

Feuille d'exercices n°63

Exercice 1 (**)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ espace probabilisé et X_1, \dots, X_n variables aléatoires indépendantes de loi $\mathcal{U}_{\{-1,1\}}$.

1. Montrer $\forall (t, \varepsilon) \in]0; +\infty[^2 \quad \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon\right) \leq e^{-t\varepsilon} \operatorname{ch}(t)^n$

2. Pour $t > 0$, comparer $e^{\frac{t^2}{2}}$ et $\operatorname{ch}(t)$.

3. En déduire $\forall \varepsilon > 0 \quad \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon\right) \leq e^{-n\varepsilon^2/2}$

Corrigé : 1. Soit $t, \varepsilon > 0$. Par croissance stricte de $u \mapsto e^{tu}$, on a

$$\left\{ \frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon \right\} = \left\{ \sum_{i=1}^n X_i \geq n\varepsilon \right\} = \left\{ \exp\left(t \sum_{i=1}^n X_i\right) \geq e^{tn\varepsilon} \right\}$$

D'après l'inégalité de Markov appliquée à la variable aléatoire finie positive $\exp\left(t \sum_{i=1}^n X_i\right)$, il vient

$$\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon\right) \leq e^{-tn\varepsilon} \mathbb{E}\left(\exp\left(t \sum_{i=1}^n X_i\right)\right)$$

Or
$$\mathbb{E}\left(\exp\left(t \sum_{i=1}^n X_i\right)\right) = \mathbb{E}\left(\prod_{i=1}^n e^{tX_i}\right)$$

et par indépendance puis égalité en loi des X_i , il vient

$$\mathbb{E}\left(\prod_{i=1}^n e^{tX_i}\right) = \prod_{i=1}^n \mathbb{E}(e^{tX_i}) = (\mathbb{E}(e^{tX_1}))^n$$

Par transfert
$$\mathbb{E}(e^{tX_1}) = e^t \mathbb{P}(X_1 = 1) + e^{-t} \mathbb{P}(X_1 = -1) = \operatorname{ch} t$$

On conclut
$$\forall (t, \varepsilon) \in]0; +\infty[^2 \quad \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon\right) \leq e^{-tn\varepsilon} \operatorname{ch}^n t$$

2. D'après les développements en série entière usuels, on a

$$\forall t \in \mathbb{R} \quad \operatorname{ch}(t) = \sum_{n=0}^{+\infty} \frac{t^{2n}}{(2n)!} \quad \text{et} \quad e^{\frac{t^2}{2}} = \sum_{n=0}^{+\infty} \frac{t^{2n}}{2^n n!}$$

Une récurrence immédiate permet de prouver que $2^n n! \leq (2n)!$ pour tout n entier d'où

$$\forall t \in \mathbb{R} \quad \operatorname{ch}(t) \leq e^{\frac{t^2}{2}}$$

Variante : On peut éviter une récurrence avec

$$(2n)! = \prod_{k=1}^{2n} k = \prod_{k=1}^n (2k) \times \prod_{k=1}^n (2k-1) = 2^n n! \prod_{k=1}^n (2k-1)$$

3. Soient $t, \varepsilon > 0$. On obtient
$$\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon\right) \leq e^{-tn\varepsilon + n\frac{t^2}{2}}$$

En choisissant $t = \varepsilon$, il vient

$$\forall \varepsilon > 0 \quad \mathbb{P} \left(\frac{1}{n} \sum_{i=1}^n X_i \geq \varepsilon \right) \leq e^{-n\varepsilon^2/2}$$

Remarque : Le choix $t = \varepsilon$ est optimal sur la dernière égalité puisqu'il s'agit de la valeur de t qui minimise le trinôme dans l'exponentielle.

Exercice 2 (**)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de loi $\mathcal{B}(p)$ avec $p \in]0; 1[$ et $Y_n = X_n + X_{n+1} + X_{n+2}$ pour tout $n \geq 1$. Montrer

$$\forall \varepsilon > 0 \quad \mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n Y_k - 3p \right| \geq \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0$$

Corrigé : Les Y_n ne sont pas indépendantes ! En effet, on a

$$\forall n \geq 1 \quad \mathbb{P}(Y_n = 3, Y_{n+1} = 0) = 0 \neq \mathbb{P}(Y_n = 3) \times \mathbb{P}(Y_{n+1} = 0)$$

Soit $\varepsilon > 0$ et $n \geq 1$. On a

$$\left\{ \left| \frac{1}{n} \sum_{k=1}^n Y_k - 3p \right| \geq \varepsilon \right\} \subset \bigcup_{\ell=0}^2 \left\{ \left| \frac{1}{n} \sum_{k=1}^n X_{k+\ell} - p \right| \geq \frac{\varepsilon}{3} \right\}$$

D'où

$$\mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n Y_k - 3p \right| \geq \varepsilon \right) \leq \sum_{\ell=0}^2 \mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n X_{k+\ell} - p \right| \geq \frac{\varepsilon}{3} \right)$$

Pour $\ell \in \llbracket 0; 2 \rrbracket$, les variables $(X_{k+\ell})_k$ sont indépendantes, de même loi, avec des moments d'ordre deux. D'après la loi faible des grands nombres, il s'ensuit

$$\forall \varepsilon > 0 \quad \mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n Y_k - 3p \right| \geq \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0$$

Exercice 3 (**)

Une urne contient n billes numérotées de 1 à n . On saisit une poignée de billes et on note X la somme des numéros des billes. En supposant que toutes les poignées sont équiprobables, que vaut $\mathbb{E}(X)$?

Corrigé : Soit $U \sim \mathcal{U}_{\mathcal{P}(\llbracket 1; n \rrbracket)}$ et $X = \sum_{x \in U} x$. On peut écrire

$$X = \sum_{k=1}^n k \mathbf{1}_{k \in U}$$

Par suite

$$\mathbb{E}(X) = \sum_{k=1}^n k \mathbb{P}(k \in U)$$

et

$$\forall k \in \llbracket 1; n \rrbracket \quad \mathbb{P}(k \in U) = \frac{\text{Card } \mathcal{P}(\llbracket 1; n \rrbracket \setminus \{k\})}{\text{Card } \mathcal{P}(\llbracket 1; n \rrbracket)} = \frac{2^{n-1}}{2^n} = \frac{1}{2}$$

On conclut

$$\mathbb{E}(X) = \sum_{k=1}^n \frac{k}{2} = \frac{n(n+1)}{4}$$

Variante : On peut aussi considérer (X_1, \dots, X_n) des variables i.i.d de loi $\mathcal{B}(1/2)$ avec le formalisme suivant : X_i vaut 1 si i pioché et 0 sinon. On a

$$U = \{i \in \llbracket 1; n \rrbracket \mid X_i = 1\} \sim \mathcal{U}_{\mathcal{P}(\llbracket 1; n \rrbracket)}$$

puisque $\forall A \in \mathcal{P}(\llbracket 1; n \rrbracket) \quad \mathbb{P}(U = A) = \mathbb{P}\left(\bigcap_{i \in A} \{X_i = 1\} \cap \bigcap_{i \in \bar{A}} \{X_i = 0\}\right) = \frac{1}{2^n}$

Puis
$$X = \sum_{k=1}^n kX_k$$

et on retrouve
$$\mathbb{E}(X) = \sum_{k=1}^n k\mathbb{E}(X_k) = \frac{n(n+1)}{4}$$

Exercice 4 (***)

Soit $p \in]0; 1[$ et $q = 1 - p$. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de même loi de Bernoulli de paramètre p . On pose

$$\forall n \geq 1 \quad S_n = \sum_{i=1}^n X_i \quad T_n = \frac{S_n - np}{\sqrt{npq}}$$

1. Donner la loi et la fonction génératrice de S_n .
2. Déterminer $\mathbb{E}(T_n)$, $\mathbb{V}(T_n)$.
3. Calculer $\mathbb{E}(x^{T_n})$ pour $x > 0$ puis $\lim_{n \rightarrow +\infty} \mathbb{E}(x^{T_n})$.

Corrigé : 1. Soit n entier non nul. La variable aléatoire S_n est somme de n variables indépendantes de loi de Bernoulli $\mathcal{B}(p)$ d'où

$$S_n \sim \mathcal{B}(n, p) \quad \text{et} \quad \forall t \in \mathbb{R} \quad G_{S_n}(t) = (pt + 1 - p)^n$$

2. Soit n entier non nul. On a par linéarité de l'espérance

$$\mathbb{E}(T_n) = \frac{1}{\sqrt{npq}} \sum_{i=1}^n \mathbb{E}(X_i - p) = 0$$

Puis, par indépendance des X_i , il vient

$$\mathbb{V}(T_n) = \frac{1}{npq} \mathbb{V}\left(\sum_{i=1}^n X_i\right) = \frac{1}{npq} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{npq}{npq}$$

Ainsi

$$\forall n \in \mathbb{N}^* \quad \mathbb{E}(T_n) = 0 \quad \text{et} \quad \mathbb{V}(T_n) = 1$$

Remarque : La variable aléatoire T_n est dite *centrée réduite*.

3. Soit $x > 0$. Par indépendance des X_i , on trouve

$$\mathbb{E}(x^{T_n}) = \mathbb{E}\left(x^{\frac{1}{\sqrt{npq}} \sum_{i=1}^n (X_i - p)}\right) = \mathbb{E}\left(\prod_{i=1}^n x^{\frac{X_i - p}{\sqrt{npq}}}\right) = \prod_{i=1}^n \mathbb{E}\left(x^{\frac{X_i - p}{\sqrt{npq}}}\right)$$

Puis par transfert, on obtient

$$\mathbb{E}(x^{T_n}) = \left(px^{\sqrt{\frac{q}{np}}} + qx^{-\sqrt{\frac{p}{nq}}}\right)^n = \left[p \exp\left(\sqrt{\frac{q}{np}} \ln(x)\right) + q \exp\left(-\sqrt{\frac{p}{nq}} \ln(x)\right)\right]^n$$

Avec un développement limité à l'ordre deux, il vient

$$\begin{aligned}\mathbb{E}(x^{T_n}) &= \left[p \left(1 + \sqrt{\frac{q}{np}} \ln(x) + \frac{q}{2np} \ln(x)^2 \right) + q \left(1 - \sqrt{\frac{p}{nq}} \ln(x) + \frac{p}{2nq} \ln(x)^2 \right) + o\left(\frac{1}{n}\right) \right]^n \\ &= \left[\frac{p+q}{2n} \ln(x)^2 + o\left(\frac{1}{n}\right) \right]^n = \exp \left[n \ln \left(1 + \frac{\ln(x)^2}{2n} + o\left(\frac{1}{n}\right) \right) \right] = \exp \left[\frac{\ln(x)^2}{2} + o(1) \right]\end{aligned}$$

On conclut

$$\forall x > 0 \quad \mathbb{E}(x^{T_n}) \xrightarrow[n \rightarrow \infty]{} e^{\frac{\ln(x)^2}{2}}$$

Remarque : Ce résultat est une conséquence du *théorème de la limite centrée* ou de son corollaire qu'est le théorème de Moivre-Laplace.

Exercice 5 (***)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de même loi $\mathcal{B}(1/2)$. On note $S_n = \sum_{i=1}^n X_i$ pour n entier non nul.

1. Déterminer une expression sommatoire de $u_n = \mathbb{P}\left(S_n < \frac{n}{2}\right)$.
2. Pour $n \geq 1$, on note $v_n = \mathbb{P}\left(S_n \geq \frac{n}{2}\right)$. Établir l'égalité

$$\forall n \geq 1 \quad v_n = u_n + \mathbb{P}\left(S_n = \frac{n}{2}\right)$$

3. En déduire le comportement asymptotique de u_n pour $n \rightarrow +\infty$.

Corrigé : 1. On a $\mathbb{P}\left(S_n < \frac{n}{2}\right) = \mathbb{P}\left(\bigsqcup_{0 \leq k < \frac{n}{2}} \{S_n = k\}\right) = \sum_{0 \leq k < \frac{n}{2}} \mathbb{P}(S_n = k)$

D'où

$$\mathbb{P}\left(S_n < \frac{n}{2}\right) = \frac{1}{2^n} \sum_{0 \leq k < \frac{n}{2}} \binom{n}{k}$$

2. On a $v_n = \mathbb{P}\left(\bigsqcup_{\frac{n}{2} \leq k \leq n} \{S_n = k\}\right) = \sum_{\frac{n}{2} \leq k \leq n} \mathbb{P}(S_n = k) = \frac{1}{2^n} \sum_{\frac{n}{2} \leq k \leq n} \binom{n}{k}$

Avec le changement d'indice $k \longleftrightarrow n - k$, on trouve

$$v_n = \frac{1}{2^n} \sum_{0 \leq k \leq \frac{n}{2}} \binom{n}{n-k} = \sum_{0 \leq k \leq \frac{n}{2}} \mathbb{P}(S_n = k) = \mathbb{P}\left(S_n < \frac{n}{2}\right) + \mathbb{P}\left(S_n = \frac{n}{2}\right)$$

Autrement dit

$$v_n = u_n + \mathbb{P}\left(S_n = \frac{n}{2}\right)$$

3. Si n est impair, alors $\mathbb{P}\left(S_n = \frac{n}{2}\right) = 0$.

Supposons $n = 2m$. On a

$$\mathbb{P}(S_{2m} = m) = \frac{\binom{2m}{m}}{2^{2m}} = \frac{(2m)!}{2^{2m}(m!)^2}$$

Avec l'équivalent de Stirling, on obtient

$$\mathbb{P}(S_{2m} = m) \underset{m \rightarrow +\infty}{\sim} \frac{1}{2^{2m}} \left(\frac{e}{m}\right)^{2m} \frac{1}{2\pi m} \left(\frac{2m}{e}\right)^{2m} 2\sqrt{\pi m} \underset{m \rightarrow +\infty}{\sim} \frac{1}{\sqrt{\pi m}}$$

Par conséquent, indépendamment de la parité de n , on a

$$u_n = v_n + o(1)$$

Enfin, les événements $\left\{S_n < \frac{n}{2}\right\}$ et $\left\{S_n \geq \frac{n}{2}\right\}$ sont complémentaires donc u_n et v_n vérifient le système

$$\begin{cases} u_n + v_n = 1 \\ u_n - v_n = o(1) \end{cases}$$

Ainsi

$$u_n = \frac{1}{2} + o(1) \xrightarrow{n \rightarrow \infty} \frac{1}{2}$$

Exercice 6 (***)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire de loi $\mathcal{U}_{[0; n-1]}$ avec n entier. On suppose n non premier avec $n = ab$ et a, b des entiers supérieurs ou égaux à 2.

1. Montrer qu'il existe un unique couple de variables aléatoires (Q, R) à valeurs dans \mathbb{N} tel que $X = aQ + R$ avec $R(\Omega) \subset [0; a-1]$.
2. Préciser la loi de (Q, R) puis de Q et de R .
3. En déduire qu'il existe Y et Z , variables aléatoires indépendantes à valeurs dans \mathbb{N} dont on précisera les lois telles que $X \sim Y + Z$.

Corrigé : 1. Soit $\omega \in \Omega$. D'après le théorème de la division euclidienne il existe un unique couple $(q, r) \in \mathbb{N} \times [0; a-1]$ tel que $X(\omega) = aq + r$. Ceci prouve l'unicité du couple de variables aléatoires solutions sous réserve d'existence. On définit l'application $\varphi : \mathbb{N} \rightarrow \mathbb{N} \times [0; a-1], x \mapsto (q, r)$ qui à x associe son couple quotient-reste. Cette application est bien définie et bijective d'après le théorème de la division euclidienne et on pose $(Q, R) = \varphi(X)$. Il s'agit bien d'une variable aléatoire discrète en tant que fonction d'une variable aléatoire discrète et par construction, on a

$$X = aQ + R \text{ avec } R(\Omega) \subset [0; a-1]$$

Remarque : On peut expliciter Q et R : on a $Q = \left\lfloor \frac{X}{a} \right\rfloor$ et $R = X - aQ$.

2. Soit $(q, r) \in [0; b-1] \times [0; a-1]$. D'après le théorème de la division euclidienne, l'application φ précédemment définie est une bijection et par suite

$$\mathbb{P}((Q, R) = (q, r)) = \mathbb{P}(\varphi(X) = \varphi(x)) = \mathbb{P}(X = x) = \frac{1}{n}$$

On détermine les lois marginales avec

$$\mathbb{P}(Q = q) = \sum_{r=0}^{a-1} \mathbb{P}((Q, R) = (q, r)) = \frac{a}{n} = \frac{1}{b} \quad \text{et} \quad \mathbb{P}(R = r) = \sum_{q=0}^{b-1} \mathbb{P}((Q, R) = (q, r)) = \frac{b}{n} = \frac{1}{a}$$

Ainsi

$$(Q, R) \sim \mathcal{U}_{[0; a-1] \times [0; b-1]} \quad Q \sim \mathcal{U}_{[0; b-1]} \quad R \sim \mathcal{U}_{[0; a-1]}$$

3. Soit $(q, r) \in [0; b-1] \times [0; a-1]$. On a

$$\mathbb{P}(Q = q, R = r) = \frac{1}{n} = \frac{1}{b} \times \frac{1}{a} = \mathbb{P}(Q = q)\mathbb{P}(R = r)$$

ce qui prouve que les variables Q et R sont indépendantes et on conclut

$$X = aQ + R \text{ avec } aQ \text{ et } R \text{ indépendantes et } aQ \sim \mathcal{U}_{a[0; b-1]}, R \sim \mathcal{U}_{[0; a-1]}.$$

Exercice 7 (***)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de même loi $\mathcal{B}(x)$ avec $x \in [0; 1]$. Soit $f : [0; 1] \rightarrow \mathbb{R}$ continue par morceaux. Pour n entier non nul, on note $S_n = \sum_{i=1}^n X_i$ et on pose

$$\forall x \in [0; 1] \quad B_n(f)(x) = \mathbb{E} \left(f \left(\frac{S_n}{n} \right) \right)$$

1. Préciser la loi de S_n puis déterminer une expression sommatoire de $\mathbb{E} \left[f \left(\frac{S_n}{n} \right) \right]$.
2. Si x est un point de continuité de f , montrer

$$B_n(f)(x) \xrightarrow[n \rightarrow \infty]{} f(x)$$

Corrigé : 1. D'après le cours et par transfert, on a

$$S_n \sim \mathcal{B}(n, x) \quad \text{et} \quad \mathbb{E} \left(f \left(\frac{S_n}{n} \right) \right) = \sum_{k=0}^n f \left(\frac{k}{n} \right) \mathbb{P}(S_n = k) = \sum_{k=0}^n \binom{n}{k} f \left(\frac{k}{n} \right) x^k (1-x)^{n-k}$$

2. Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que

$$\forall t \in [0; 1] \quad |x - t| \leq \eta \implies |f(x) - f(t)| \leq \varepsilon$$

On a $|B_n(f)(x) - f(x)| = \left| \mathbb{E} \left(f \left(\frac{S_n}{n} \right) - f(x) \right) \right| \leq \mathbb{E} \left(\left| f \left(\frac{S_n}{n} \right) - f(x) \right| \right)$

Posons $A_n = \left\{ \left| \frac{S_n}{n} - x \right| > \eta \right\}$. Ainsi

$$\begin{aligned} |B_n(f)(x) - f(x)| &\leq \mathbb{E} \left(\left| f \left(\frac{S_n}{n} \right) - f(x) \right| \mathbb{1}_{A_n} \right) + \mathbb{E} \left(\left| f \left(\frac{S_n}{n} \right) - f(x) \right| \mathbb{1}_{A_n^c} \right) \\ &\leq 2\|f\|_\infty \mathbb{P}(A_n) + \varepsilon \end{aligned}$$

D'après la loi faible des grands nombres, on a $\mathbb{P}(A_n) = o(1)$ et par conséquent

$$B_n(f)(x) \xrightarrow[n \rightarrow \infty]{} f(x)$$