Inégalités en analyse et en probabilités

Corrigé partiel des exercices

Exercice 6 (Covariance et corrélation)**.**

1. On vérifie facilement que $\langle X,Y\rangle = \operatorname{cov}(X,Y)$ définit un forme bilinéaire symétrique positive sur l'espace vectoriel des variables aléatoires $X : \Omega \to \mathbb{R}$. Toutefois, on a

$$\langle X,X\rangle = 0 \quad \Leftrightarrow \quad \mathbb{E}\left([X-\mathbb{E}(X)]^2\right) = 0 \quad \Leftrightarrow \quad \exists c \in \mathbb{R} : \mathbb{P}\{X = c\} = 1 \;.$$

   La covariance n'est donc pas une forme définie en général.
   Toutefois $\langle X,Y\rangle = \operatorname{cov}(X,Y)$ est un produit scalaire si l'on se restreint aux variables aléatoires centrées (d'espérance nulle), et que l'on identifie à 0 toutes les variables aléatoires valant 0 avec probabilité 1.
2. La preuve de l'inégalité de Cauchy–Schwarz donnée dans l'exercice 1 n'utilise en fait pas le caractère défini du produit scalaire (celui-ci n'intervient que dans la caractérisation de l'égalité). On peut donc conclure que

$$|\operatorname{cov}(X,Y)|^2 \leqslant \operatorname{cov}(X,X)\operatorname{cov}(Y,Y) = \operatorname{Var}(X)\operatorname{Var}(Y) \;.$$

3. Par définition du coefficient de corrélation et l'inégalité ci-dessus, on a

$$\rho_{X,Y} \in [-1,1] \;.$$

4. On a $|\rho_{X,Y}| = 1$ si et seulement si X et Y sont liés dans le sens qu'il existe $a,b,c \in \mathbb{R}$ tels que

$$\mathbb{P}\{aX + bY = c\} = 1 \;.$$

   C'est encore une fois une conséquence de la preuve de l'exercice 1: l'égalité

$$|\operatorname{cov}(X,Y)|^2 = \operatorname{Var}(X)\operatorname{Var}(Y)$$

   a lieu si et seulement s'il existe $a,b \in \mathbb{R}$ tels que $\operatorname{Var}(aX+bY) = 0$, ce qui équivaut à la condition ci-dessus.

Exercice 7 (Inégalité de Bienaymé–Tchebychev)**.**

1. Une première méthode consiste à observer que

$$\mathbb{E}(X)^2 \leqslant \mathbb{E}(X^2) < \infty$$

   en vertu de l'inégalité de Jensen, appliquée à la fonction convexe $\varphi(t) = t^2$.
   Une seconde possibilité consiste à appliquer l'inégalité de Cauchy–Schwarz à la forme bilinéaire symétrique $\langle X,Y\rangle = \mathbb{E}(XY)$ pour conclure que

$$|\mathbb{E}(X)| = |\mathbb{E}(X\cdot 1)| \leqslant \sqrt{\mathbb{E}(X^2)\mathbb{E}(1)} = \sqrt{\mathbb{E}(X^2)} < \infty \;,$$

   où 1 désigne la variable aléatoire constante égale à 1.
2. Il suffit d'appliquer l'inégalité de Markov pour obtenir

$$\mathbb{P}\left\{|X-\mathbb{E}(X)| > a\right\} = \mathbb{P}\left\{(X-\mathbb{E}(X))^2 > a^2\right\} \leqslant \frac{1}{a^2}\mathbb{E}\left((X-\mathbb{E}(X))^2\right) = \frac{1}{a^2}\operatorname{Var}(X) \;.$$

Exercice 8 (La loi faible des grands nombres)**.**

1. Par linéarité de l'espérance, on a

$$\mathbb{E}\left(S_n\right) = n\mu \ .$$

De plus, comme la variance d'une somme de variables aléatoires *indépendantes* est égale à la somme de leurs variances, on a

$$\mathrm{Var}(S_n) = n\sigma^2 \ .$$

2. L'inégalité de Bienaymé–Tchebychev implique

$$\mathbb{P}\left\{\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right\} = \mathbb{P}\left\{|S_n - \mathbb{E}\left(S_n\right)| > n\varepsilon\right\} \leqslant \frac{1}{n^2\varepsilon^2}\,\mathrm{Var}(S_n) = \frac{\sigma^2}{n\varepsilon^2} \ .$$

Lorsque $\varepsilon > 0$ est fixé, $\sigma^2/(n\varepsilon^2)$ tend bien vers 0 lorsque $n \to \infty$.

Exercice 9 (Equivalence entre inégalités de Cauchy–Schwarz et de Bessel*)**.**

1. L'inégalité de Bessel pour $n = 1$ s'écrit

$$|\langle x, e_1\rangle|^2 \leqslant \langle x, x\rangle$$

pour tout vecteur e_1 tel que $\langle e_1, e_1\rangle = 1$. Si $y \neq 0$, l'inégalité appliquée au vecteur $e_1 = y/\sqrt{\langle y, y\rangle}$ implique l'inégalité de Cauchy–Schwarz. Si $y = 0$, l'inégalité est clairement satisfaite.
De plus, on a égalité si et seulement si x est colinéaire à e_1, ce qui est vrai si et seulement si x est colinéaire à y.

2. En supposant vraie l'inégalité de Cauchy–Schwarz, on a

$$\begin{aligned}
\sum_{i=1}^{n}|\langle x, e_i\rangle|^2 &= \sum_{i=1}^{n}\Big|\langle x, e_i\rangle - \sum_{j\neq i}\langle x, e_j\rangle \underbrace{\langle e_j, e_i\rangle}_{=0}\Big|^2 \\
&= \sum_{i=1}^{n}\Big|\Big\langle x - \sum_{j\neq i}\overline{\langle x, e_j\rangle}e_j, e_i\Big\rangle\Big|^2 \\
&\leqslant \sum_{i=1}^{n}\Big\langle x - \sum_{j\neq i}\overline{\langle x, e_j\rangle}e_j, x - \sum_{k\neq i}\overline{\langle x, e_k\rangle}e_k\Big\rangle \underbrace{\langle e_i, e_i\rangle}_{=1} \\
&= \sum_{i=1}^{n}\Big[\langle x, x\rangle - \sum_{j\neq i}|\langle x, e_j\rangle|^2 - \sum_{k\neq i}|\langle x, e_k\rangle|^2 + \sum_{j,k\neq i}\langle x, e_j\rangle\overline{\langle x, e_k\rangle}\underbrace{\langle e_j, e_k\rangle}_{=\delta_{jk}}\Big] \\
&= n\langle x, x\rangle - \sum_{i=1}^{n}\sum_{j\neq i}|\langle x, e_j\rangle|^2 \\
&= n\langle x, x\rangle - (n-1)\sum_{i=1}^{n}|\langle x, e_i\rangle|^2 \ .
\end{aligned}$$

En ajoutant $(n-1)\sum_{i=1}^{n}|\langle x, e_i\rangle|^2$ des deux côtés et en divisant par n, on obtient l'inégalité de Bessel.

De plus, on a égalité si et seulement si $x - \sum_{j \neq i} \overline{\langle x, e_j \rangle} e_j$ est colinéaire à e_i pour tout i. Dans ce cas il existe des scalaires α_i tels que

$$x = \alpha_i e_i + \sum_{j \neq i} \langle e_j, x \rangle e_j$$

pour tout i. En prenant le produit scalaire avec e_i, on voit que $\alpha_i = \langle e_i, x \rangle$. Par conséquent, on a bien égalité si et seulement si

$$x = \sum_{i=1}^{n} \langle e_i, x \rangle e_i \ .$$

Exercice 10 (Intégration par parties*)**.**

1. Considérons tout d'abord le cas où X admet une densité f. Notons

$$R(x) = \mathbb{P}\{X > x\} = \int_x^\infty f(t) \, \mathrm{d}t \ .$$

On a $R'(x) = -f(x)$ et $\lim_{x \to \infty} R(x) = 0$. Par conséquent, une intégration par parties nous fournit

$$\mathbb{E}(\varphi(X)) = \lim_{b \to \infty} \int_0^b \varphi(x) f(x) \, \mathrm{d}x = \lim_{b \to \infty} \Big[-\varphi(x) R(x) \Big]_0^b + \lim_{b \to \infty} \int_0^b \varphi'(x) R(x) \, \mathrm{d}x \ .$$

Le terme de bord $\varphi(0)R(0)$ est nul par hypothèse. Par ailleurs, pour tout $b > 0$, nous avons

$$\varphi(b) R(b) = \int_b^\infty \varphi(b) f(x) \, \mathrm{d}x \leqslant \int_b^\infty \varphi(x) f(x) \, \mathrm{d}x \leqslant \mathbb{E}(\varphi(X)) < \infty \ .$$

Par conséquent, le terme de bord $\lim_{b \to \infty} \varphi(b) R(b)$ est également nul, et le résultat suit.
Dans le cas général, on peut utiliser la théorème de Fubini justifiant l'échange des ordres d'intégration pour écrire

$$\begin{aligned}
\mathbb{E}(\varphi(X)) &= \int_\Omega \varphi(X(\omega)) \, \mathrm{d}\mathbb{P}(\omega) \\
&= \int_\Omega \int_0^{X(\omega)} \varphi'(t) \, \mathrm{d}t \, \mathrm{d}\mathbb{P}(\omega) \\
&= \int_\Omega \int_0^\infty 1_{\{t < X(\omega)\}} \varphi'(t) \, \mathrm{d}t \, \mathrm{d}\mathbb{P}(\omega) \\
&= \int_0^\infty \int_\Omega 1_{\{X(\omega) > t\}} \varphi'(t) \, \mathrm{d}\mathbb{P}(\omega) \, \mathrm{d}t \\
&= \int_0^\infty \mathbb{P}\{X > t\} \varphi'(t) \, \mathrm{d}t \ .
\end{aligned}$$

2. En découpant l'intégrale en $t = \mathbb{E}(X)$ puis en utilisant l'inégalité de Markov, on a

$$\begin{aligned}
\mathbb{E}(\varphi(X)) &= \int_0^{\mathbb{E}(X)} \varphi'(t) \mathbb{P}\{X > t\} \, \mathrm{d}t + \int_{\mathbb{E}(X)}^\infty \varphi'(t) \mathbb{P}\{X > t\} \, \mathrm{d}t \\
&\leqslant \int_0^{\mathbb{E}(X)} \varphi'(t) \, \mathrm{d}t + \int_{\mathbb{E}(X)}^\infty \varphi'(t) \frac{1}{t} \mathbb{E}(X) \, \mathrm{d}t \\
&= \varphi(\mathbb{E}(X)) + \mathbb{E}(X) \int_{\mathbb{E}(X)}^\infty \frac{\varphi'(t)}{t} \, \mathrm{d}t \ .
\end{aligned}$$

3. Pour $\varphi(t) = t^p$ on obtient

$$\mathbb{E}\left(X^p\right) \leqslant \mathbb{E}\left(X\right)^p + \mathbb{E}\left(X\right) \int_{\mathbb{E}\,(X)}^{\infty} p t^{p-2} \, \mathrm{d}t \; .$$

L'intégrale converge si et seulement si $p < 1$, et dans ce cas on obtient

$$\mathbb{E}\left(X^p\right) \leqslant \frac{\mathbb{E}\left(X\right)^p}{1-p} \; .$$

Si $0 < p < 1$, la fonction $t \mapsto t^p$ est concave, et l'inégalité de Jensen nous fournit $\mathbb{E}\left(X^p\right) \leqslant \mathbb{E}\left(X\right)^p$. L'inégalité de Jensen est donc plus précise dans ce cas. Toutefois la majoration obtenue ci-dessus par l'inégalité de Markov ne requiert pas la concavité, elle est donc plus générale.

Exercice 11 (Inégalité de type Chernoff–Cramér pour la loi binomiale*)**.**

1. La somme S_n suit une loi binomiale d'espérance $\frac{n}{2}$ et de variance $\frac{n}{4}$.
2. On a

$$\mathbb{E}\left(Z_i\right) = \sum_{n=0}^{1} \mathbb{P}\left\{X_i = n\right\} Z_i(n) = \frac{1}{2}\,\mathrm{e}^{-\lambda/2} + \frac{1}{2}\,\mathrm{e}^{\lambda/2} = \cosh\biggl(\frac{\lambda}{2}\biggr) \; .$$

3. On a

$$\mathbb{E}\left(\mathrm{e}^{\lambda[S_n - \mathbb{E}\,(S_n)]}\right) = \mathbb{E}\left(\prod_{i=1}^{n} Z_i\right) = \cosh\biggl(\frac{\lambda}{2}\biggr)^n \; .$$

4. On peut utiliser l'inégalité de Markov en écrivant

$$\begin{aligned}\mathbb{P}\bigl\{S_n - \mathbb{E}\left(S_n\right) > nt\bigr\} &= \mathbb{P}\left\{\mathrm{e}^{\lambda[S_n - \mathbb{E}\,(S_n)]} > \mathrm{e}^{\lambda n t}\right\} \\ &\leqslant \mathrm{e}^{-\lambda n t}\, \mathbb{E}\left(\mathrm{e}^{\lambda[S_n - \mathbb{E}\,(S_n)]}\right) \\ &= \mathrm{e}^{-\lambda n t} \cosh\biggl(\frac{\lambda}{2}\biggr)^n = \mathrm{e}^{-n f_t(\lambda)} \; ,\end{aligned}$$

où nous avons posé

$$f_t(\lambda) = \lambda t - \log\biggl(\cosh\biggl(\frac{\lambda}{2}\biggr)\biggr) \; .$$

5. Comme la dérivée

$$f_t'(\lambda) = t - \frac{1}{2}\tanh\biggl(\frac{\lambda}{2}\biggr)$$

est positive pour $\lambda \to -\infty$ et négative pour $\lambda \to \infty$ en ne changeant de signe qu'une seule fois, la fonction $\lambda \mapsto f_t(\lambda)$ admet un unique maximum, en λ^* solution de

$$\tanh\biggl(\frac{\lambda^*}{2}\biggr) = 2t \; .$$

On en déduit

$$\lambda^* = \log\biggl(\frac{1+2t}{1-2t}\biggr)$$

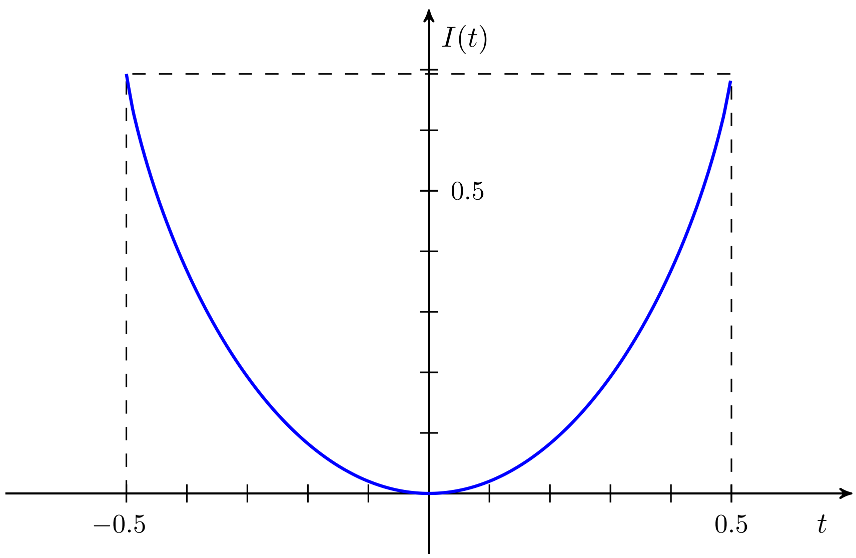
et

$$\begin{aligned}I(t) = f_t(\lambda^*) &= t\log\biggl(\frac{1+2t}{1-2t}\biggr) - \log\frac{1}{\sqrt{(1-2t)(1+2t)}} \\ &= \biggl(\frac{1}{2}+t\biggr)\log(1+2t) + \biggl(\frac{1}{2}-t\biggr)\log(1-2t) \; .\end{aligned}$$

La fonction $t \mapsto I(t)$ est définie sur l'intervalle $[-\frac{1}{2}, \frac{1}{2}]$, paire, et admet le développement limité en 0

$$I(t) = 2t^2 + \mathcal{O}(t^4) .$$

De plus, on vérifie qu'elle est croissante sur $[0, \frac{1}{2}]$ et admet une tangente verticale en $t = \frac{1}{2}$.



6. La probabilité cherchée est

$$\mathbb{P}\{S_{1000} > 600\} = \mathbb{P}\{S_{1000} - 500 > 0.1 \cdot 1000\} \leqslant \mathrm{e}^{-1000 \cdot I(0.1)}$$

On trouve $I(0.1) \cong 0.0201$ et par conséquent

$$\mathbb{P}\{S_{1000} > 600\} \leqslant \mathrm{e}^{-20.1} = 1.8 \cdot 10^{-9} .$$

L'inégalité de Bienaymé–Tchebychev fournit l'estimation beaucoup moins bonne

$$\begin{aligned} \mathbb{P}\{S_{1000} > 600 \text{ ou } S_{1000} < 400\} &= \mathbb{P}\left\{\left|\frac{S_{1000}}{1000} - \frac{1}{2}\right| > 0.1\right\} \\ &\leqslant \frac{1}{0.1^2} \operatorname{Var}\left(\frac{S_{1000}}{1000}\right) \\ &= \frac{1}{40} = 0.025 . \end{aligned}$$