

ECOLES NORMALES SUPERIEURES

CONCOURS D'ADMISSION 2025

VENDREDI 18 AVRIL 2025

08h00 - 12h00

FILIERES MP et MPI

Epreuve n° 9

MATHEMATIQUES C (ULSR)

Fonctions d'un grand nombre de variables aléatoires

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Les variables aléatoires sur Ω intervenant dans le problème sont discrètes, à valeurs dans des sous-ensembles dénombrables de \mathbb{R} .

Rappels, notations

Soit X une variable aléatoire positive ou nulle, à valeurs dans l'ensemble $\{x_n; n \in \mathbb{N}\} \subset \mathbb{R}_+$. L'espérance de X est

$$\mathbb{E}[X] := \sum_{n=0}^{\infty} x_n \mathbb{P}(X = x_n). \quad (1)$$

Dans le cas général où X est une variable aléatoire à valeurs dans l'ensemble $\{x_n; n \in \mathbb{N}\} \subset \mathbb{R}$, et si $\mathbb{E}[|X|]$ est finie, l'espérance de X est par définition la quantité

$$\mathbb{E}[X] := \sum_{n=0}^{\infty} x_n \mathbb{P}(X = x_n) = \lim_{m \rightarrow +\infty} \sum_{n=0}^m x_n \mathbb{P}(X = x_n). \quad (2)$$

Si $\mathbb{E}[X^2]$ est finie, alors $\mathbb{E}[|X|]$ est finie et on a

$$(\mathbb{E}[|X|])^2 \leq \mathbb{E}[X^2]. \quad (3)$$

Dans ce cas, la *variance* de X est la quantité

$$\text{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2]. \quad (4)$$

On rappelle l'expression alternative

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2. \quad (5)$$

On rappelle aussi l'*inégalité de Markov* : si X est une variable aléatoire positive ou nulle d'espérance finie et $t > 0$, alors

$$\mathbb{P}(X \geq t) \leq \frac{1}{t} \mathbb{E}[X]. \quad (6)$$

On dit que des variables aléatoires X_1, X_2, \dots, X_N sont *indépendantes* si

Soit $N \in \mathbb{N}^*$. On considère N variables aléatoires X_1, \dots, X_N identiquement distribuées et satisfaisant, pour une certaine constante $K \geq 1$, et pour tout $n \in \{1, \dots, N\}$,

$$\mathbb{P}(|X_n| \leq K) = 1, \quad (8)$$

et

$$\mathbb{E}[X_n] = 0, \quad \text{Var}(X_n) \leq 1. \quad (9)$$

L'objet de cette partie est d'estimer la taille de la somme

$$S_N := X_1 + \dots + X_N, \quad (10)$$

sous diverses hypothèses de dépendance entre les variables X_n pour $n = 1, \dots, N$.

I.1) Pour $N = 1$, parmi les variables aléatoires de lois usuelles, donner sans justification un exemple de variable aléatoire satisfaisant (8) et deux exemples de variables aléatoires *ne satisfaisant pas* (8).

I.2) Pour tout $N \geq 1$, donner un exemple de variables aléatoires satisfaisant les hypothèses (8)-(9) et telles que $\mathbb{P}(|S_N| \geq N) \geq 1/2$.

I.3) On suppose que les variables aléatoires X_1, \dots, X_N sont deux à deux décorréelées, c'est-à-dire :

$$\forall 1 \leq m, n \leq N, \quad n \neq m \Rightarrow \mathbb{E}[X_n X_m] = 0. \quad (11)$$

Démontrer que

$$\mathbb{E}[|S_N|^2] \leq N. \quad (12)$$

En déduire que, pour tout $t > 0$,

$$\mathbb{P}(|S_N| > t\sqrt{N}) \leq \frac{1}{t^2}. \quad (13)$$

I.4) Soit k un entier supérieur ou égal à 2. On dit que des variables aléatoires $(Y_n)_{n \geq 1}$ sont k -indépendantes si

$$\mathbb{E}[\psi_1(Y_{n_1}) \cdots \psi_k(Y_{n_k})] = \mathbb{E}[\psi_1(Y_{n_1})] \cdots \mathbb{E}[\psi_k(Y_{n_k})] \quad (14)$$

pour tous indices $1 \leq n_1 < \dots < n_k$ et pour toutes fonctions bornées $\psi_1, \dots, \psi_k: \mathbb{R} \rightarrow \mathbb{R}$.

I.4.a) Démontrer que la k -indépendance implique la j -indépendance si $j \leq k$.

I.4.b) Qu'est-ce que la N -indépendance pour N variables aléatoires Y_1, \dots, Y_N ?

I.4.c) Soit Y_1 et Y_2 des variables aléatoires indépendantes de loi uniforme sur $\{0, 1\}$: pour $n = 1, 2$,

On introduit les notations suivantes : \mathcal{T} désigne l'ensemble $\{1, \dots, N\}^k$. Si $T = (n_1, \dots, n_k) \in \mathcal{T}$ et $n \in \{1, \dots, N\}$, on note $m_T(n)$ la multiplicité de n dans T , c'est-à-dire

$$m_T(n) = \text{Card} \{i \in \{1, \dots, k\}; n_i = n\}. \quad (17)$$

Pour $\ell \in \{1, \dots, k\}$, on note \mathcal{T}_ℓ l'ensemble des T dans \mathcal{T} faisant intervenir exactement ℓ indices distincts, et où chacun a une multiplicité au moins 2, à savoir : $T \in \mathcal{T}_\ell$ si

$$\text{Card}(\{n \in \{1, \dots, N\}; m_T(n) > 0\}) = \ell, \quad (18)$$

et

$$\forall n \in \{1, \dots, N\}, \quad m_T(n) > 0 \Rightarrow m_T(n) \geq 2. \quad (19)$$

Enfin, on notera $|\mathcal{T}_\ell|$ le cardinal de \mathcal{T}_ℓ .

I.5.a) Déterminer $|\mathcal{T}_1|$ et $|\mathcal{T}_\ell|$ pour $\ell > k/2$.

I.5.b) Justifier

$$\mathbb{E} \left[(S_N)^k \right] = \sum_{T \in \mathcal{T}} \prod_{n=1}^N \mathbb{E} \left[X_n^{m_T(n)} \right], \quad (20)$$

puis

$$\mathbb{E} \left[(S_N)^k \right] = \sum_{\ell=1}^{k/2} \sum_{T \in \mathcal{T}_\ell} \prod_{n=1}^N \mathbb{E} \left[X_n^{m_T(n)} \right]. \quad (21)$$

I.5.c) Démontrer que

$$\mathbb{E} \left[(S_N)^k \right] \leq \sum_{\ell=1}^{k/2} K^{k-2\ell} |\mathcal{T}_\ell|. \quad (22)$$

I.5.d) Soit $\ell \in \{1, \dots, k/2\}$. Justifier l'estimation suivante :

$$|\mathcal{T}_\ell| \leq \binom{N}{\ell} \ell^k \leq \frac{N^\ell}{\ell!} \ell^k. \quad (23)$$

On pourra considérer l'ensemble des $T \in \mathcal{T}$ faisant intervenir au plus ℓ éléments distincts.

I.5.e) Pour $\ell \in \{1, \dots, k/2\}$, démontrer que

$$\ell! \geq \ell^\ell e^{-\ell}, \quad (24)$$

puis en déduire que

$$\dots / k \setminus^{k-\ell} \dots$$

Démontrer que

$$\mathbb{E} \left[(S_N)^k \right] \leq \frac{\theta}{\theta - 1} \left(\frac{Nek}{2} \right)^{k/2} \leq 2 \left(\frac{Nek}{2} \right)^{k/2}, \quad (28)$$

où

$$\theta := \frac{2Ne}{kK^2}. \quad (29)$$

I.5.h) Démontrer (sous l'hypothèse (27)) l'estimation suivante : pour tout $t > 0$,

$$\mathbb{P}(|S_N| \geq t\sqrt{N}) \leq 2 \left(\frac{\sqrt{ek/2}}{t} \right)^k. \quad (30)$$

I.6) On suppose maintenant que les variables aléatoires X_1, \dots, X_N sont indépendantes, de sorte qu'elles sont k -indépendantes pour tout $k \in \{2, \dots, N\}$. On veut maintenant établir la borne suivante : il existe des constantes numériques $\alpha, \beta > 0$ (indépendantes de $K \geq 1$ et N) telles que pour tout $t \geq 0$,

$$\mathbb{P}(|S_N| \geq t\sqrt{N}) \leq \beta \exp(-\alpha t^2/K^2). \quad (31)$$

I.6.a) Justifier qu'il suffit de considérer le cas $K = 1$, ce que l'on fera dans les trois questions suivantes.

I.6.b) Soit k le plus grand entier pair de $\{1, \dots, N\}$ inférieur ou égal à $\frac{2t^2}{e^2}$. Justifier que (27) est satisfaite si

$$e \leq t \leq \frac{e}{\sqrt{2}} \sqrt{N}. \quad (32)$$

I.6.c) Sous l'hypothèse (32), démontrer qu'on a (31) avec

$$\beta = 2e, \quad \alpha = e^{-2}. \quad (33)$$

I.6.d) Conclure qu'il existe des constantes numériques $\alpha, \beta > 0$ telles que (31) est vérifiée pour tout $t \geq 0$.

II Concentration de combinaisons de variables aléatoires

L'espace \mathbb{R}^N est muni du produit scalaire canonique noté $\langle \cdot, \cdot \rangle$ et de la norme euclidienne associée, notée $|\cdot|$: pour $x \in \mathbb{R}^N$,

$$|x|^2 := \langle x, x \rangle^{1/2}. \quad (34)$$

Soit $(e_i)_{1 \leq i \leq N}$ la base canonique de \mathbb{R}^N . On rappelle que, pour tout $x \in \mathbb{R}^N$, on a

$$N \qquad N$$

$(X_i)_{1 \leq i \leq N}$ le vecteur aléatoire de composantes X_1, \dots, X_N . Alors il existe des constantes numériques $\alpha, \beta > 0$ telles que pour toute fonction 1-lipschitzienne et convexe $F: \mathbb{R}^N \rightarrow \mathbb{R}$, pour tout $m \in \mathbb{R}$, pour tout $t \geq 0$,

$$\mathbb{P}(F(X) \geq m) \geq \frac{1}{2} \implies \mathbb{P}(F(X) \leq m - t) \leq \beta e^{-\alpha t^2 / K^2}. \quad (37)$$

On introduit les notations et conventions suivantes. On désigne par Q^N l'hypercube $Q^N := [-K, K]^N$. Notons que le diamètre de Q^N est

$$\text{diam}(Q^N) := \max \{|x - y|; x, y \in Q^N\} = 2\sqrt{N}K. \quad (38)$$

Soit $x \in Q^N$ et $A \subset Q^N$. Si A est non-vide, on note $d(x, A)$ la distance de x à A définie par

$$d(x, A) := \inf_{a \in A} |x - a|. \quad (39)$$

On adopte par ailleurs la convention suivante : si A est vide, alors on pose $d(x, A) = 4NK$ (la valeur précise choisie ici n'a pas grande importance, tant qu'elle reste strictement plus grande que $\text{diam}(Q^N)$).

II.1) Soit γ une constante numérique positive. On suppose vérifiée la propriété suivante : pour tout ensemble $A \subset Q^N$ convexe,

$$\mathbb{P}(X \in A) \mathbb{E} \left[\exp \left(\gamma \frac{d(X, A)^2}{4K^2} \right) \right] \leq 1. \quad (40)$$

On se donne une fonction 1-lipschitzienne et convexe $F: \mathbb{R}^N \rightarrow \mathbb{R}$. Le but de cette question est de démontrer que (37) est alors vérifiée.

II.1.a) Soient $s, \sigma \in \mathbb{R}$ avec $s < \sigma$. En considérant l'ensemble

$$A_s = \{x \in Q^N; F(x) \leq s\}, \quad (41)$$

montrer que

$$\mathbb{P}(F(X) \leq s) \mathbb{P}(F(X) \geq \sigma) \leq \exp \left(-\gamma \frac{(\sigma - s)^2}{4K^2} \right). \quad (42)$$

II.1.b) Démontrer que (37) est vérifiée.

II.2) Soit x un point arbitraire de Q^N . Soit P^N l'ensemble des sommets de l'hypercube $[0, 1]^N$, c'est-à-dire l'ensemble des combinaisons linéaires des e_i pour $i \in \{1, \dots, N\}$ à coefficients 0 ou 1. Si A est une partie non-vide de Q^N , on définit les sous-ensembles $P_A(x)$ et $R_A(x)$ de P^N comme suit : soit H_i l'hyperplan orthogonal à e_i , engendré par les e_j pour $j \neq i$. Alors $z \in P_A(x)$ si il existe $a \in A$ tel que

II.2.c) En dimension $N = 3$, donner un exemple d'ensemble A pour lequel $e_3 \notin P_A(0)$ et décrire précisément les ensembles $R_A(0)$ et $P_A(0)$ correspondant.

Si B est un sous-ensemble non-vide de \mathbb{R}^N , on note $\Gamma(B)$ l'enveloppe convexe de B , qui est l'ensemble des combinaisons convexes (finies) d'éléments de B :

$$\Gamma(B) := \left\{ \sum_{j=1}^m \theta_j z_j; m \in \mathbb{N}^*, \theta_j \in [0, 1], z_j \in B, \sum_{j=1}^m \theta_j = 1 \right\}. \quad (47)$$

Étant donné $A \subset Q^N$ non-vide et $x \in Q^N$, on définit aussi la quantité

$$q(x, A) := \inf \{ |z|; z \in \Gamma(P_A(x)) \}. \quad (48)$$

On adopte par ailleurs la convention suivante : si A est l'ensemble vide, on pose $q(x, A) = 2N$.

II.2.d) Soit B un sous-ensemble non-vide de \mathbb{R}^N . On note $\Gamma_0(B)$ l'ensemble des combinaisons convexes d'au plus $N + 1$ éléments de B :

$$\Gamma_0(B) := \left\{ \sum_{j=1}^{N+1} \theta_j z_j; \theta_j \in [0, 1], z_j \in B, \sum_{j=1}^{N+1} \theta_j = 1 \right\}. \quad (49)$$

Démontrer que $\Gamma(B) = \Gamma_0(B)$.

Indication : on pourra démontrer que toute combinaison convexe de $m + 1$ éléments de B avec $m > N$ peut se réécrire comme combinaison convexe d'au plus m éléments de B .

II.2.e) Soit B un sous-ensemble non-vide de \mathbb{R}^N . Démontrer que $\Gamma(B)$ est un ensemble convexe, et qu'il est compact si B l'est.

II.2.f) Représenter graphiquement et nommer (en tant qu'objet géométrique) l'enveloppe convexe $\Gamma(B)$ en dimension $N = 3$, dans les trois cas suivants :

$$B = \{e_1, e_2, e_1 + e_2\}, \quad B = \{e_1, e_2, e_1 + e_2, e_2 + e_3\}, \quad B = P^3. \quad (50)$$

Pour chacun de ces exemples, dire si B peut correspondre à un ensemble $P_A(x)$.

II.2.g) Soit $A \subset Q^N$ non-vide et $x \in Q^N$. Justifier que l'inf dans (48) est atteint.

II.2.h) Soit $A \subset Q^N$ non-vide et $x \in Q^N$. Justifier que $q(x, A) \leq \sqrt{N}$. A quelle condition a-t-on $q(x, A) = 0$?

II.2.i) Soit $x \in Q^N$ et $A \subset Q^N$ non-vide. Justifier que

$$q(x, A) = \inf \{ |z|; z \in \Gamma(R_A(x)) \}. \quad (51)$$

L'objet de la fin de cette partie II est la *preuve par récurrence sur la dimension N* de la propriété (53), pour $\gamma = \gamma_0$. Pour N entier naturel non nul, on introduit donc l'hypothèse de récurrence H_N suivante :

« Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit N variables aléatoires X_1, \dots, X_N à valeurs dans un ensemble fini, indépendantes et identiquement distribuées, satisfaisant (8) et soit $X = (X_i)_{1 \leq i \leq N}$ le vecteur aléatoire de composantes X_1, \dots, X_N . Alors (53) est vérifiée pour $\gamma = \gamma_0 = \frac{1}{4}$ ».

II.3.a) On considère le cas $N = 1$. Démontrer que (53) est satisfait lorsque $\gamma \leq \ln(2)$, et donc pour $\gamma = \gamma_0$.

On suppose maintenant $N > 1$, et on se fixe $A \subset Q^N$ convexe. On adopte les notations suivantes : on décompose

$$x = (\bar{x}, x_N) \quad \text{avec} \quad \bar{x} = (x_i)_{1 \leq i \leq N-1} \in \mathbb{R}^{N-1}. \quad (55)$$

Si $A \subset \mathbb{R}^N$ et $\theta \in \mathbb{R}$, on note

$$A_\theta := \{b \in \mathbb{R}^{N-1}; (b, \theta) \in A\} \quad (56)$$

la section de A au niveau θ . On note aussi

$$\bar{A} := \{\bar{a} \in \mathbb{R}^{N-1}; \exists \theta \in \mathbb{R}, (\bar{a}, \theta) \in A\} \quad (57)$$

la projection de A sur \mathbb{R}^{N-1} .

II.3.b) Soit $x \in Q^N$. Soit $A \subset Q^N$ tel que A_{x_N} soit non-vidé. Soit $\bar{z} \in P^{N-1}$. Démontrer que

$$\bar{z} \in P_{A_{x_N}}(\bar{x}) \implies (\bar{z}, 0) \in P_A(x) \quad (58)$$

et

$$\bar{z} \in P_{\bar{A}}(\bar{x}) \implies (\bar{z}, 1) \in P_A(x). \quad (59)$$

II.3.c) Démontrer que, pour tout $\lambda \in [0, 1]$, on a

$$q(x, A)^2 \leq (1 - \lambda)^2 + \lambda q(\bar{x}, A_{x_N})^2 + (1 - \lambda) q(\bar{x}, \bar{A})^2. \quad (60)$$

On fixe $x_N \in \mathbb{R}$ tel que $\mathbb{P}(X_N = x_N) > 0$ et on considère la probabilité

$$\bar{\mathbb{P}}(B) := \mathbb{P}(B | X_N = x_N) = \frac{\mathbb{P}(B \cap \{X_N = x_N\})}{\mathbb{P}(X_N = x_N)} \quad \text{pour } B \in \mathcal{A}, \quad (61)$$

ainsi que l'espérance associée

$$\bar{\mathbb{E}}[Z] := \frac{1}{\mathbb{P}(X_N = x_N)} \mathbb{E}[Z \mathbf{1}_{\{X_N = x_N\}}] \quad (62)$$

$$\bar{\mathbb{E}} \left[e^{\lambda Y} e^{(1-\lambda)Z} \right] \leq \{ \bar{\mathbb{E}} [e^Y] \}^\lambda \{ \bar{\mathbb{E}} [e^Z] \}^{1-\lambda} \quad (65)$$

pour $\lambda \in [0, 1]$ et Y, Z des variables aléatoires. On prendra par ailleurs soin de bien préciser quel est l'espace probabilisé considéré dans l'application de H_{N-1} .

II.3.e) On suppose

$$\bar{\mathbb{P}}(\bar{X} \in \bar{A}) > 0, \quad (66)$$

et on définit

$$r = \frac{\bar{\mathbb{P}}(\bar{X} \in A_{x_N})}{\bar{\mathbb{P}}(\bar{X} \in \bar{A})}. \quad (67)$$

Démontrer que

$$r^\lambda e^{-\gamma_0(1-\lambda)^2} \bar{\mathbb{P}}(\bar{X} \in \bar{A}) \bar{\mathbb{E}} [\exp(\gamma_0 q(X, A)^2)] \leq 1. \quad (68)$$

II.3.f) On admet provisoirement l'inégalité suivante : pour tout $\gamma \in [0, \gamma_0]$, pour tout $r \in]0, 1]$,

$$\frac{1}{2-r} \leq \sup_{\lambda \in [0,1]} r^\lambda e^{-\gamma(1-\lambda)^2}. \quad (69)$$

Justifier que

$$\bar{\mathbb{P}}(\bar{X} \in \bar{A}) \bar{\mathbb{E}} [\exp(\gamma_0 q(X, A)^2)] \leq (2-r). \quad (70)$$

On distinguera les cas $\bar{\mathbb{P}}(\bar{X} \in A_{x_N}) > 0$ et $\bar{\mathbb{P}}(\bar{X} \in A_{x_N}) = 0$.

II.3.g) Démontrer que

$$\mathbb{P}(X \in A) \mathbb{E} [\exp(\gamma_0 q(X, A)^2) \mathbf{1}_{\{X_N = x_N\}}] \leq R(2-r) \mathbb{P}(X_N = x_N), \quad (71)$$

où

$$R = \frac{\mathbb{P}(X \in A)}{\bar{\mathbb{P}}(\bar{X} \in \bar{A})}. \quad (72)$$

II.3.h) Démontrer que

$$\mathbb{P}(X \in A) \mathbb{E} [\exp(\gamma_0 q(X, A)^2)] \leq R(2-R), \quad (73)$$

où R est défini dans (72), puis démontrer (53) et conclure l'hérédité $H_{N-1} \Rightarrow H_N$. On prendra bien garde à tenir compte du cas où (66) n'est pas vérifié.

II.3.i) Justifier (69).