n)}\} = \{(X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)})\ ;\ n \geqslant 1\}$ be a triangular array of independent Bernoulli RVs, where, for each row $n \geqslant 1$, the RVs $(X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)})$ are IID $\text{Bern}(\lambda/n)$ for some fixed $\lambda > 0$ independent of n.

Let P_n denote the joint PMF of $(X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)})$, $n \geqslant 1$. Find a sequence of constants $\{c_n\}$ and a RV Z such that the following version of the asymptotic equipartition property holds in this case: As $n \to \infty$:

$$-\frac{1}{c_n} \log P_n(X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}) \to Z \quad \text{in distribution.}$$

You may assume the following without proof: If $\{Z_n\}$ and Z are RVs and $\{a_n\}$ and $\{b_n\}$ are sequences of real numbers such that, as $n \to \infty$, (i) the PMFs P_{Z_n} of Z_n converge to the PMF P_Z of Z in that $\|P_{Z_n} - P_Z\|_{\text{TV}} \to 0$, (ii) $a_n \to 1$, and (iii) $b_n \to 0$, then $a_n Z_n + b_n \to Z$ in distribution, as $n \to \infty$.

END OF PAPER