UNIVERSITY OF CAMBRIDGE

MAMA/208, NST3AS/208, MAAS/208

MAT3

MATHEMATICAL TRIPOS Part III

Friday 12 June 2026 2:00pm to 4:00pm

PAPER 208

CONCENTRATION INEQUALITIES

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 Generalised Marton's Transport Cost Inequality

Let $P = P_1 \otimes P_2 \otimes \cdots \otimes P_N$ be a product measure on a discrete space $\mathcal{X}^N$. Consider a non-negative function $w : \mathcal{X} \times \mathcal{X} \to [0,\infty)$ and a convex function $\phi : [0,\infty) \to [0,\infty)$. Suppose for every $i = 1, 2, \ldots, N$ and for every probability measure $Q_i \ll P_i$, the following inequality is known to hold:

$$\inf_{\pi \in \Pi(P_i,Q_i)} \phi\left(\mathbb{E}_\pi[w(X_i,Y_i)]\right) \leqslant D(Q_i \| P_i).$$

Here, $\Pi(P_i, Q_i)$ is the set of all couplings of P_i and Q_i. Show that for any $Q \ll P$,

$$\inf_{\pi \in \Pi(P,Q)} \sum_{i=1}^{N} \phi\left(\mathbb{E}_\pi[w(X_i,Y_i)]\right) \leqslant D(Q \| P).$$

[You may assume without proof that all optimising couplings exist.]

2 Entropy method for self-bounding functions

Let $\mathcal{X}$ be a discrete set and let $f : \mathcal{X}^n \to [0,\infty)$ be a *self-bounding* function, that is, there exist functions $f_i : \mathcal{X}^{n-1} \to \mathbb{R}$ such that for all $x = (x_1, \ldots, x_n)$ and all $i = 1, 2, \ldots, n$, the following inequalities hold:

$$\begin{aligned} &0 \leqslant f(x) - f_i(x^{(i)}) \leqslant 1, \quad \text{and} \\ &\sum_{i=1}^{n} \left(f(x) - f_i(x^{(i)})\right) \leqslant f(x). \end{aligned}$$

Let $X_1, \ldots, X_n$ be independent random variables on $\mathcal{X}$ and let $Z = f(X_1, \ldots, X_n)$.

(a) Suppose that for all $\lambda \in \mathbb{R}$, the following inequality holds:

$$\log \mathbb{E} e^{\lambda(Z - \mathbb{E}Z)} \leqslant \phi(\lambda)\mathbb{E}Z, \tag{$\star$}$$

where $\phi(x) = e^x - x - 1$. Bound the left and right tail probabilities of $Z - \mathbb{E}Z$ using the Chernoff–Crámer method.

(b) Prove the inequality ($\star$) from part (a). You may use any results from the lectures without proof as long as you state or quote them clearly. [*Hint: Convexity of ϕ gives $\phi(tx) \leqslant t\phi(x)$ for all $t \in [0,1]$. Establish the differential inequality $\left(\frac{\psi(\lambda)}{e^\lambda - 1}\right)' \leqslant \mathbb{E}Z \left(\frac{-\lambda}{e^\lambda - 1}\right)'$, where $\psi(\lambda) = \log \mathbb{E} e^{\lambda(Z - \mathbb{E}Z)}$.*]

3 Concentration around the mean for certifiable functions

Let $X_1, X_2, \ldots, X_n$ be independent random variables on a discrete space $\mathcal{X}$ and let $f : \mathcal{X}^n \to \mathbb{N} = \{0, 1, 2, \ldots\}$. Let $Z = f(X_1, \ldots, X_n)$.

(a) Explain what it means to say that f satisfies the *bounded differences property* with $c_i = 1$ for $i = 1, 2, \ldots, n$. Explain what it means to say f is *g-certifiable* for some $g : \mathbb{N} \to \mathbb{N}$.

(b) Suppose f satisfies the bounded differences property with $c_i = 1$ for $i = 1, 2, \ldots, n$ and is g-certifiable for a $g : \mathbb{N} \to \mathbb{N}$. Prove that for all $t > 0$,

$$\mathbb{P}(Z - \mathbb{E}Z \leqslant -t) \leqslant e^{-\frac{t^2}{2\mathbb{E}[g(f(X))]}}.$$

You may use any results from the lectures provided you quote or state them clearly.

4 Maximum matching number in random graphs

For $n \geqslant 2$, consider $\binom{n}{2}$ random variables $\{X_{ij}\}_{1 \leqslant i < j \leqslant n}$ that are i.i.d. Bernoulli(p). Consider a random graph G on the vertex set $V = \{1, 2, \ldots, n\}$ such that each edge (i, j) is present if $X_{ij} = 1$ and absent otherwise, for $1 \leqslant i < j \leqslant n$.

A *matching* in a graph G is a collection of edges with the property that no two of them share a vertex. A *maximum matching* is a matching containing the largest possible number of edges. We denote this number by $f(G)$.

(a) Show that $\mathrm{Var}(f(G)) \leqslant \frac{\binom{n}{2}}{4}$.

(b) Show that for all $t > 0$,

$$\mathbb{P}(f(G) - \mathbb{E}[f(G)] \geqslant t) \leqslant \exp\left(-\frac{2t^2}{\binom{n}{2}}\right), \quad \text{and}$$
$$\mathbb{P}(f(G) - \mathbb{E}[f(G)] \leqslant -t) \leqslant \exp\left(-\frac{2t^2}{\binom{n}{2}}\right).$$

(c) Improve on the above bounds to show that for all $t > 0$,

$$\mathbb{P}(f(G) - \mathbb{E}[f(G)] \geqslant t) \leqslant \exp\left(-\frac{2t^2}{n-1}\right), \quad \text{and}$$
$$\mathbb{P}(f(G) - \mathbb{E}[f(G)] \leqslant -t) \leqslant \exp\left(-\frac{2t^2}{n-1}\right).$$

Now suppose $p = n^{-3/2}$. It is known (and you may assume without proof) that for all sufficiently large n, both the mean and the median of $f(G)$ lie in the interval $[c_1\sqrt{n}, c_2\sqrt{n}]$, for some constants $0 < c_1 < c_2 < 1/2$.

(d) Are the bounds in parts (a), (b), and (c) strong enough to show that $f(G)$ is tightly concentrated around its mean when $p = n^{-3/2}$? Justify your answer.

(e) Prove that when $p = n^{-3/2}$, $f(G)$ is much more tightly concentrated around its median than the bounds from parts (a)–(c) suggest. In particular, show deviations of larger order than $n^{1/4}$ from the median are unlikely.

[*You may use any results from the lectures provided you quote or state them clearly.*]

END OF PAPER