MAMA/208, NST3AS/208, MAAS/208

MAT3

MATHEMATICAL TRIPOS Part III

Friday 7 June 2024 1:30 pm to 3:30 pm

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 Sub-Gamma Random Variables

A real-valued random variable X with $\mathbb{E}[X] = 0$ is said to be *sub-gamma on the right tail with variance parameter* $\nu > 0$ *and scale parameter* $c > 0$ if

$$\psi_X(\lambda) \leqslant \frac{\lambda^2 \nu}{2(1 - c\lambda)} \quad \text{for every } \lambda \text{ such that} \quad 0 < \lambda < 1/c,$$

where $\psi_X(\lambda) = \log \mathbb{E}[e^{\lambda X}]$. The set of all such random variables is denoted by $\Gamma_+(\nu, c)$.

(a) Show that if $X \in \Gamma_+(\nu, c)$, then $\mathrm{Var}(X) \leqslant \nu$.

(b) Suppose $X_1, \ldots, X_n$ are independent random variables such that $X_i \in \Gamma_+(\nu_i, c_i)$. Show that $Z = \sum_{i=1}^n X_i$ satisfies $Z \in \Gamma_+(\nu, c)$, for suitable values of ν and c that you should determine.

(c) Let $X \in \Gamma_+(\nu, c)$. Show that for $t > 0$, the following right tail bound holds

$$\mathbb{P}(X \geqslant t) \leqslant \exp\left(-\frac{\nu}{c^2} h\left(\frac{ct}{\nu}\right)\right),$$

where h is a function that you should determine.

(d) Let $X \in \Gamma_+(\nu, c)$. Using part (c) or otherwise, show that for all $t > 0$,

$$\mathbb{P}(X \geqslant \sqrt{2\nu t} + ct) \leqslant e^{-t}.$$

2 Han's Inequality for KL-divergence

Let $\mathcal{X}$ be a finite discrete space, and let $P_{X_{1:n}}$ and $Q_{X_{1:n}}$ be two measures on $\mathcal{X}^n$ such that $P_{X_{1:n}}$ is a product measure; i.e., $P_{X_{1:n}} = P_{X_1} \otimes P_{X_2} \otimes \cdots \otimes P_{X_n}$ for measures P_{X_i} over $\mathcal{X}$, for $1 \leqslant i \leqslant n$.

Let $X_{1:n} \sim P_{X_{1:n}}$; i.e., $X_1, \ldots, X_n$ are independent random variables such that $X_i \sim P_{X_i}$ for $1 \leqslant i \leqslant n$. Let $f : \mathcal{X}^n \to [0, \infty)$ and set $Z = f(X_{1:n})$. Assume that $\mathbb{E}Z = 1$.

(a) State and prove *Han's inequality for KL-divergence*. If you use the chain rule of KL-divergence, you should prove it. [You may use the convexity of KL-divergence without proof.]

(b) Assuming Han's inequality for KL-divergence, state and prove the *tensorisation of entropy* theorem for upper-bounding $\mathrm{Ent}(Z)$.

(c) Prove the following inequality:

$$D(Q_{X_{1:n}} \| P_{X_{1:n}}) \leqslant \frac{1}{(n-1)} \sum_{i=1}^{n} D(Q_{X^{(i)}|X_i} \| P_{X^{(i)}} | Q_{X_i}).$$

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

(d) Assuming the inequality in part (c), derive a new upper bound on $\mathrm{Ent}(Z)$. [You may find it useful to introduce the notation $E^{[i]}Z = \mathbb{E}[Z|X_i]$ and $\mathrm{Ent}^{[i]}(Z) = E^{[i]}[Z \log Z] - E^{[i]}[Z] \log \mathbb{E}^{[i]}[Z]$.]

(e) Is the upper bound on $\mathrm{Ent}(Z)$ derived in part (d) tighter than the one from part (b)? Justify your answer.

3 Poisson log-Sobolev inequality

(a) State the *Gaussian log-Sobolev inequality.* Show that it implies the following inequality:

$$\operatorname{Ent}(f(X)) \leqslant 2\mathbb{E}\left[\frac{\|\nabla f(X)\|^2}{f(X)}\right],$$

where $X \sim N(0, I)$ is the standard Gaussian in $\mathbb{R}^n$ and $f : \mathbb{R}^n \to (0, \infty)$ is a continuously differentiable function.

Let $f : \mathbb{N} \to \mathbb{R}$, and define the discrete derivative of f evaluated at $x \in \mathbb{N}$ as $Df(x) := f(x+1) - f(x)$.

(b) Let X be a Poisson(λ) random variable; i.e., $\mathbb{P}(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}$ for $i \in \mathbb{N}$. Consider the following inequality, which is suggested as a version of the log-Sobolev inequality for the Poisson measure:

$$\operatorname{Ent}(f(X)^2) \leqslant C\mathbb{E}\left[|Df(X)|^2\right],$$

where $C > 0$ is a constant. Show that the above inequality cannot hold for all $f : \mathbb{N} \to \mathbb{R}$, regardless of the choice of C. [*Hint: Consider indicator functions of intervals* $[k+1, \infty)$.]

Let $\varepsilon \sim$ Bernoulli(p); i.e. $\mathbb{P}(\varepsilon = 1) = p$ and $\mathbb{P}(\varepsilon = 0) = 1 - p$ for some $p \in (0, 1)$. The following inequality is known to hold for all $f : \{0, 1\} \to (0, \infty)$:

$$\operatorname{Ent}(f(\varepsilon)) \leqslant p(1-p)\mathbb{E}\left[\frac{|Df(\varepsilon)|^2}{f(\varepsilon)}\right]. \tag{$\star$}$$

Here, we interpret $Df(0) = f(1) - f(0)$ and $Df(1) = f(0) - f(1)$.

(c) Let $X \sim$ Poisson(λ) for some $\lambda > 0$. Let $f : \mathbb{N} \to (0, \infty)$ be such that for all $x \in \mathbb{N}$, we have $0 < K_1 \leqslant f(x) \leqslant K_2$ and $|Df(x)| \leqslant K_3$, for some constants K_1, K_2, and K_3.

Using inequality ($\star$), prove that

$$\operatorname{Ent}(f(X)) \leqslant \lambda\mathbb{E}\left[\frac{|Df(X)|^2}{f(X)}\right]. \tag{$\star\star$}$$

[You may use any results from the lectures, provided you state or quote them clearly. You may also use the following fact without proof: if $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. Bernoulli(λ/n), then $S_n := \sum_{i=1}^n \varepsilon_i$ converges in distribution to a Poisson(λ) random variable as $n \to \infty$.]

[QUESTION CONTINUES ON THE NEXT PAGE]

(d) Let $X \sim \nu$ be a random variable that takes values in $\mathbb{N}$. Suppose ν satisfies the following inequality for all $f : \mathbb{N} \to (0, \infty)$:

$$\mathrm{Ent}(f(X)) \leqslant C\mathbb{E}\left[\frac{|Df(X)|^2}{f(X)}\right], \qquad (\star\star\star)$$

for some $C > 0$. Let $g : \mathbb{N} \to \mathbb{R}$ be a (discrete) 1-Lipschitz function; i.e., $|Dg(x)| \leqslant 1$ for all $x \in \mathbb{N}$ (note that g may take negative values). Using inequality $(\star\star\star)$, show that for $t > 0$,

$$\mathbb{P}(g(X) - \mathbb{E}g(X) \geqslant t) \leqslant \exp(-\psi^*(t)),$$

where ψ^* is the Chernoff-Cramér transform of $\psi(\lambda) = \frac{C\lambda}{2}(e^{2\lambda} - 1)$. [You may use the following fact without proof: if $|a| \leqslant \lambda$, then $|e^a - 1| \leqslant \lambda e^\lambda$.]

4 Convex and concave 1-Lipschitz functions

(a) State the *modified log-Sobolev inequality* and *Talagrand's one-sided bounded differences inequality.*

Let $f : [0,1]^n \to \mathbb{R}$ be a *convex* function whose partial derivatives exist. Let $g : [0,1]^n \to \mathbb{R}$ be a *concave* function whose partial derivatives exist. Suppose that both f and g are 1-Lipschitz; i.e. $\|\nabla f(x)\| \leqslant 1$ and $\|\nabla g(x)\| \leqslant 1$ for all $x \in [0,1]^n$.

Let $X_1, \ldots, X_n$ be independent random variables supported on $[0,1]$. Let $Z = f(X_{1:n})$ and $W = g(X_{1:n})$.

(b) Show that $\mathrm{Var}(Z) \leqslant 1$ and $\mathrm{Var}(W) \leqslant 1$. [If you use the convex Poincaré inequality, you should prove it. You may use any other results from the lectures provided you state or quote them clearly.]

(c) Show that for $t > 0$, each of the following probabilities is upper-bounded by $e^{-t^2/2}$: (i) $\mathbb{P}(Z - \mathbb{E}Z \geqslant t)$; (ii) $\mathbb{P}(Z - \mathbb{E}Z \leqslant -t)$; (iii) $\mathbb{P}(W - \mathbb{E}W \geqslant t)$; and (iv) $\mathbb{P}(W - \mathbb{E}W \leqslant -t)$. [You may use any results from the lectures, except those specifically asked for, without proof provided you state or quote them clearly.]

END OF PAPER