

# Examen

## Exercice 1

On dit qu'une variable aléatoire réelle  $X$  est symétrique si  $X$  et  $-X$  ont la même loi, ou, de façon équivalente, si  $X = \varepsilon|X|$ , avec  $\varepsilon = \text{sign}(X)$  valant  $+1$  ou  $-1$  avec probabilité  $1/2$ , indépendante de  $|X|$ .

Soient  $X_1, \dots, X_n$  des variables aléatoires réelles indépendantes symétriques. On suppose de plus que pour tout  $i = 1, \dots, n$ ,  $\mathbf{P}(X_i = 0) = 0$ . Montrer que pour tout  $t \geq 0$ ,

$$\mathbf{P}\left(\frac{X_1 + \dots + X_n}{\sqrt{X_1^2 + \dots + X_n^2}} \geq t\right) \leq \exp\left\{-\frac{t^2}{2}\right\}.$$

## Exercice 2

Soit  $G \sim \mathcal{G}(n, p)$  un graphe aléatoire d'Erdős-Rényi : l'ensemble de sommets correspond à  $\{1, \dots, n\}$  et, indépendamment pour chaque paire  $i < j$  de sommets distincts, on place une arête entre  $i$  et  $j$  avec probabilité  $p$ . Un sous-ensemble  $I \subset \{1, \dots, n\}$  de sommets est dit *indépendant* si aucune paire de sommets de  $I$  n'est reliée par une arête. On note  $\mathcal{I}$  l'ensemble des sous-ensembles indépendants de  $G$  et l'on s'intéresse à la variable

$$Z = \max_{I \in \mathcal{I}} |I|,$$

le cardinal maximal d'un sous-ensemble indépendant. Montrer que pour tout  $t \geq 0$ ,

$$\mathbf{P}(|Z - \mathbf{E}[Z]| \geq t) \leq 2 \exp\left\{-\frac{2t^2}{n}\right\}.$$

## Exercice 3

Soient  $X_1, \dots, X_n$  des variables aléatoires de Bernoulli indépendantes, de paramètres respectifs  $p_1, \dots, p_n \in [0, 1]$ . En identifiant un vecteur  $a \in \{0, 1\}^n$  avec le sous-ensemble des  $i \in \{1, \dots, n\}$  tels que  $a_i = 1$ , on définit, pour  $a \in \{0, 1\}^n$ , la variable

$$Y_a = \prod_{i \in a} X_i.$$

Soit maintenant  $\mathcal{A} \subset \{0, 1\}^n$  une collection de sous-ensembles. On s'intéresse à la variable

$$Z = \sum_{a \in \mathcal{A}} Y_a.$$

(1) Montrer que

$$\mathbf{Var}(Z) = \sum_{\substack{a, b \in \mathcal{A}, \\ a \cap b \neq \emptyset}} \left\{ \mathbf{E}[Y_a Y_b] - \mathbf{E}[Y_a] \mathbf{E}[Y_b] \right\}.$$

(2) Le but de cette question est de montrer que pour tout  $\lambda \leq 0$ , on a

$$\psi(\lambda) := \log \mathbf{E} \left[ e^{\lambda(Z - \mathbf{E}Z)} \right] \leq \frac{v}{c^2} \phi(c\lambda),$$

avec

$$v = \sum_{\substack{a, b \in \mathcal{A}, \\ a \cap b \neq \emptyset}} \mathbf{E}[Y_a Y_b] \quad \text{et} \quad c = \frac{v}{\mathbf{E}[Z]},$$

et  $\phi(u) = e^u - u - 1$ .

(a) Soit  $a \in \mathcal{A}$ . On définit

$$Z_a = \sum_{\substack{b \in \mathcal{A}, \\ a \cap b = \emptyset}} Y_b \quad \text{et} \quad \bar{Z}_a = \sum_{\substack{b \in \mathcal{A}, \\ a \cap b \neq \emptyset}} Y_b,$$

de telle sorte que  $Z = Z_a + \bar{Z}_a$ . En utilisant l'inégalité de Harris rappelée en fin d'énoncé, montrer que pour tout  $\lambda \leq 0$ ,

$$\mathbf{E} \left[ e^{\lambda Z} \mid Y_a = 1 \right] \geq \mathbf{E} \left[ e^{\lambda \bar{Z}_a} \mid Y_a = 1 \right] \mathbf{E} \left[ e^{\lambda Z_a} \right].$$

(b) En déduire que pour tout  $\lambda \leq 0$ , on a

$$\frac{\mathbf{E} [Z e^{\lambda Z}]}{\mathbf{E} [e^{\lambda Z}]} \geq \sum_{a \in \mathcal{A}} \mathbf{E}[Y_a] e^{\lambda \mathbf{E}[\bar{Z}_a \mid Y_a=1]},$$

puis que

$$\frac{\mathbf{E} [Z e^{\lambda Z}]}{\mathbf{E} [e^{\lambda Z}]} \geq \mathbf{E}[Z] e^{\lambda c}.$$

(c) En utilisant le fait que  $\psi(\lambda) = -\int_{\lambda}^0 \psi'(u) du$ , en déduire la majoration souhaitée sur  $\psi(\lambda)$  pour  $\lambda \leq 0$ .

(3) **Application.** Soit  $G \sim \mathcal{G}(n, p)$  un graphe aléatoire d'Erdős-Rényi, comme dans l'exercice 2 et soit  $Z$  le nombre de triangles dans  $G$ , c'est-à-dire le nombre de triplets de sommets  $i < j < k$  tels que  $\{i, j\}$ ,  $\{j, k\}$  et  $\{i, k\}$  appartiennent à  $E$ , l'ensemble des arêtes de  $G$ . Soit  $t \geq 0$ . En utilisant la question précédente, donner une majoration sous-gaussienne pour  $\mathbf{P}(Z - \mathbf{E}Z \leq -t)$ .

**Inégalité de Harris :** Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  et  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  deux fonctions croissantes et  $X = (X_1, \dots, X_n)$  un vecteur de variables aléatoires réelles indépendantes. On a

$$\mathbf{E} [f(X)g(X)] \geq \mathbf{E} [f(X)] \mathbf{E} [g(X)].$$

#### Exercice 4

Une fonction  $f : \mathcal{X}^n \rightarrow \mathbb{R}_+$  est dite  $(a, b)$ -auto-bornée, pour  $a, b \geq 0$ , s'il existe des fonctions  $f_i : \mathcal{X}^{n-1} \rightarrow \mathbb{R}_+$  telles que pour tout  $x \in \mathcal{X}^n$ , on a

$$0 \leq f(x) - f_i(x^{(i)}) \leq 1,$$

et

$$\sum_{i=1}^n \left( f(x) - f_i(x^{(i)}) \right) \leq af(x) + b.$$

Soit  $X = (X_1, \dots, X_n)$  un vecteur de variables aléatoires indépendantes à valeurs dans  $\mathcal{X}$  et  $Z = f(X)$  pour  $f$  une fonction  $(a, b)$ -auto-bornée.

(1) Montrer que  $\mathbf{Var}(Z) \leq a\mathbf{E}[Z] + b$ .

Pour  $\lambda \in \mathbb{R}$ , on note  $\psi(\lambda) = \log \mathbf{E} [e^{\lambda(Z - \mathbf{E}Z)}]$ .

- (2) Montrer que pour tout  $\lambda \in \mathbb{R}$ , on a

$$\lambda\psi'(\lambda) - \psi(\lambda) \leq (a\psi'(\lambda) + c)\phi(-\lambda),$$

où  $\phi(u) = e^u - u - 1$  et  $c = a\mathbf{E}[Z] + b$  (on pourra commencer par invoquer la Proposition 4.9 des notes de cours).

- (3) Montrer que  $\psi'(\lambda)$  est du même signe que  $\lambda$  (on pourra utiliser l'inégalité de Harris rappelée dans l'exercice précédent).
- (4) On considère d'abord le cas  $\lambda \geq 0$ . En utilisant le fait que  $\phi(u) \leq u^2/2$  pour  $u \leq 0$ , montrer que, pour tout  $\lambda \in [0, 2/a[$ , on a

$$\psi(\lambda) \leq \frac{c\lambda^2}{2 - a\lambda}.$$

On pourra introduire la fonction  $\Phi(u) = \frac{\psi(u)}{u}$ .

- (5) On considère maintenant le cas  $\lambda \leq 0$ .

- (a) Montrer que pour tout  $\lambda \leq 0$ , on a

$$\lambda\psi'(\lambda) - \psi(\lambda) \leq c\phi(-\lambda).$$

- (b) En utilisant le fait que la fonction  $u \mapsto \frac{\phi(u)}{u^2}$  est croissante sur  $\mathbb{R}$ , montrer que pour tout  $\lambda \leq 0$ , on a

$$\psi(\lambda) \leq c\phi(-\lambda).$$

- (6) En déduire que pour tout  $t \geq 0$ , on a

$$\mathbf{P}(Z - \mathbf{E}Z \geq t) \leq \exp\left(-\frac{t^2}{2\left(c + \frac{at}{2}\right)}\right),$$

et

$$\mathbf{P}(Z - \mathbf{E}Z \leq -t) \leq \exp\left(-\frac{t^2}{2\left(c + \frac{t}{3}\right)}\right).$$

- (7) **Application.** Soit  $g : \mathcal{X}^n \rightarrow \mathbb{R}_+$  telle que (i) faire varier la  $i^{\text{ème}}$  coordonnée fait varier la fonction d'au plus 1, (ii) si  $g(x) = y$ , alors il existe un sous-ensemble de coordonnées de taille au plus  $ay + b$  qui certifient que  $g(x) \geq y$  (une telle fonction est dite certifiable). Montrer que  $g$  est  $(a, b)$ -auto-bornée.