260 ESPÉRANCE, VARIANCE ET MOMENTS D'UNE VARIABLE ALÉATOIRE.

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et $(E, \mathcal{E})$ un espace probabilisable.

I. Autour de l'espérance et de la variance [BL07, §III.4/5, p52–71]

I. A. Espérance

Définition 1. [Variable aléatoire]
Une variable aléatoire X est une fonction mesurable $X : (\Omega, \mathcal{A}) \longrightarrow (E, \mathcal{E})$.
On note $\mathbb{P}_X$ la mesure image de $\mathbb{P}$ par X et on l'appelle loi de X.
Si on a $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on dit que X est une variable aléatoire réelle.

Soit X une variable aléatoire à valeurs dans $E = \mathbb{R}^d$ où $d \in \mathbb{N}^*$.

Définition 2. [Espérance]
- Si $E = \mathbb{R}_+$, on définit l'espérance de X par $\mathbb{E}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega) \in [0, +\infty]$,
- Si $E = \mathbb{R}$ et $\mathbb{E}[|X|] < +\infty$, on définit l'espérance de X par la même formule,
- Si $E = \mathbb{R}^d$, on définit l'espérance de $X = (X_1, \dots, X_d)$ lorsque $\mathbb{E}[\|X\|] < +\infty$ par $\mathbb{E}[X] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])$.

Exemple 3. Espérance d'une loi binomiale et d'une loi normale, … **[voir annexe]**
Les lois de Cauchy n'admettent pas d'espérance.

Remarque 4. L'espérance est une forme linéaire positive et les résultats classiques d'intégration sont valables pour l'espérance : lemme de Fatou, théorèmes de convergence dominée et convergence monotone, inégalité de Cauchy-Schwarz.

Proposition 5. [Lemme de transfert]
Si X est à valeurs dans $\mathbb{R}^d$, soit $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ une fonction borélienne.
Alors $f(X) \in L^1(\Omega, \mathcal{A}, \mathbb{P}) \iff f \in L^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_X)$ et dans ce cas :

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega)) d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x)$$

Application 6. Si $X \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$ et $\theta > 0$, alors $\theta X \sim \mathcal{E}(\lambda/\theta)$.
Si $X \sim \mathcal{G}(p)$ et $f : n \longmapsto \mathbb{1}_{n \text{ est pair}}$, alors $\mathbb{E}[f(X)] = 1/3$.

Remarque 7. Si X est réelle positive, on a $\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > t) dt$. Dans le cas où X est à valeurs dans $\mathbb{N}$, on a $\mathbb{E}[X] = \sum_{n \geq 1} \mathbb{P}(X \geq n)$.

Proposition 8. [Espérance conditionnelle]
Soit $\mathcal{B}$ une sous-tribu de $\mathcal{A}$. Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$, intégrable.
Alors il existe une unique (presque sûrement) variable aléatoire Z telle que :
(i) Z est $\mathcal{B}$-mesurable,
(ii) pour tout $B \in \mathcal{B}$, $\mathbb{E}[X \mathbb{1}_B] = \mathbb{E}[Z \mathbb{1}_B]$.
De plus, Z est intégrable.

Définition 9. [Espérance conditionnelle]
Z est appelée espérance conditionnelle de X sachant $\mathcal{B}$, et est notée $\mathbb{E}[X \mid \mathcal{B}]$.

Application 10. Soient X une variable aléatoire indépendante de $\mathcal{B}$ et Y une variable aléatoire $\mathcal{B}$-mesurable. Alors pour toute fonction ψ telle que $\psi(X, Y)$ est intégrable :

$$\mathbb{E}[\psi(X, Y) \mid \mathcal{B}] = \int \psi(x, Y)\, d\mu_X(x) = h(Y) \qquad \text{où} \qquad h(y) = \mathbb{E}[\psi(X, y)]$$

Exemple 11. Soit $(X_i)_{i \in \mathbb{N}^*}$ des v.a. i.i.d. intégrables. Alors si $S_n = \sum_{k=1}^n X_i$, on a pour $p < q$, $\mathbb{E}[S_q \mid S_p] = S_p + (q - p)\mathbb{E}[X_1]$.

I. B. Variance

Soient X et Y deux variables aléatoires à valeurs dans $E = \mathbb{R}^d$ de carré intégrable.

Définition 12. [Covariance, variance]
Si $d = 1$, on définit la covariance de X et Y par $\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$, la variance de X par $\mathrm{Var}(X) = \mathrm{Cov}(X, X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$.
Sinon, on définit $\Sigma = (\mathrm{Cov}(X_i, X_j))_{1 \leq i, j \leq d}$ la matrice de covariance de $X = (X_1, \dots, X_d)$.

Application 13. Si $A \in \mathcal{M}_d(\mathbb{R})$ et $b \in \mathbb{R}^d$ alors $\Sigma_{AX+b} = A \Sigma_X A^\top$.

Proposition 14. *Cov est une forme bilinéaire symétrique positive sur L^2 et on a $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ et $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.*

Remarque 15. $\mathrm{Var}(X)$ s'interprète comme la projection de X sur l'espace des fonctions constantes presque sûrement.

Exemple 16. Variance d'une loi de poisson et d'une loi exponentielle, … **[voir annexe]**

Proposition 17. *Si $X \perp\!\!\!\perp Y$ alors $\mathrm{Cov}(X,Y)=0$ et $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$.*

Exemple 18. La réciproque est fausse : prendre $X \sim \mathcal{N}(0,1)$ ou $X \sim \mathcal{U}([-1,1])$ et $Y = X^2$.

I. C. Inégalités importantes en probabilités

Soit X une variable aléatoire réelle.

Proposition 19. [Inégalité de Markov]
Si $X \geq 0$ p.s., alors $\forall t > 0, \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$.

Proposition 20. [Inégalité de Bienaymé-Tchebychev]
On a $\forall t > 0, \mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}$.

Proposition 21. [Inégalité de Jensen]
Si $X \in L^1$ et ϕ est convexe, alors $\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$. De plus, si ϕ est strictement converxe, on a égalité si et seulement si X est constante presque sûrement.

Remarque 22. En particulier $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$ et $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$. L'inégalité de Jensen s'applique en dehors du cadre des probabilités, par exemple pour montrer que $f * g$ est bien défini et $\|f * g\|_p \leq \|f\|_1 \|g\|_p$ lorsque $f \in L^1, g \in L^p$.

Application 23. Si $(X_i)_{1 \leq i \leq n}$ sont des v.a. i.i.d. d'espérance m et de variance σ^2 finies, alors

$$\forall t > 0, \mathbb{P}\Big(\frac{1}{n}\Big|\sum_{k=1}^{n} X_i - m\Big| \geq t\Big) \leq \frac{\sigma^2}{nt^2}$$

Application 24. [Théorème de Weierstrass]
$\mathbb{R}[X]$ est dense dans $\mathcal{C}([a,b], \mathbb{C}, \|.\|_\infty)$ pour $a < b \in \mathbb{R}$.

II. Moments d'une variable aléatoire

[BL07, §III.4/5, p52–71]

Soit X une variable aléatoire réelle. Soit $p \in \mathbb{N}^*$.

II. A. Généralités

Définition 25. [Moment et moment centré d'ordre p]
Si $\mathbb{E}[|X|^p] < +\infty$, on dit que X admet un moment d'ordre p (ou $X \in L^p$) et alors on appelle $\mathbb{E}[X^p]$ le moment d'ordre p et $\mathbb{E}[(X - \mathbb{E}[X])^p]$ le moment centré d'ordre p de X.

Remarque 26. Si X admet un moment d'ordre p, l'inégalité de Markov assure que $\mathbb{P}(|X - \mathbb{E}[X]| \geq x) = \mathbb{P}(|X - \mathbb{E}[X]|^p \geq x^p) \leq \frac{\mathbb{E}[|X-\mathbb{E}[X]|^p]}{x^p}$. Plus X admet de moments d'ordre élevé, plus on a un contrôle sur la queue de distribution de X.

Proposition 27. *Si $X \in L^p$ alors $X \in L^q$ pour tout $1 \leq q \leq p$.*

Définition 28. [Fonction génératrice des moments]
On définit M_X la fonction génératrice des moments de X par $M_X(\lambda) = \mathbb{E}[\mathrm{e}^{\lambda X}]$ pour $\lambda \in \mathbb{R}$ tel que cette quantité soit définie.

Proposition 29. *Si M_X est définie sur un voisinage de 0, alors M_X caractérise la loi de X et alors on a $M_X(\lambda) = \sum_{k \geq 0} \frac{t^k}{k!}\mathbb{E}[X^k]$ au voisinage de 0, en particulier X admet des moments de tout ordre et $\mathbb{E}[X^p] = M_X^{(p)}(0)$ pour tout p.*

Exemple 30. $M_{\mathcal{N}(0,1)} : \lambda \longmapsto \exp(\lambda^2/2)$ et $M_{\mathcal{R}(1/2)} : \lambda \longmapsto \mathrm{ch}(\lambda)$.

Proposition 31. $\forall t \in \mathbb{R}, \mathbb{P}(X \geq t) \leq \inf_{\lambda \geq 0}\{\mathrm{e}^{-\lambda t} M_X(\lambda)\}$ **[Borne de Chernoff]**

Application 32. [Variables aléatoires sous-gaussiennes]
On dit que X est K-sous-gaussienne si $M_X(\lambda) \leq \exp(\lambda^2 K^2/2)$ pour un $K > 0$.
On a alors X centrée, $\mathbb{V}(X) \leq \mathbb{E}(X^2) \leq K^2$ et pour $t \geq 0, \mathbb{P}(|X| \geq t) \leq 2\exp(-t^2/2K^2)$.
Si $X \sim \mathcal{N}(0,1)$, X est 1-sous-gaussienne.
Si $X \in [a,b]$ p.s., $X - \mathbb{E}[X]$ est $(b-a)/2$-sous-gaussienne (lemme de Hoeffding, admis).
Si $(X_i)_{1 \leq i \leq n}$ est une famille de variables aléatoires indépendantes et si pour tout i, X_i est K_i-sous gaussienne, alors $\sum_{i=1}^n X_i$ est $\sqrt{\sum_{i=1}^n K_i^2}$ sous-gaussienne.
En particulier, on obtient l'inégalité de Hoeffding : si pour tout i, $X_i \in [a_i, b_i]$ p.s., alors :

$$\forall t \geq 0, \mathbb{P}\left(|\textstyle\sum_{i=1}^n (X_i - \mathbb{E}[X_i])| \geq t\right) \leq 2\exp(-2t^2/\textstyle\sum_{i=1}^n (b_i - a_i)^2)$$

Théorème 33. [Théorème des moments] *(admis)*
Si X et Y sont bornées p.s. et si leurs moments sont égaux à tout ordre, alors $X \sim Y$.

Remarque 34. C'est faux si X et Y ne sont pas bornées : les considérer de densités $f_X : x \longmapsto \frac{1}{\sqrt{2\pi}x}\mathrm{e}^{-\ln(x)^2/2}\mathbb{1}_{\mathbb{R}_+}(x)$ et $f_Y : x \longmapsto f_X(x)(1 + a\sin(2\pi\ln(x)))$ pour un $a \in [-1,1]$.

II. B. Liens avec les fonctions génératrices et caractéristiques

[BL07, §III.5, p61]

Soient X, Y des variables aléatoires à valeurs dans $\mathbb{R}^d$ ou $\mathbb{N}$ selon le contexte.

DÉFINITION 35. [FONCTION GÉNÉRATRICE, FONCTION CARACTÉRISTIQUE]
Si X est à valeurs dans $\mathbb{N}$ presque sûrement, on définit g_X la fonction génératrice de X par $g_X(s) = \mathbb{E}[s^X] = \sum_{k\geq 0} \mathbb{P}(X = k)s^k$ (lorsque la série converge).
On définit ϕ_X la fonction caractéristique de X par $\phi_X(\lambda) = \mathbb{E}\left[e^{i\langle\lambda|X\rangle}\right]$ pour $\lambda \in \mathbb{R}^d$.

REMARQUE 36. Le rayon de convergence de cette série est au moins égal à 1 et $G_X(1) = 1$.

EXEMPLE 37. Fonctions génératrices de $\mathcal{B}(p), \mathcal{B}(n,p), \mathcal{P}(\lambda), \mathcal{G}(p)$ ... Fonctions caractéristiques de $\mathcal{B}(n,p), \mathcal{E}(\lambda)$ ...**[VOIR ANNEXE]**

PROPOSITION 38. *ϕ_X et G_X caractérisent la loi de X. En particulier, $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{G^{(k)}(0)}{k!}$.*

PROPOSITION 39.
- *X admet un moment d'ordre $p \in \mathbb{N}^*$ si et seulement si g_X est p fois dérivable en 1, et dans ce cas $G_X^{(p)}(1) = \mathbb{E}[X(X-1)\dots(X-p)]$,*
- *Si X est réelle et admet un moment d'ordre p, alors ϕ_X est p fois dérivable et on a $\phi_X^{(p)}(\lambda) = i^p\mathbb{E}[X^p e^{i\lambda X}]$. En particulier, $\phi^{(p)}(0) = i^p\mathbb{E}[X^p]$.*

APPLICATION 40. Expression de l'espérance, la variance, ou plus généralement d'un moment en fonction des dérivées en 0 de G_X.

PROPOSITION 41. *Si $X \perp\!\!\!\perp Y$ alors $G_{X+Y} = G_X G_Y$ ou selon le contexte $\forall(\lambda_1,\lambda_2) \in \mathbb{R}^d \times \mathbb{R}^\ell, \phi_{(X+Y)}(\lambda_1,\lambda_2) = \phi_X(\lambda_1)\phi_Y(\lambda_2)$.*

APPLICATION 42. [PROCESSUS DE BRANCHEMENT DE GALTON-WATSON]
Soit $(Z_n)_{n\in\mathbb{N}}$ un processus de branchement de GALTON-WATSON de loi μ. Alors $g_{Z_n} = g_\mu^{\otimes n}$ pour tout $n \in \mathbb{N}$ et si $\mu \neq \delta_0$, on a $\mathbb{P}(\text{survie du processus}) > 0 \iff \mathbb{E}[\mu] > 1$.

APPLICATION 43. Les composantes d'un vecteur gaussien sont indépendantes si et seulement si sa matrice de covariance est diagonale.

APPLICATION 44. Calcul de la loi d'une variable aléatoire multinomiale dont le paramètre n n'est pas déterministe mais suit une loi $\mathcal{P}(\lambda)$.

III. Théorèmes limites

[BL07, §V.5, p131–147]

Des conditions sur les moments d'une loi apparaissent naturellement dans plusieurs énoncés fondamentaux de convergence de variables aléatoires.

III. A. Loi des grands nombres

THÉORÈME 45. [LOI DES GRANDS NOMBRES]
Soit $(X_i)_{i\in\mathbb{N}^*}$ des variables aléatoires réelles i.i.d. intégrables. Alors :

$$\frac{1}{n}\sum_{i=1}^n X_i \underset{n\to+\infty}{\longrightarrow} \mathbb{E}[X_1] \quad \text{p.s.}$$

L'hypothèse d'intégrabilité est cruciale :

PROPOSITION 46. *Soient $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d. telle que $\mathbb{E}[|X_1|] = +\infty$. Alors la probabilité que la limite de $\frac{1}{n}\sum_{i=1}^n X_i$ existe dans $\mathbb{R}$ est nulle.*

APPLICATION 47. [MÉTHODE DE MONTE-CARLO] [Mé11]
Soit $f : [0,1] \longrightarrow \mathbb{R}$ intégrable par rapport à la mesure de LEBESGUE et $(X_i)_{n\in\mathbb{N}^*}$ une suite de variables aléatoiresi.i.d. de loi $\mathcal{U}([0,1])$. Alors $I_n = \frac{1}{n}\sum_{i=1}^n f(X_i) \underset{n\to+\infty}{\longrightarrow} \int_0^1 f(x)dx$ p.s..

III. B. Théorème central limite

THÉORÈME 48. [THÉORÈME DE LÉVY (FAIBLE)]
Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Alors $X_n \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} X$ si et seulement si $\phi_{X_n} \underset{n\to+\infty}{\xrightarrow{s}} \phi_X$.

THÉORÈME 49. [THÉORÈME CENTRAL LIMITE]
Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires réelles i.i.d. de carré intégrable. Alors :

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - \mathbb{E}[X_1]) \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} Z \sim \mathcal{N}(0, \mathrm{Var}(X_1))$$

APPLICATION 50. [INTERVALLE DE CONFIANCE ASYMPTOTIQUE] [RS12, §3.4, p25]
Soient $(X_i)_{i\in\mathbb{N}^*}$ des réalisations i.i.d. de $\mathcal{B}(p)$ pour un $p \in [0,1]$ inconnu. Le théorème central limite donne un intervalle de confiance asymptotique de niveau α pour p en fonction de la moyenne empirique $\hat{p}_n = \frac{1}{n}\sum_{k=1}^n X_i : \mathrm{IC}_\alpha(p) = \left[\hat{p}_n \pm \frac{q_{1-\alpha/2}}{2}\frac{1}{\sqrt{n}}\right]$, où q_t est le quantile d'ordre t de $\mathcal{N}(0,1)$.

ANNEXE

[BL07, Appendice, p229]

Tableau des principales lois discrètes

Soit X une loi discrète et $p \in [0,1], q \in]0,1[, n \in \mathbb{N}, \lambda > 0$.

$\mathcal{L}(X)$	$\mathbb{P}_X$	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	$\phi_X(t)$	$G_X(t)$
$\mathcal{U}([\![1,n]\!])$	$\sum_{k=1}^{n} \frac{1}{n}\delta_k$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{1}{n}\sum_{k=1}^{n} \mathrm{e}^{ikt}$	$\frac{1}{n}\frac{t-t^{n+1}}{1-t} \quad (t \neq 1)$
$\mathcal{B}(p)$	$p\delta_1 + (1-p)\delta_0$	p	$p(1-p)$	$p\,\mathrm{e}^{it} + (1-p)$	$1-p+pt$
$\mathcal{B}(n,p)$	$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}\delta_k$	np	$np(1-p)$	$p\,\mathrm{e}^{it} + (1-p)^n$	$(1-p+pt)^n$
$\mathcal{P}(\lambda)$	$\sum_{k=0}^{\infty} \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}\delta_k$	λ	λ	$\mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}$	$\mathrm{e}^{(\lambda-1)t}$
$\mathcal{G}(q)$	$\sum_{k=1}^{\infty} q(1-q)^{k-1}\delta_k$	$\frac{1}{q}$	$\frac{1-q}{q^2}$	$\frac{q\,\mathrm{e}^{it}}{1-(1-q)\,\mathrm{e}^{it}}$	$\frac{pt}{1-(1-p)t}$

Tableau des principales lois à densité

Soit X une loi admettant une densité f par rapport à la mesure de LEBESGUE. Soit $a < b \in \mathbb{R}, m \in \mathbb{R}, \sigma > 0, \lambda > 0, \alpha, \beta > 0$.

$\mathcal{L}(X)$	$f(t)$	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	$\phi_X(t)$
$\mathcal{U}([a,b])$	$\frac{\mathbb{1}_{[a,b]}(t)}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{\mathrm{e}^{itb}-\mathrm{e}^{ita}}{it(b-a)}$
$\mathcal{N}(m,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	m	σ^2	$\mathrm{e}^{imt-\frac{t^2\sigma^2}{2}}$
$\mathcal{C}(m,\sigma^2)$	$\frac{1}{\pi\sigma}\frac{1}{1+(\frac{x-m}{\sigma})^2}$	/	/	$\mathrm{e}^{-\lvert t\rvert}$
$\mathcal{E}(\lambda)$	$\lambda \exp(-\lambda t)\mathbb{1}_{[0,+\infty[}(t)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{1-\frac{it}{\lambda}}$
$\Gamma(\alpha,\beta)$	$\frac{\beta^\alpha}{\Gamma(\alpha)} \mathrm{e}^{-\beta t} t^{\alpha-1}\mathbb{1}_{[0,+\infty[}(t)$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(1-\frac{it}{\beta}\right)^{-\alpha}$

QUESTIONS

Q Donner une variable aléatoire sans moment.

R Une variable aléatoire de loi de CAUCHY.

Q Quelles sont les applications de l'inégalité de MARKOV ?

R De nombreuses autres inégalités classiques, le fait que la convergence L^1 implique la convergence en probabilité. Aussi, si X centrée admet un moment d'ordre p, alors $\mathbb{P}(X \geq x) = \mathbb{P}(X^p \geq x) \leq \mathbb{E}[X^p]/x^p$, et donc plus X admets des moments d'ordres élevés, plus on a une décroissance rapide de la queue de distribution.

Q Quel est l'intérêt de regarder la matrice de covariance d'un vecteur gaussien ?

R Les coordonnées du vecteur sont indépendantes si et seulement si la matrice de covariance est diagonale (le sens réciproque n'est vrai que pour des vecteurs gaussiens).

Q Démontrer le théorème central limite en dimension d ?

R On se ramène au cas réel à partir du théorème de LÉVY.

Q Soit $\theta > 0$ et $(X_k)_{k\geq 1}$ i.i.d. de loi $\mathcal{U}([0,\theta])$. On pose $\hat{\theta}_n = \max_{1\leq i\leq n} X_k$. Calculer la fonction de répartition, la densité, l'espérance et la variance de $\hat{\theta}_n$.

Q Soit $X \in L^2$ d'espérance m, de variance σ^2. Montrer que $\mathbb{P}(X - m \geq \alpha) \leq \frac{\sigma^2+\lambda^2}{(\alpha+\lambda)^2}$ pour tout $\alpha, \lambda > 0$.

R $\mathbb{P}(X - m + \lambda \geq \alpha + \lambda) \leq \mathbb{P}((X - m + \lambda)^2 \geq (\alpha + \lambda)^2)$ puis on applique l'inégalité de MARKOV et on développe l'espérance.

Q Dans le même contexte, montrer que $\mathbb{P}(X - m \geq \alpha) \leq \frac{\sigma^2}{\alpha^2+\sigma^2}$ pour tout $\alpha > 0$.

R On minimise la fonction précédemment obtenue en $\lambda = \sigma^2/\alpha$.

Q Montrer que $\mathbb{P}(|X - m| \geq \alpha) \leq 2\frac{\sigma^2}{\alpha^2+\sigma^2}$. Comparer ce résultat avec l'inégalité de BIENAYMÉ-TCHEBYCHEV.

R Pour $\sigma \geq \alpha$, cette inégalité est meilleure que l'inégalité de BIENAYMÉ-TCHEBYCHEV.

BIBLIOGRAPHIE

[BL07] P. BARBE et M. LEDOUX : *Probabilité*. EDP Sciences, 2007.

[Mé11] S. MÉLÉARD : *Aléatoire : introduction à la théorie et au calcul des probabilités*. Les éditions de l'École Polytechnique, 2011.

[RS12] V. RIVOIRARD et G. STOLTZ : *Statistique mathématique en action*. Vuibert, 2 ème édition, 2012.