3

PROBABILITÉS

Exercice 3.1.

On cherche à évaluer le nombre N de lions d'Asie, espèce en voie de disparition, encore en vie dans la forêt de Gir. Pour cela, on capture d'abord, en une seule fois, m lions (avec $m \in \mathbb{N}^*$) que l'on tatoue avant de les relâcher dans la nature, et on admet que pendant toute la durée de l'étude il n'y a ni décès ni naissance, puis on utilise l'une des deux méthodes suivantes ...

Première méthode

On capture successivement, au hasard (donc avec équiprobabilité) et avec remise en liberté après l'observation du sujet, n lions. Soit Y_n le nombre de lions tatoués parmi eux.

1. Déterminer la loi de Y_n. En déduire que $\frac{1}{nm}Y_n$ est un estimateur sans biais et convergent de $\frac{1}{N}$.

2. Pourquoi ne peut-on pas prendre $\frac{nm}{Y_n}$ comme estimateur de N ?

3. On pose $B_n = \frac{m(n+1)}{Y_n+1}$. Calculer l'espérance de B_n et montrer que B_n est un estimateur asymptotiquement sans biais de N.

Seconde méthode

On se donne $n \in \mathbb{N}^*$. On capture également, un par un, des lions de Gir au hasard et avec remise en liberté après l'observation du sujet.

On note X_n, la variable aléatoire égale au nombre de lions qu'il a été juste nécessaire de capturer pour en obtenir n tatoués.

On pose $D_1 = X_1$, et pour tout i de $[\![2, n]\!]$, $D_i = X_i - X_{i-1}$. On admet que les D_i sont des variables indépendantes deux-à-deux.

4. a) Pour tout i de $[\![2,n]\!]$, que représente concrètement D_i ?

b) Déterminer, pour tout i de $[\![2,n]\!]$, la loi de D_i, son espérance et sa variance. En déduire l'espérance et la variance de X_n.

c) On pose $A_n = \frac{m}{n} X_n$. Montrer que A_n est un estimateur sans biais et convergent de N et déterminer son risque quadratique.

5. a) Pour n assez grand, par quelle loi peut-on approcher la loi de la variable aléatoire $\widetilde{X_n} = \frac{X_n}{n}$?

b) On a tatoué $m = 200$ lions, puis capturé 450 lions, pour obtenir $n = 50$ lions marqués. On note σ l'écart-type de A_{50}. On a pu prouver que $\sigma \leqslant 100$. Déterminer en fonction de σ, un intervalle de confiance pour N au seuil de confiance $0,9$ (on rappelle que $\Phi(1,64) \simeq 0.95$).

Solution :

1. La probabilité de succès à chaque capture vaut $\frac{m}{N}$, donc $Y_n \hookrightarrow \mathcal{B}(n, \frac{m}{N})$.
Comme $E\big(\frac{Y_n}{nm}\big) = \frac{1}{N}$: $\frac{Y_n}{nm}$ est un estimateur sans biais de $\frac{1}{N}$.
De plus :
$V\big(\frac{Y_n}{nm}\big) = \frac{1}{(nm)^2} V(Y_n) = \frac{1}{(nm)^2} \times n\frac{m}{N}(1 - \frac{m}{N}) = \big(1 - \frac{m}{N}\big) \times \frac{1}{nmN} \underset{n\to\infty}{\longrightarrow} 0.$
L'inégalité de Bienaymé-Tchebicheff permet d'affirmer que cet estimateur est convergent.

2. Comme $P(Y_n = 0) \neq 0$, $\frac{nm}{Y_n}$ aurait une probabilité non nulle de ne pas être défini : aussi ne peut-on pas le choisir comme estimateur.

3. Par le théorème de transfert :

$$\begin{aligned}
E(B_n) &= \sum_{k=0}^{n} \frac{m(n+1)}{k+1} \binom{n}{k} \big(\frac{m}{N}\big)^k \big(1 - \frac{m}{N}\big)^{n-k} \\
&= \sum_{k=0}^{n} m \binom{n+1}{k+1} \big(\frac{m}{N}\big)^k \big(1 - \frac{m}{N}\big)^{n-k} \\
&= N \sum_{k=0}^{n} \binom{n+1}{k+1} \big(\frac{m}{N}\big)^{k+1} \big(1 - \frac{m}{N}\big)^{n-k} \\
&= N \sum_{k=1}^{n+1} \binom{n+1}{k} \big(\frac{m}{N}\big)^k \big(1 - \frac{m}{N}\big)^{n+1-k},
\end{aligned}$$

soit, en utilisant la formule du binôme :
$E(B_n) = N\big(1 - \big(1 - \frac{m}{N}\big)^{n+1}\big) \underset{n\to\infty}{\to} N$. Ainsi B_n est un estimateur asymptotiquement sans biais de N.

4. a) Pour tout $i \in [\![2, n]\!]$, D_i représente le nombre de lions supplémentaires, à partir du $(i-1)^{\text{ème}}$ lion tatoué obtenu, qu'il faut capturer pour obtenir un $i^{\text{ème}}$ lion tatoué (qui a pu être déjà recapturé), nous sommes donc dans le schéma géométrique ...

b) Ainsi $D_i(\Omega) = \mathbb{N}^*$ et D_i suit la loi géométrique de paramètre $\frac{m}{N}$. Ainsi :

$$E(D_i) = \frac{N}{m} \text{ et } V(D_i) = \frac{N(N-m)}{m^2}$$

Or $X_n = \sum_{i=1}^{n} D_i$. Donc, par indépendance :

$$E(X_n) = \frac{nN}{m} \text{ et } V(X_n) = \frac{nN(N-m)}{m^2}$$

c) On a $E(A_n) = N$ et $V(A_n) = \frac{N(N-m)}{n}$. Ainsi A_n est un estimateur sans biais de N, convergent et de risque quadratique égal à $V(A_n)$.

5. a) En utilisant le théorème de la limite centrée :

$$\frac{X_n - (nN/m)}{\sqrt{n}\frac{\sqrt{N(N-m)}}{m}} = \frac{mX_n - nN}{\sqrt{nN(N-m)}} \hookrightarrow \mathcal{N}(0,1)$$

Ainsi, pour n assez grand, on peut approcher la loi de $\frac{X_n}{n}$ par la loi normale $\mathcal{N}\left(\frac{N}{m}, \frac{N(N-m)}{nm^2}\right)$

b) Ici $n = 50, m = 200, X_n = 450$. On pose $\sigma = \sigma(A_n)$. On considère que : $\frac{A_n - N}{\sigma} \hookrightarrow \mathcal{N}(0,1)$.

En utilisant cette approximation, pour tout $t > 0$:

$$P\Big(\Big|\frac{A_n - N}{\sigma}\Big| \leqslant t\Big) = \Phi(t) - \Phi(-t) = 2\Phi(t) - 1$$

Donc : $P\Big(\Big|\frac{A_n - N}{\sigma}\Big| \leqslant t\Big) \geqslant 0.9 \iff \Phi(t) \geqslant 0.95 \iff t \geqslant 1.64$

Comme $\sigma \leqslant 100$, $(A_n - t\sigma \leqslant N \leqslant A_n + t\sigma) \subset (A_n - 100t \leqslant N \leqslant A_n + 100t)$.

Si l'on prend $I = [A_n - 164, A_n + 164]$, alors :

$$P(N \in I) = P\Big(\Big|\frac{A_n - N}{\sigma}\Big| \leqslant 1.64\Big) \geqslant 0.9$$

Comme $A_{50} = 1800$, il vient $I = [1636, 1964]$.

Exercice 3.2.

Une urne contient initialement n boules numérotées depuis 1 jusqu'à n, avec $n \geqslant 2$. On vide l'urne en extrayant toutes les boules une à une, au hasard et sans remise.

1. Pour i compris entre 1 et n, on note X_i la variable aléatoire qui vaut 1 si la boule obtenue au $i^{\text{ème}}$ tirage porte le numéro i et 0 dans le cas contraire. Quelle est la loi de X_i ?

2. En déduire l'espérance du nombre de fois où il y a concidence entre le rang du tirage et le numéro de la boule obtenue, lorsque l'on vide l'urne.

3. Pour k compris entre 1 et n, on dit que le résultat du $k^{\text{ème}}$ tirage est un «record» si la boule obtenue à ce tirage porte un numéro strictement supérieur à tous les numéros obtenus jusqu'alors. (par convention, le résultat du premier tirage sera toujours considéré comme un record).

a) Combien existe-t-il de façons de vider l'urne et pour lesquelles il n'y a qu'un seul record ? Pour lesquelles il y a n records ?

b) Montrer que pour p et q entiers naturels, on a la relation suivante :

$$\binom{p}{p} + \binom{p+1}{p} + \cdots + \binom{p+q}{p} = \binom{p+q+1}{p+1}$$

(où $\binom{n}{m}$ est le nombre de parties à m éléments d'un ensemble à n éléments).

4. Soit k fixé entre 2 et n et j fixé entre k et n. Combien existe-t-il de façons de vider l'urne, pour lesquelles la $k^{\text{ème}}$ boule obtenue porte le numéro j et le $k^{\text{ème}}$ tirage constitue un record ?

5. Combien existe-t-il de façons de vider l'urne pour lesquelles le $k^{\text{ème}}$ tirage est un record ? En déduire la probabilité que le $k^{\text{ème}}$ tirage soit un record. Pouvait-on avoir ce résultat directement ?

6. Pour k compris entre 1 et n, soit Y_k la variable aléatoire qui prend la valeur 1 si le résultat du $k^{\text{ème}}$ tirage est un record et 0 sinon. Déterminer la loi de Y_k. Soit R le nombre aléatoire de records obtenus lorsque l'on vide l'urne. Déterminer l'espérance de R.

Solution :

1. Parmi tous les $n!$ tirages possibles, il suffit de compter ceux pour lesquels on a obtenu i au $i^{\text{ème}}$ tirage : il y en a $(n-1)!$. Donc, X_i suit la loi de Bernoulli de paramètre $1/n$.

2. On a : $N = \sum\limits_{i=1}^{n} X_i$ (nombre de coïncidences). Par linéarité de l'espérance, on a : $E(N) = n \times \dfrac{1}{n} = 1$.

3. a) Le premier tirage étant considéré comme un record, s'il n'y a eu qu'un seul record, c'est que la boule numéro n est sortie au premier tirage. Ensuite, peu importe ce qui se passe. Il y a donc $(n-1)!$ façons de vider l'urne pour lesquelles il n'y a qu'un record.

S'il y eu n records, cela signifie qu'à chaque tirage, on a obtenu un numéro supérieur à celui obtenu au tirage précédent ; la seule possibilité est d'avoir tiré les boules par ordre croissant, donc un seul tirage convient.

b) La formule demandée :

$$\binom{p}{p}+\binom{p+1}{p}+\cdots+\binom{p+q}{p}=\binom{p+q+1}{p+1}$$

s'obtient, par exemple, par récurrence en utilisant la formule de Pascal.

4. Si le $k^{\text{ème}}$ tirage est un record avec le numéro j sorti, cela signifie qu'au cours des $(k-1)$ tirages précédents, on n'a obtenu que des numéros inférieurs ou égaux à $j-1$.

- si $j<k$, c'est impossible ;
- si $j \geqslant k$ il y a $\binom{j-1}{k-1}$ façons de choisir les boules sorties lors des $(k-1)$ premiers tirages (elles doivent avoir toutes un numéro compris entre 1 et $j-1$), $(k-1)!$ façons de les ordonner, puis une façon de placer la boule j en place k, et la fin est une permutation quelconque des boules restantes. Finalement, il y a :

$$\binom{j-1}{k-1}\times(k-1)!\times 1\times(n-k)!$$

façons convenables de vider l'urne

(on pourrait simplifier, mais ce n'est pas intéressant pour la question suivante).

5. Par la question précédente, le nombre de façons d'avoir un record au $k^{\text{ème}}$ tirage est égal à :

$$(n-k)!(k-1)!\sum_{j=k}^{n}\binom{j-1}{k-1}=(n-k)!(k-1)!\binom{n}{k}$$

et la probabilité demandée est donc :

$$\frac{(n-k)!(k-1)!}{n!}\binom{n}{k}=\frac{1}{k}$$

On peut retrouver ce résultat directement avec le raisonnement suivant :

dire que le $k^{\text{ème}}$ tirage est un record, c'est dire que le plus grand des k premiers résultats est le dernier, c'est donc dire qu'en ordonnant k nombres deux à deux distincts le plus grand est le dernier. La probabilité est donc de $\frac{1}{k}$.

6. La variable aléatoire Y_k suit la loi de Bernoulli de paramètre $\frac{1}{k}$. La variable aléatoire R est égale à $\sum_{k=1}^{n} Y_k$ et $E(R)=\sum_{k=1}^{n} E(Y_k)=n\sum_{k=1}^{n}\frac{1}{k}$.

Exercice 3.3.

1. On considère la fonction g définie sur $\mathbb{R}$ par $g(x)=\dfrac{2}{\pi(\mathrm{e}^{x}+\mathrm{e}^{-x})}$.

Montrer que g est une fonction densité de probabilité.

Dans la suite, on note X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ admettant g pour densité (on dit que X suit la loi d'Euler).

2. Déterminer la fonction de répartition F_X de X.

3. Montrer que X admet des moments de tous ordres et calculer son espérance.

4. On pose $Y = \mathrm{e}^X$.

a) Montrer que Y est une variable aléatoire à densité et déterminer une densité de Y.

b) La variable aléatoire Y admet-elle une espérance ?

5. On considère une suite de variables aléatoires $(Y_n)_{n\in\mathbb{N}^*}$ définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et suivant toutes la même loi que Y. Pour $n \in \mathbb{N}^*$, on pose $M_n = \sup(Y_1, Y_2, \dots, Y_n)$.

a) Déterminer la fonction de répartition de M_n.

b) Pour $n \geqslant 1$, on pose $Z_n = \dfrac{n}{M_n}$. Montrer que la suite $(Z_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire dont on donnera la loi.

Solution :

1. La fonction g est continue sur $\mathbb{R}$ et positive. De plus, en abrégeant :

$$\int_{-\infty}^{+\infty} g(x)dx = \frac{2}{\pi}\int_{-\infty}^{+\infty} \frac{\mathrm{e}^x}{\mathrm{e}^{2x}+1}dx = \Big[\frac{2}{\pi}\operatorname{Arc\,tan}(\mathrm{e}^x)\Big]_{-\infty}^{+\infty} = \frac{2}{\pi}(\frac{\pi}{2} - 0) = 1.$$

g est bien une densité de probabilité.

2. De la même façon, pour tout réel t :

$$F_X(t) = \Big[\frac{2}{\pi}\operatorname{Arc\,tan}(\mathrm{e}^x)\Big]_{-\infty}^{t} = \frac{2}{\pi}\operatorname{Arc\,tan}(\mathrm{e}^t).$$

3. a) Soit $n \geqslant 1$, et $h_n : x \mapsto x^n g(x)$. La fonction h_n est continue sur $\mathbb{R}$.

- au voisinage de $+\infty$, $h_n(x) \sim \dfrac{2}{\pi}x^n\mathrm{e}^{-x} = o(1/x^2)$,
- au voisinage de $-\infty$, $|h_n(x)| \sim \dfrac{2}{\pi}|x|^n\mathrm{e}^{x} = o(1/x^2)$.

(on aurait pu invoquer la parité de la fonction g)

Ainsi la convergence (absolue) de l'intégrale sur $\mathbb{R}$ de h_n est acquise et X admet des moments de tous ordres.

b) La fonction h_1 est impaire d'intégrale sur $\mathbb{R}$ convergente, donc $E(X) = 0$.

4. a) On a $Y(\Omega) =]0, +\infty[$ et pour $y > 0$, $(Y \leqslant y) = (\mathrm{e}^X \leqslant y) = (X \leqslant \ln y) \in \mathcal{A}$, donc Y est une variable aléatoire et :

Pour $y > 0$, $P(Y \leqslant y) = P(X \leqslant \ln(y)) = \dfrac{2}{\pi}\operatorname{Arc\,tan}(y)$. Ainsi :

$$F_Y(y) = \begin{cases} 0 & \text{si } y < 0 \\ \dfrac{2}{\pi}\operatorname{Arc\,tan}(y) & \text{si } y \geqslant 0 \end{cases}$$

F_Y est de classe $\mathcal{C}^1$ par morceaux, donc Y est une variable aléatoire à densité et une densité f_Y de Y s'obtient par exemple par dérivation sur $\mathbb{R}^*$ et en faisant un choix arbitraire en 0 :

$$f_Y(y) = \begin{cases} 0 & \text{si } y < 0 \\ \dfrac{2}{\pi}\dfrac{1}{1+y^2} & \text{si } y \geqslant 0 \end{cases}$$

b) Il est clair que Y n'admet pas d'espérance, puisque $yf_Y(y) \sim \dfrac{C}{y}$ au voisinage de $+\infty$ et donc déjà l'intégrale $\displaystyle\int^{+\infty} yf_Y(y)\,dy$ diverge.

5. a) On a : $[M_n \leqslant x] = \displaystyle\bigcap_{i=1}^{n}[Y_i \leqslant x]$, ce qui montre que M_n est une variable aléatoire, clairement à valeurs dans $\mathbb{R}^+$. De plus, pour $x \geqslant 0$, par indépendance des (Y_i), on a :

$$P(M_n \leqslant x) = \prod_{i=1}^{n} P(Y_i \leqslant x) = \Big(\frac{2}{\pi}\operatorname{Arc\,tan}(x)\Big)^n.$$

Donc : $F_{M_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ \Big(\dfrac{2}{\pi}\operatorname{Arc\,tan}(x)\Big)^n & \text{si } x \geqslant 0 \end{cases}$

b) On a $Z_n(\Omega) =]0, +\infty[$ et pour $x > 0$:

$P(Z_n \geqslant x) = P(M_n \leqslant \dfrac{n}{x}) = \Big(\dfrac{2}{\pi}\operatorname{Arc\,tan}(\dfrac{n}{x})\Big)^n = \Big(\dfrac{2}{\pi}(\dfrac{\pi}{2} - \operatorname{Arc\,tan}(\dfrac{x}{n}))\Big)^n$

car pour $t > 0$, on a $\operatorname{Arc\,tan} t + \operatorname{Arc\,tan}\dfrac{1}{t} = \dfrac{\pi}{2}$

Donc : $P(Z_n \geqslant x) = \big(1 - \dfrac{2}{\pi}\operatorname{Arc\,tan}(\dfrac{x}{n})\big)^n = \exp(n\ln(1 - \dfrac{2}{\pi}\operatorname{Arc\,tan}(\dfrac{x}{n})))$

Or : $n\ln(1 - \dfrac{2}{\pi}\operatorname{Arc\,tan}(\dfrac{x}{n})) \underset{(n\to\infty)}{\sim} -\dfrac{2}{\pi}n\operatorname{Arc\,tan}(\dfrac{x}{n}) \underset{(n\to\infty)}{\sim} -\dfrac{2x}{\pi}$

Donc $\forall\, x > 0$, $\displaystyle\lim_{n\to\infty} P(Z_n \leqslant x) = 1 - \mathrm{e}^{-\frac{2x}{\pi}}$ et comme Z_n est à valeurs dans $\mathbb{R}^+$, ceci prouve que Z_n tend en loi vers la loi exponentielle de paramètre $\dfrac{2}{\pi}$.

Exercice 3.4.

Une urne contient n boules numérotées de 1 à n et k boules bleues non numérotées. Les boules sont tirées avec remise jusqu'à ce qu'une boule bleue soit tirée. Au cours de ces tirages, on définit le nombre R de répétitions de la manière suivante :

au début, $R = 0$. Ensuite, on ajoute 1 à R dès que l'on obtient une boule numérotée qui avait été déjà tirée précédemment.

1. Déterminer les probabilités des événements suivants :

$A_1 =$ « la première boule tirée est la boule numéro 1 ».

$A_2 =$ « la première boule tirée est une boule portant un numéro strictement supérieur à 1 ».

$A_3 =$ « la première boule tirée est une boule bleue ».

2. On note A_0 l'événement « la boule numéro 1 n'est jamais tirée lors du jeu ». En utilisant la formule des probabilités totales avec les événements précédents, montrer que $P(A_0) = \dfrac{k}{k+1}$.

3. On note X le nombre de fois où l'on a tiré la boule 1 au cours du jeu. En utilisant un raisonnement analogue à celui de la question précédente, montrer que $E(X) = \dfrac{1}{k}$.

4. On définit la variable aléatoire Y par :

$$\begin{cases} \text{Si } X \geqslant 1, \text{ alors } Y = X - 1 \\ \text{Si } X = 0, \text{ alors } Y = 0 \end{cases}$$

(Y est donc le nombre de répétitions de la boule numérotée 1.)
Montrer que $E(Y) = \sum\limits_{m \geq 1} (m-1)P(X = m)$ puis que $E(Y) = \dfrac{1}{k(k+1)}$.

Soit r un entier naturel. On recherche la valeur minimale de k (en fonction de n et r) de manière à ce que le nombre moyen t de répétitions soit inférieur ou égal à r.

5. Montrer que $t = nE(Y)$.

6. En déduire que la valeur minimale recherchée est $k_0 = \lfloor \sqrt{\frac{n}{r} + \frac{1}{4}} - \frac{1}{2} \rfloor$.

Solution :

1. On a clairement : $P(A_1) = \dfrac{1}{n+k}$ et $P(A_2) = \dfrac{n-1}{n+k}$. La probabilité que le jeu s'arrête dès la première étape vaut $P(A_3) = \dfrac{k}{n+k}$.

2. Les événements A_1, A_2, A_3 forment une partition de l'ensemble Ω. De plus, les événements A_0 et A_1 sont incompatibles. Si la première boule tirée est une boule bleue, le jeu s'arrête et $A_3 \subset A_0$. Les tirages étant indépendants, $P_{A_2}(A_0) = P(A_0)$.

Ainsi, d'après la formule des probabilités totales :

$$\begin{aligned} P(A_0) &= \sum_{i=1}^{3} P(A_i)P_{A_i}(A_0) = P(A_1)\times 0 + P(A_2)P_{A_2}(A_0) + P(A_3)\times 1 \\ &= \frac{n-1}{n+k}P(A_0) + \frac{k}{n+k} \end{aligned}$$

Soit : $\dfrac{n+k-n+1}{n+k}P(A_0) = \dfrac{k}{n+k}$, d'où $P(A_0) = \dfrac{k}{k+1}$.

3. On utilise la décomposition selon les événements A_1, A_2, A_3. Ainsi :

$$\begin{aligned} E(X) &= E(X|A_1)P(A_1) + E(X|A_2)P(A_2) + E(X|A_3)P(A_3) \\ &= (E[X]+1)\times P(A_1) + E(X)\times P(A_2) + 0, \end{aligned}$$

soit : $E(X) = \dfrac{E(X)+1}{n+k} + \dfrac{n-1}{n+k}E(X)$, d'où : $E(X)\bigl(1 - \dfrac{n}{n+k}\bigr) = \dfrac{1}{n+k}$ et $E(X) = \dfrac{1}{k}$.

4. D'après la définition, $Y = \max\{0, X-1\}$. Ainsi :

$$\begin{aligned} E(Y) &= \sum_{k=1}^{\infty} kP(Y=k) = \sum_{k=1}^{\infty} kP(Y=k) = \sum_{k=1}^{\infty} kP(X=k+1) \\ &= \sum_{m=2}^{\infty} (m-1)P(X=m), \text{ soit :} \end{aligned}$$

$$\begin{aligned} E(Y) &= \sum_{m=1}^{\infty} (m-1)P(X=m) = E(X) - \sum_{m=1}^{\infty} P(X=m) = E(X) - P(A_0) \\ &= \frac{1}{k} - \frac{k}{k+1} = \frac{1}{k(k+1)} \end{aligned}$$

5. Pour $i \in [\![1, n]\!]$, en notant Y_i le nombre de répétitions de la boule numérotée i, on a clairement, par symétrie, $E(Y_i) = E(Y)$. Ainsi :

$$t = E\Bigl(\sum_{i=1}^{n} Y_i\Bigr) = \sum_{i=1}^{n} E(Y_i) = nE(Y).$$

6. Finalement, $t \leqslant r \iff nE(Y) \leqslant r \iff \dfrac{n}{k(k+1)} \leqslant r \iff rk^2 + rk - n \leqslant 0$.

Soit, comme k est entier, $k > \left\lfloor \sqrt{\dfrac{n}{r} + \dfrac{1}{4}} - \dfrac{1}{2} \right\rfloor$.

Exercice 3.5.

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soient $\theta > 0$ et $X, X_1, \ldots, X_n$ des variables aléatoires indépendantes de même loi uniforme sur $[0, \theta]$.

1. Montrer que $\ln X$ admet une espérance et calculer son espérance m.

2. Pour tout $n \in \mathbb{N}^*$, on pose $Y_n = \bigl(X_1 \times \cdots \times X_n\bigr)^{\frac{1}{n}}$.

   a) Montrer que la suite $(\ln Y_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers m.

   b) En déduire que $(Y_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers e^m.

3. Soit $M_n = \max\limits_{1\leq i\leq n} X_i$. Calculer, pour tout $x \in \mathbb{R}$, $P(M_n \leqslant x)$.

4. En déduire que $(M_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers une variable aléatoire constante C que l'on déterminera en fonction de θ.

5. a) Donner, pour tout réel x, la nature de la série $\sum_{n\in\mathbb{N}} P(M_n \leqslant x)$.

b) Montrer que pour tout $\omega \in \Omega$, la suite $(M_n(\omega))_{n\in\mathbb{N}^*}$ converge.

c) Soit $\varepsilon > 0$. Montrer que $P(\lim\limits_{n\to+\infty} M_n \leqslant \theta - \varepsilon) = 0$.

Solution :

1. La variable aléatoire X admet pour densité $\dfrac{1}{\theta}\mathbf{1}_{[0,\theta]}$. De plus, pour $a > 0$, $\displaystyle\int_0^a \ln x\, dx$ converge. Ainsi, d'après le théorème de transfert :

$$m = E(X) = \int_0^\theta \ln x \times \frac{1}{\theta}\, dx = \frac{1}{\theta}\big[x\ln x - x\big]_{\to 0}^{\theta} = \ln\theta - 1.$$

2. a) D'après la loi faible des grands nombres, la suite de variables aléatoires $\big(\frac{1}{n}\sum\limits_{i=1}^{n} \ln Y_i\big)_n$ converge en probabilité vers m. Soit :

$$\forall\, \varepsilon > 0,\ \lim_{n\to+\infty} P(|\ln Y_n - m| \geqslant \varepsilon) = 0$$

b) Soient ε, $\eta > 0$. Notons $\varepsilon_0 = \min\{\ln(1+\varepsilon e^{-m}), -\ln(1-\varepsilon e^{-m})\}$.
D'après la question précédente, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, $P(|\ln Y_n - m| \geqslant \varepsilon_0) \leqslant \eta$. Par suite :

$$\begin{aligned} P(|Y_n - e^m| \leqslant \varepsilon) &= P\big(\ln(1-\varepsilon e^{-m}) \leqslant \ln(Y_n) - m \leqslant \ln(1+\varepsilon e^{-m})\big) \\ &\geqslant P(|\ln Y_n - m| \leqslant \varepsilon_0) \geqslant 1 - \eta. \end{aligned}$$

Soit, $P(|Y_n - e^m| \geqslant \varepsilon) \leqslant \eta$.
Ainsi, la suite $(Y_n)_n$ converge en probabilité vers e^m.

3. D'après la définition de M_n, la positivité et l'indépendance des (X_i),

- Si $x \leqslant 0$, $P(M_n \leqslant x) = 0$.
- Si $x \geqslant \theta$, $P(M_n \leqslant x) = 1$.
- Si $x \in [0,\theta]$,

$$\begin{aligned} P(M_n \leqslant x) &= P(\forall\, i \in [\![1,n]\!],\, X_i \leqslant x) = \prod_{i=1}^{n} P(X_i \leqslant x) = \Big(\int_0^x \frac{1}{\theta} dx\Big)^n \\ &= \big(\frac{x}{\theta}\big)^n. \end{aligned}$$

4. Soit C la variable aléatoire définie par $P(C = \theta) = 1$. Soit $\varepsilon > 0$. On a :

$$\begin{aligned} P(|M_n - C| \geqslant \varepsilon) &= P(|M_n - \theta| \geqslant \varepsilon) = P(\theta - M_n \geqslant \varepsilon) \\ &= P(M_n \leqslant \theta - \varepsilon) \underset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Ainsi, la suite $(M_n)_n$ converge en probabilité vers la variable aléatoire constante égale à θ.

5. a) D'après la question 3., la série de terme général $P(M_n \leqslant x)$ converge si et seulement si $x \leqslant \theta$.

b) La suite $(M_n(\omega))$ étant croissante et majorée par θ, elle converge.

c) En utilisant la question précédente :

$$P(\lim_{n\to\infty} M_n \leqslant \theta-\varepsilon) = P(\forall\, n\in\mathbb{N}, M_n \leqslant \theta-\varepsilon) = P(\bigcap_{n\in\mathbb{N}} \{M_n \leqslant \theta-\varepsilon\})$$
$$\leqslant P(M_k \leqslant \theta-\varepsilon), \text{ pour tout entier naturel } k.$$

Or, d'après les questions précédentes, $\lim_{k\to\infty} P(M_k \leqslant \theta-\varepsilon) = 0$, soit :

$$P(\lim_{n\to\infty} M_n \leqslant \theta-\varepsilon) = 0.$$

Exercice 3.6.

Soient $X_1,\dots,X_n,\dots$ des variables aléatoires, définies sur le même espace probabilisé, indépendantes qui suivent toutes la loi de Poisson de paramètre 1. Pour tout $n\in\mathbb{N}^*$, on note $S_n = X_1+\cdots+X_n$. Soit M un nombre réel tel que $M>0$.

1. Donner la loi de S_n, son espérance et sa variance. Justifier la convergence en loi de la suite $\big(\frac{S_n-n}{\sqrt{n}}\big)_{n\in\mathbb{N}^*}$ vers une variable aléatoire N dont on précisera la loi.

2. a) Pour tout $n\in\mathbb{N}^*$ et tout $k\in[\![1,n]\!]$, montrer que :

$$\prod_{i=0}^{k}\big(1-\frac{i}{n}\big) \leqslant \exp\big(-\frac{k^2}{2n}\big).$$

b) On note β_n la partie entière de $M\sqrt{n}$. Montrer qu'il existe $n_1\in\mathbb{N}$ tel que :

$$n\geqslant n_1 \implies n\geqslant \beta_n+1.$$

Dans toute la suite on se place dans le cas où $n\geqslant n_1$.

3. Soit la variable aléatoire $Y_n = h\big(\frac{S_n-n}{\sqrt{n}}\big)$, où la fonction réelle h est définie par :

$$h(x)=\begin{cases} -x & \text{si } x\in[-M,0]\\ 0 & \text{sinon}\end{cases}$$

a) Montrer que :

$$E(Y_n) = \frac{\mathrm{e}^{-n}}{\sqrt{n}}\big(\frac{n^{n+1}}{n!} - \frac{n^{n-\beta_n}}{(n-\beta_n-1)!}\big).$$

b) Puis que :

$$E(Y_n) \leqslant \frac{\mathrm{e}^{-n}n^{n+\frac12}}{n!} \leqslant E(Y_n) + \frac{\mathrm{e}^{-n}n^{n+\frac12}}{n!}\,\mathrm{e}^{-\frac{\beta_n^2}{2n}}.$$

4. En étudiant la série de terme général $u_n = \ln(a_{n+1}) - \ln(a_n)$, montrer la convergence de la suite de terme général $a_n = \dfrac{\mathrm{e}^{-n} n^{n+\frac{1}{2}}}{n!}$ vers une limite $K > 0$.

5. En admettant que $E(Y_n)$ converge vers $E(h(N))$ quand n tend vers $+\infty$, montrer que :

$$\frac{1-\mathrm{e}^{-\frac{M^2}{2}}}{\sqrt{2\pi}} \leqslant K \leqslant \frac{1-\mathrm{e}^{-\frac{M^2}{2}}}{\sqrt{2\pi}} + K\,\mathrm{e}^{-\frac{M^2}{2}}.$$

En déduire que : $n! \underset{n\to\infty}{\sim} \sqrt{2\pi}\, n^{n+\frac{1}{2}}\mathrm{e}^{-n}$.

Solution :

1. S_n suit la loi de Poisson de paramètre n, d'espérance n et de variance n. D'après le théorème central limite, la suite $\big(\dfrac{S_n - n}{\sqrt{n}}\big)_n$ converge en loi vers N qui suit la loi normale centrée réduite.

2. a) D'après l'inégalité de convexité $1 - x \leqslant \mathrm{e}^{-x}$, on a : (tout est positif) :

$$\prod_{i=0}^{k}\big(1-\frac{i}{n}\big) \leqslant \prod_{i=0}^{k}\mathrm{e}^{-\frac{i}{n}} = \exp\big(-\sum_{i=0}^{k}\frac{i}{n}\big) = \exp\big(-\frac{k(k+1)}{2n}\big) \leqslant \exp\big(-\frac{k^2}{2n}\big).$$

b) Comme $\beta_n \leqslant M\sqrt{n}$, pour avoir $n \geqslant \beta_n + 1$ il suffit que $n - M\sqrt{n} - 1 \geqslant 0$. Cette inéquation du second degré en $\sqrt{n}$ possède un coefficient dominant positif donc elle est vraie au voisinage de $+\infty$.
On peut aussi dire que $\beta_n = \lfloor M\sqrt{n}\rfloor \underset{(\infty)}{\sim} M\sqrt{n} \underset{(\infty)}{=} o(n)$

3. a) Par théorème de transfert, sous réserve de convergence de la série, on a :

$$E(Y_n) = \sum_{k=0}^{+\infty} h\big(\frac{k-n}{\sqrt{n}}\big)\frac{\mathrm{e}^{-n}n^k}{k!}, \text{ or } h\big(\frac{k-n}{\sqrt{n}}\big) = \begin{cases} -\dfrac{k-n}{\sqrt{n}} & \text{si } \dfrac{k-n}{\sqrt{n}} \in [-M, 0] \\ 0 & \text{sinon} \end{cases}.$$

Et : $\dfrac{k-n}{\sqrt{n}} \in [-M, 0] \iff k \in [n - M\sqrt{n}, n] \iff k \in [n - \beta_n, n]$.
Donc la série converge (il n'y a qu'un nombre fini de termes non nuls) et, comme $[n-\beta_n, n] \subset \mathbb{N}^*$ (d'après la condition $n \geqslant n_1$), on a :

$$E(Y_n) = \sum_{k=n-\beta_n}^{n} -\frac{k-n}{\sqrt{n}} \times \frac{\mathrm{e}^{-n}n^k}{k!} = \frac{\mathrm{e}^{-n}}{\sqrt{n}}\sum_{k=n-\beta_n}^{n}\big(\frac{n^{k+1}}{k!} - \frac{n^k}{(k-1)!}\big)$$
$$= \frac{\mathrm{e}^{-n}}{\sqrt{n}}\big(\frac{n^{n+1}}{n!} - \frac{n^{n-\beta_n}}{(n-\beta_n-1)!}\big)$$

b) La première inégalité est évidente d'après l'égalité ci-dessus. Pour la seconde on utilise 2. a) avec $k = \beta_n$:

$$\frac{n^{n-k}}{(n-k-1)!}\times\frac{n!}{n^{n+1}}=\prod_{i=0}^{k}\big(1-\frac{i}{n}\big)\leqslant\exp\big(-\frac{k^2}{2n}\big).$$

D'où :

$$\frac{n^{n-\beta_n}}{(n-\beta_n-1)!}\leqslant\frac{n^{n+1}}{n!}\exp\big(-\frac{\beta_n^2}{2n}\big).$$

4. Le calcul donne :

$$u_n=\big(n+1+\frac{1}{2}\big)\ln(n+1)-(n+1)-\ln((n+1)!)-\big((n+\frac{1}{2})\ln n-n-\ln(n!)\big)$$
$$u_n=\big(n+1+\frac{1}{2}\big)\ln(n+1)-1-\ln(n+1)-\big(n+\frac{1}{2}\big)\ln n=\big(n+\frac{1}{2}\big)\ln\big(1+\frac{1}{n}\big)-1$$
$$=\big(n+\frac{1}{2}\big)\big(\frac{1}{n}-\frac{1}{2n^2}+\frac{1}{3n^3}+o(\frac{1}{n^3})\big)-1=\frac{1}{12n^2}+o(\frac{1}{n^2})\underset{(\infty)}{\sim}\frac{1}{12n^2}$$

Ainsi, par la règle de Riemann $\sum u_n$ converge ; la suite $\ln(a_n)$ converge donc par télescopage ; ainsi (a_n) a une limite K strictement positive (c'est une exponentielle).

5. On passe à la limite quand n tend vers l'infini dans l'encadrement 3. b) en utilisant :

- $\lim\limits_{n\to+\infty}-\frac{\beta_n^2}{2n}=-\frac{M^2}{2}$, d'après $M\sqrt{n}-1<\beta_n\leqslant M\sqrt{n}$.

Donc $\lim\limits_{n\to+\infty}\mathrm{e}^{-\frac{\beta_n^2}{2n}}=\mathrm{e}^{-\frac{M^2}{2}}$.

- $\lim\limits_{n\to+\infty}E(Y_n)=E(h(N))$ (admis) et par le théorème de transfert :

$$E(h(N))=\int_{-\infty}^{+\infty}h(x)\frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx=\int_{-M}^{0}-x\frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx=\frac{1-\mathrm{e}^{-\frac{M^2}{2}}}{\sqrt{2\pi}}$$

$$\frac{1-\mathrm{e}^{-\frac{M^2}{2}}}{\sqrt{2\pi}}\leqslant K\leqslant\frac{1-\mathrm{e}^{-\frac{M^2}{2}}}{\sqrt{2\pi}}+K\,\mathrm{e}^{-\frac{M^2}{2}}.$$

L'encadrement obtenu est valable pour tout $M>0$. En faisant tendre M vers $+\infty$ on a $K=\frac{1}{\sqrt{2\pi}}$, d'où l'équivalence d'après le résultat de la question 4.

Exercice 3.7.

Soit a et c deux paramètres réels avec $c>0$. Pour n entier supérieur ou égal à 2, soit $x_1,\dots,x_n$ des réels fixés non tous nuls. On considère n variables aléatoires $Y_1,\dots,Y_n$ mutuellement indépendantes telles que pour tout $i\in[\![1,n]\!]$, Y_i suit une loi normale avec $E(Y_i)=ax_i$ et $V(Y_i)=c$.

Pour tout $i\in[\![1,n]\!]$, on note y_i une réalisation de Y_i et f_{Y_i} la densité continue sur $\mathbb{R}$ de Y_i.

Soit F la fonction définie sur $\mathbb{R}\times\mathbb{R}_*^+$, à valeurs réelles, telle que :

$$F(a,c)=\ln\big(\prod_{i=1}^{n}f_{Y_i}(y_i)\big).$$

1. Montrer que F est de classe C^2 sur $\mathbb{R}\times\mathbb{R}_*^+$.

2. Montrer que F admet sur $\mathbb{R}\times\mathbb{R}_*^+$ un unique point critique $(\widehat{a},\widehat{c})$ que l'on déterminera.

3. a) Montrer que pour tout $a\in\mathbb{R}$, on a : $\sum\limits_{i=1}^{n}(y_i-ax_i)^2\geqslant\sum\limits_{i=1}^{n}(y_i-\widehat{a}x_i)^2$.

b) En déduire que pour tout $(a,c)\in\mathbb{R}\times\mathbb{R}_*^+$, on a : $F(a,c)-F(\widehat{a},\widehat{c})\leqslant 0$. Conclure.

4. On pose : $s_n=\sum\limits_{i=1}^{n}x_i^2$, $A_n=\dfrac{1}{s_n}\sum\limits_{i=1}^{n}x_iY_i$ et $C_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}(Y_i-A_nx_i)^2$.

a) Que représentent $\widehat{a}$ et $\widehat{c}$ pour A_n et C_n ?

b) Calculer $E(A_n)$ et $V(A_n)$.

c) On admet que la définition et les propriétés de la covariance de deux variables aléatoires discrètes s'appliquent aux variables aléatoires à densité.

Montrer que pour tout $i\in[\![1,n]\!]$, on a : $\mathrm{Cov}(Y_i,A_n)=\dfrac{x_i}{s_n}c$.

En remarquant que pour tout $i\in[\![1,n]\!]$, la variable aléatoire $(Y_i-A_nx_i)$ est centrée, calculer $E(C_n)$.

Solution :

1. On a : $f_{Y_i}(y_i)=\dfrac{1}{\sqrt{2\pi c}}\exp\Big(-\dfrac{1}{2c}(y_i-ax_i)^2\Big)$, donc

$$F(a,c)=-\frac{n}{2}\ln(2\pi c)-\frac{1}{2c}\sum_{i=1}^{n}(y_i-ax_i)^2$$

Les fonctions polynomiales sont de classe C^2 sur $\mathbb{R}$ et les fonctions $c\mapsto 1/c$ et $c\mapsto\ln c$ sont de classe C^2 sur $\mathbb{R}_*^+$, donc F est de classe C^2 sur $\mathbb{R}\times\mathbb{R}_*^+$.

2. $\dfrac{\partial F}{\partial a}(a,c)=\dfrac{1}{c}\sum\limits_{i=1}^{n}x_iy_i-\dfrac{a}{c}s_n$ et $\dfrac{\partial F}{\partial c}(a,c)=-\dfrac{n}{2c}+\dfrac{1}{2c^2}\sum\limits_{i=1}^{n}(y_i-ax_i)^2$.

L'unique point critique $(\widehat{a},\widehat{c})$ est donné par : $\widehat{a}=\dfrac{1}{s_n}\sum\limits_{i=1}^{n}x_iy_i$ et $\widehat{c}=\dfrac{1}{n}\sum\limits_{i=1}^{n}(y_i-\widehat{a}x_i)^2$.

3. a) $\sum\limits_{i=1}^{n}(y_i-ax_i)^2-\sum\limits_{i=1}^{n}(y_i-\widehat{a}x_i)^2=-2a\sum\limits_{i=1}^{n}x_iy_i+a^2s_n+2\widehat{a}\sum\limits_{i=1}^{n}x_iy_i-\widehat{a}^2s_n$

donc $\sum\limits_{i=1}^{n}(y_i-ax_i)^2-\sum\limits_{i=1}^{n}(y_i-\widehat{a}x_i)^2=s_n(a-\widehat{a})^2\geqslant 0$, car $s_n>0$.

Bilan : pour tout $a\in\mathbb{R}$, $\sum\limits_{i=1}^{n}(y_i-ax_i)^2\geqslant\sum\limits_{i=1}^{n}(y_i-\widehat{a}x_i)^2$.

b) Compte tenu de l'expression de $\widehat{c}$, on a : $F(\widehat{a},\widehat{c}) = -\frac{n}{2}\big(1+\ln(2\pi\widehat{c})\big)$. La question a) permet d'écrire :

$$F(a,c)-F(\widehat{a},\widehat{c}) = \frac{n}{2}\Big(1+\ln\Big(\frac{\widehat{c}}{c}\Big)\Big)-\frac{1}{2c^2}\sum_{i=1}^{n}(y_i-ax_i)^2 \leqslant \frac{n}{2}\Big(1+\ln\Big(\frac{\widehat{c}}{c}\Big)-\frac{\widehat{c}}{c}\Big) \leqslant 0$$

(car $u > 0 \Longrightarrow \ln u \leqslant u-1$).

Bilan : $\forall\,(a,c) \in \mathbb{R}\times\mathbb{R}_*^+, F(a,c) - F(\widehat{a},\widehat{c}) \leqslant 0$. Le couple $(\widehat{a},\widehat{c})$ est le point de $\mathbb{R}\times\mathbb{R}_*^+$ en lequel F admet un maximum global.

4. a) Les réels $\widehat{a}$ et $\widehat{c}$ sont les réalisations des variables aléatoires A_n et C_n respectivement.

b) $E(A_n) = \frac{1}{s_n}\sum_{i=1}^{n} x_i(ax_i) = \frac{s_n}{s_n}\,a = a.$

Par indépendance des Y_i, on a : $V(A_n) = \frac{1}{s_n^2}\sum_{i=1}^{n} x_i^2 c = \frac{c}{s_n}$.

c) $\forall\, i \in [\![1,n]\!]$,

$$\mathrm{Cov}(Y_i,A_n) = \mathrm{Cov}\Big(Y_i, \frac{1}{s_n}\sum_{i=1}^{n} x_iY_i\Big) = \frac{1}{s_n}\,\mathrm{Cov}(Y_i,x_iY_i) = \frac{x_i}{s_n}c.$$

D'autre part, $E(C_n) = \frac{1}{n}\sum_{i=1}^{n} E\big((Y_i-A_nx_i)^2\big) = \frac{1}{n}\sum_{i=1}^{n} V(Y_i-A_nx_i)$.

Or :

$$\begin{aligned} V(Y_i-A_nx_i) &= V(Y_i) + x_i^2V(A_n) - 2x_i\,\mathrm{Cov}(Y_i,A_n) = c + \frac{x_i^2c}{s_n} - \frac{2{x_i}^2c}{s_n} \\ &= c - \frac{x_i^2c}{s_n}. \end{aligned}$$

Par suite :

$$E(C_n) = \frac{1}{n}\sum_{i=1}^{n}\Big(c - \frac{x_i^2c}{s_n}\Big) = \frac{n-1}{n}\,c$$

Exercice 3.8.

Soit $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n,\dots$ une suite de variables de Bernoulli indépendantes telles que :

$$\text{pour tout } n \in \mathbb{N}^*,\ P(\varepsilon_n = +1) = P(\varepsilon_n = -1) = 1/2.$$

1. Calculer l'espérance $E\big((\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_n)^2\big)$ en fonction de n.

2. Soit $a \in\,]0,1[$ fixé.

a) Montrer l'inégalité $P(|\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_n| \geqslant an) \leqslant \frac{1}{a^2n}$.

b) Montrer que

$$P(|\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_n| \geqslant an) = \frac{1}{2^n}\sum_{\ell=0}^{n}\binom{n}{\ell}\mathbf{1}_{\{|2\ell-n|\geq an\}}$$

où $\mathbf{1}_{\{|2\ell-n|\geq an\}} = 1$ si ℓ satisfait $|2\ell-n| \geqslant an$ et 0 sinon.

c) Montrer que $\lim\limits_{n\to+\infty} \dfrac{1}{2^n} \sum\limits_{\ell=0}^{n} \dbinom{n}{\ell} \mathbf{1}_{\{|2\ell-n|\geqslant an\}} = 0$.

3. Soit N une variable aléatoire suivant la loi de Poisson, de paramètre $\theta > 0$, indépendante de la suite $(\varepsilon_n)_{n\geq 1}$. Calculer, en fonction de θ, l'espérance :

$$E\Big(\big(\sum_{n=1}^{N+1} \varepsilon_n\big)^2\Big).$$

Solution :

1. On a $E(\varepsilon_k) = 0$ et $V(\varepsilon_k) = 1$. D'après l'hypothèse d'indépendance :

$$V(\varepsilon_1 + \cdots + \varepsilon_n) = V(\varepsilon_1) + \cdots + V(\varepsilon_n) = n$$

et, puisque les variables aléatoires sont centrées :

$$E\big((\varepsilon_1 + \cdots + \varepsilon_n)^2\big) = V(\varepsilon_1 + \cdots + \varepsilon_n) = n.$$

2. a) D'après l'inégalité de Bienaymé-Tchebychev :

$$P(|\varepsilon_1 + \cdots + \varepsilon_n| \geqslant an) \leqslant \frac{V(\varepsilon_1 + \cdots + \varepsilon_n)}{a^2 n^2} = \frac{1}{a^2 n}.$$

b) Si parmi les variables aléatoires $\varepsilon_1, \ldots, \varepsilon_n$ il y en a exactement ℓ qui valent 1 (les autres valant -1), alors on a : $\varepsilon_1 + \cdots + \varepsilon_n = \ell - (n-\ell) = 2\ell - n$. Comme il y a $\dbinom{n}{\ell}$ façons de choisir les places des 1 et que les événements du type $\{\varepsilon_1 = \pm 1, \ldots, \varepsilon_n = \pm 1\}$ sont tous de probabilité $1/2^n$, on en déduit que :

$$P(\varepsilon_1 + \cdots + \varepsilon_n = 2\ell - n) = \frac{1}{2^n}\binom{n}{\ell}$$

Par suite :

$$\begin{aligned} P(|\varepsilon_1 + \cdots + \varepsilon_n| \geqslant an) &= \sum_{\ell=0}^{n} P(\varepsilon_1 + \cdots + \varepsilon_n = 2\ell - n)\, \mathbf{1}_{\{|2\ell-n|\geq an\}} \\ &= \frac{1}{2^n} \sum_{\ell=0}^{n} \binom{n}{\ell}\, \mathbf{1}_{\{|2\ell-n|\geq an\}}. \end{aligned}$$

c) D'après la question précédente, quand n tend vers l'infini, la limite de cette expression vaut 0.

3. Suivant les valeurs de N, nous avons en vertu de l'indépendance de N et de la suite (ε_n), la décomposition :

$$E\big((\sum_{n=1}^{N+1} \varepsilon_n)^2\big) = E\Big(\sum_{k=0}^{\infty} \big(\sum_{n=1}^{k+1} \varepsilon_n\big)^2 \mathbf{1}_{\{N=k\}}\Big) = \sum_{k=0}^{\infty} E\Big(\big(\sum_{n=1}^{k+1} \varepsilon_n\big)^2\Big) P(N=k)$$

Par ailleurs, comme $E\Big(\big(\sum\limits_{n=1}^{k} \varepsilon_n\big)^2\Big) = \sum\limits_{n=1}^{k} V(\epsilon_n) = k$, on obtient :

$$E\Big(\big(\sum_{n=1}^{N+1} \varepsilon_n\big)^2\Big) = \sum_{n=0}^{\infty} (n+1)\, \frac{\theta^n}{n!} \mathrm{e}^{-\theta} = \theta + 1.$$

Exercice 3.9.

Soit $n \geqslant 1$ un entier naturel. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soient X et Y deux variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, P)$ telles que $X(\Omega) = [\![1, n+1]\!]$ et $Y(\Omega) = [\![1, n+1]\!]$.

Pour tout $(i, j) \in [\![1, n+1]\!]^2$, on pose

$$a_{i,j} = P((X = j) \cap (Y = i)) \text{ et } b_{i,j} = P_{(X=j)}(Y = i).$$

1. Dans cette première partie, on suppose qu'il existe un réel λ tel que :

$$\forall\, (i, j) \in [\![1, n+1]\!]^2, a_{i,j} = \lambda \binom{n}{i-1}\binom{n}{j-1}.$$

a) Calculer la valeur de λ.

b) Déterminer les lois marginales de X et de Y.

c) Montrer que les variables X et Y sont indépendantes.

d) Soit $Z = X - 1$. Reconnaître dans la loi de Z une loi usuelle. En déduire l'espérance et la variance de X.

2. On note $B \in \mathcal{M}_{n+1}(\mathbb{R})$ la matrice dont le coefficient de la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne est $b_{i,j}$.

a) Calculer la dimension de l'image et celle du noyau de B.

b) Calculer B^p, pour tout entier $p \in \mathbb{N}^*$.

c) Déterminer l'ensemble des valeurs propres de B. La matrice B est-elle diagonalisable ?

3. On suppose à présent que :

$$\forall\, (i, j) \in [\![1, n+1]\!]^2, a_{i,j} = \begin{cases} \alpha & \text{si } |i + j - n - 2| = 1 \\ 0 & \text{sinon} \end{cases}.$$

a) Déterminer α.

b) Déterminer les lois marginales de X et de Y. Les variables X et Y sont-elles indépendantes ?

c) On note $B \in \mathcal{M}_{n+1}(\mathbb{R})$ la matrice dont le coefficient de la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne est $b_{i,j}$. Écrire B dans le cas particulier $n = 4$.

Solution :

1. a) Pour que les $(a_{i,j})$ définissent une loi conjointe, il faut que pour tout $(i, j) \in [\![1, n+1]\!]^2$, $a_{i,j} \geqslant 0$ et $1 = \sum\limits_{i=1}^{n+1} \sum\limits_{j=1}^{n+1} a_{i,j} = \lambda \sum\limits_{i=1}^{n+1} \binom{n}{i-1} \sum\limits_{j=1}^{n+1} \binom{n}{j-1}$.

En utilisant la formule $\sum\limits_{p=0}^{n} \binom{n}{p} = 2^n$, on trouve : $1 = \lambda.2^n.2^n$. Ainsi, $\lambda = 2^{-2n}$.

b) On a $X(\Omega) = [\![1, n+1]\!]$ et pour tout $j \in [\![1, n+1]\!]$,

$$P(X = j) = \sum_{i=1}^{n+1} a_{i,j} = \lambda\binom{n}{j-1}\sum_{i=1}^{n+1}\binom{n}{i-1} = 2^{-n}\binom{n}{j-1}.$$

Par symétrie, Y suit la même loi ($X - 1$ et $Y - 1$ suivent la loi $\mathcal{B}(n, 1/2)$).

c) Pour tout $(i, j) \in [\![1, n+1]\!]^2$:

$$P(X = j)P(Y = i) = 2^{-n}\binom{n}{j-1}2^{-n}\binom{n}{i-1} = P(X = j \cap Y = i).$$

Les variables X et Y sont donc indépendantes.

d) On l'a dit : $Z \hookrightarrow \mathcal{B}(n, 1/2)$. On en déduit :

$$E(X) = E(Z+1) = E(Z) + 1 = \frac{n}{2} + 1 \text{ et } V(X) = V(Z+1) = V(Z) = \frac{n}{4}.$$

2. a) Comme X et Y sont indépendantes, pour tout $(i, j) \in [\![1, n+1]\!]^2$,

$$b_{i,j} = P_{(X=j)}(Y = i) = P(Y = i) = 2^{-n}\binom{n}{i-1}.$$

On en déduit l'écriture de la matrice B. On remarque en particulier que toutes les colonnes de B sont identiques. Par suite, $\mathrm{Im}(B)$ est de dimension 1. Par le théorème du rang, $\mathrm{Ker}(B)$ est de dimension n.

b) On pose $B^2 = [c_{i,j}]$. Pour tout $(i, j) \in [\![1, n+1]\!]^2$, on a :

$$c_{i,j} = \sum_{k=1}^{n+1} b_{i,k}b_{k,j} = 2^{-2n}\sum_{k=1}^{n+1}\binom{n}{i-1}\binom{n}{k-1} = 2^{-n}\binom{n}{i-1} = b_{i,j}$$

D'où : $B^2 = B$. Par suite, pour tout $p \in \mathbb{N}^*$, $B^p = B$.

c) La question précédente montre que $P(X) = X^2 - X$ est un polynôme annulateur de B. Les seules valeurs propres possibles pour B sont donc 0 et 1. Comme $E_0 = \mathrm{Ker}(B)$ est de dimension $n > 0$, 0 est bien valeur propre de B. On remarque d'autre part que comme $B^2 = B$, on a $BX = X$, où X est le vecteur correspondant à la première colonne de B. Ainsi, 1 est également valeur propre de B et $E_1 = \mathrm{Ker}(B - I_{n+1})$ est de dimension au moins égale à 1. Il s'ensuit que $\dim(E_0) + \dim(E_1) \geqslant n+1$, donc $\dim(E_0) + \dim(E_1) = n+1$. Ainsi, la matrice B est diagonalisable.

On peut aussi dire que B ne vaut ni 0 ni I, donc représente un projecteur non trivial.

3. a) Pour tout $(i, j) \in [\![1, n+1]\!]^2$, on doit avoir $a_{i,j} \geqslant 0$ et $1 = \sum_{i=1}^{n+1}\sum_{j=1}^{n+1} a_{i,j} = N\alpha$, où N désigne le nombre de couples $(i, j) \in [\![1, n+1]\!]^2$ tel que $|i + j - n - 2| = 1$. Facilement $N = 2n$, donc $\alpha = \dfrac{1}{2n}$.

b) $\forall j \in [\![1, n+1]\!]$, $P(X = j) = \sum_{i=1}^{n+1} a_{i,j}$.

Si $j \in [\![2, n]\!]$, on a donc $P(X = j) = a_{n+3-j,j} + a_{n+1-j,j} = \alpha + \alpha = \dfrac{1}{n}$.
D'autre part, $P(X = 1) = P(X = n+1) = \dfrac{1}{2n}$.

On remarque que $P(X=1)P(Y=1)=\dfrac{1}{4n^2}$ alors que $P(X=1\cap Y=1)=0$. Ainsi, X et Y ne sont pas indépendantes.

c) La matrice B s'écrit : $B=\begin{pmatrix} 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 1 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \end{pmatrix}$

Exercice 3.10.

Pour tout entier naturel n, on définit l'ensemble $\mathcal{D}_n$ des densités de probabilité $f:\mathbb{R}\to\mathbb{R}^+$ nulles sur $\mathbb{R}^-$ et telles que $\displaystyle\int_0^{+\infty} x^n f(x)dx$ converge. On note alors $m_n(f)$ la valeur de cette intégrale. On note enfin : $\mathcal{D}_\infty=\bigcap\limits_{k=0}^{+\infty}\mathcal{D}_k$

1. Soit n un entier naturel. Déterminer tous les couples de réels positifs (α,β) tels que :

$$f_1 \text{ et } f_2\in\mathcal{D}_n \Longrightarrow \alpha f_1+\beta f_2\in\mathcal{D}_n$$

2. a) Soit $f\in\mathcal{D}_n$.
Montrer que pour tout entier k tel que $0\leqslant k\leqslant n$, on a :

$$f\in\mathcal{D}_k \text{ et } 0<m_k(f)\leqslant 1+m_n(f).$$

b) Soit la fonction $c:x\in\mathbb{R}\mapsto\begin{cases}\dfrac{2}{\pi}\times\dfrac{1}{x^2+1} & \text{si } x>0\\ 0 & \text{si } x\leqslant 0\end{cases}$. Montrer que $c\in\mathcal{D}_0\setminus\mathcal{D}_1$.

Dans toute la suite, on note $T:\mathcal{D}_1\to\mathcal{D}_0$, l'application définie pour tout $f\in\mathcal{D}_0$ par $T(f)(x)=\dfrac{x}{m_1(f)}f$.

3. On note $f_{\mathcal{E},\lambda}$ la densité de probabilité d'une loi exponentielle de paramètre $\lambda>0$, nulle sur $\mathbb{R}^-$ et continue sur $\mathbb{R}_+^*$. Montrer que $f_{\mathcal{E},\lambda}\in\mathcal{D}_\infty$ et calculer $T(f_{\mathcal{E},\lambda})$.

4. Soit α et β deux réels positifs tels que $\alpha+\beta=1$ et $f,g\in\mathcal{D}_1$. Exprimer $T(\alpha f+\beta g)$ en fonction de $T(f)$ et de $T(g)$.

Soit $f_0\in\mathcal{D}_\infty$. On définit pour tout entier $n\geqslant 0$: $f_{n+1}=T(f_n)$.

5. Dans le cas général, exprimer f_n en fonction de f_0 et $m_n(f)$ puis calculer et reconnaître f_n lorsque f_0 est la fonction $f_{\mathcal{E},\lambda}$.

6. Soit X et Y deux variables aléatoires continues de densités respectives $f\in\mathcal{D}_1$ pour X et $T(f)$ pour Y.
On note F et G leurs fonctions de répartition respectives.

Comparer F et G. Dans quelle portion du plan est situé l'ensemble des points $\{(F(t),G(t)),t\in\mathbb{R}\}$?

Solution :

1. Comme $\alpha\geqslant 0$ et $\beta\geqslant 0$, la seule condition à réaliser est $\displaystyle\int_0^{+\infty}\alpha f_1+\beta f_2=1$.
Les couples solutions sont : $(\alpha,1-\alpha),\alpha\in[0,1]$ et $(\mathcal{D}_n,+,.)$ n'est pas un espace vectoriel sur $\mathbb{R}$.

2. a) Soit $f\in\mathcal{D}_n$. Comme pour $x\geqslant 1$, on a $k\leqslant n\implies x^kf(x)\leqslant x^nf(x)$, l'existence du moment d'ordre n entraîne celle du moment d'ordre k.
On a $\displaystyle\int_0^{+\infty}f(x)dx=1$, donc il existe un intervalle non vide $[a,b]\subset\mathbb{R}_+$ sur lequel $f>0$ et $\displaystyle 0<\int_a^b x^kf(x)\,dx\leqslant\int_0^{+\infty}x^kf(x)\,dx$, donc $m_k>0$. Enfin :

$$\begin{aligned}\int_0^{+\infty}x^kf(x)dx&\leqslant\int_0^1 f(x)dx+\int_1^{+\infty}x^kf(x)dx\\&\leqslant\int_0^{+\infty}f(x)dx+\int_0^{+\infty}x^nf(x)dx.\end{aligned}$$

Donc $0<m_k\leqslant 1+m_n$.

b) La fonction $c:x\mapsto\dfrac{2}{\pi}\times\dfrac{1}{x^2+1}\,\mathbf{1}_{\mathbb{R}_+}(x)$ est continue et positive et :

$$\int_{-\infty}^{+\infty}c(x)dx=\frac{2}{\pi}\int_0^{+\infty}\frac{dx}{x^2+1}=\big[\frac{2}{\pi}\operatorname{Arc\,tan}x\big]_0^{\to+\infty}=1.$$

En revanche $x\times c(x)\underset{(+\infty)}{\sim}\dfrac{2}{\pi}\times\dfrac{1}{x}$ et l'intégrale $\displaystyle\int_{-\infty}^{+\infty}xc(x)\,dx$ diverge, donc $c\in\mathcal{D}_0$, mais $c\notin\mathcal{D}_1$.

3. On a $f\in\mathcal{D}_1\implies m_1>0$: $T(f)$ est bien défini et on vérifie aisément que $T(f)$ appartient bien à $\mathcal{D}_0$.
Or $f_{\mathcal{E},\lambda}:x\mapsto\lambda\mathrm{e}^{-\lambda x}\mathbf{1}_{\mathbb{R}_+}$ est continue et positive sur $\mathbb{R}_+$.
Soit $k\in\mathbb{N},x^kf_{\mathcal{E},\lambda}(x)=o\big(\dfrac{1}{x^2}\big)$ au voisinage de $+\infty$, d'où l'existence de m_n.
Donc $f_{\mathcal{E},\lambda}\in\mathcal{D}_\infty$ et comme $m_1=\dfrac{1}{\lambda}$: $T(f_{\mathcal{E},\lambda})=\lambda^2x\mathrm{e}^{-\lambda x}\mathbf{1}_{\mathbb{R}_+}$.

4. Soit $\alpha\geqslant 0$ et $\beta\geqslant 0$ tels que $\alpha+\beta=1$ et $f,g\in\mathcal{D}_1$. On a :

$$T(\alpha f+\beta g)=\frac{\alpha m_1(f)}{\alpha m_1(f)+\beta m_1(g)}T(f)+\frac{\beta m_1(g)}{\alpha m_1(f)+\beta m_1(g)}T(g)$$

5. On a : $f_1=\dfrac{x}{m_1(f)}f_0$. Soit $n\geqslant 0$, on suppose que $f_n=\dfrac{x^n}{m_n(f)}f_0$; alors

$m_1(f_n) = \displaystyle\int_0^{+\infty} x f_n(x)dx = \int_0^{+\infty} \frac{x^{n+1}}{m_n(f)} f_0(x)dx = \frac{m_n+1(f)}{m_n(f)}$ et :

$$f_{n+1} = \frac{x m_n(f)}{m_{n+1}(f)} \times \frac{x^n}{m_n(f)} f_0 = \frac{x^{n+1}}{m_{n+1}(f)} f_0.$$

Pour la loi exponentielle de paramètre λ :

$$m_n(f) = \int_0^{+\infty} x^n \lambda \mathrm{e}^{-\lambda x} dx = \frac{1}{\lambda^n} \int_0^{+\infty} u^n \mathrm{e}^{-u} du = \frac{n!}{\lambda^n}.$$

et $f_n(x) = \dfrac{x^n}{n!} \lambda \mathrm{e}^{-\lambda x} \mathbf{1}_{[0,+\infty[}(x)$, qui est une densité de la loi $\gamma(\lambda, n+1)$

6. Soit $f \in \mathcal{D}_1$,

$$F(x) = \mathbf{1}_{\mathbb{R}^+}(x) \int_0^x f(t)dt \text{ et } T(f) \in \mathcal{D}_0,\ G(x) = \mathbf{1}_{\mathbb{R}^+(x)} \int_0^x \frac{t}{m_1(f)} f(t)dt.$$

$\rightarrow$ Si $x \leqslant m_1(f)$, on a :$F(x) - G(x) = \displaystyle\int_0^x \frac{m_1 - t}{m_1} f(t)dt \geqslant 0$ car $0 \leqslant t \leqslant x \leqslant m_1$.

$\rightarrow$ Si $x \geqslant m_1(f) > 0$, on a :

$$\begin{aligned} F(x) - G(x) &= \big(1 - \int_x^{+\infty} f(t)dt\big) - \big(\frac{1}{m_1}[m_1 - \int_x^{+\infty} t f(t)dt]\big) \\ &= \int_x^{+\infty} \frac{t - m_1}{m_1} f(t)dt \geqslant 0. \text{ D'où } F \geqslant G. \end{aligned}$$

On a ainsi : $\forall t \in \mathbb{R}, 0 \leqslant G(t) \leqslant F(t) \leqslant 1$: la courbe est située dans le triangle de sommets O, A, B, avec $A : (1,0)$ et $B : (1,1)$.

Exercice 3.11.

Une grenouille se déplace par bonds mutuellement *indépendants de longueur fixe* 1 en restant dans un plan muni d'un repère orthonormé $(O, \overrightarrow{i}, \overrightarrow{j})$.
On note (X_n, Y_n) ses coordonnées après n bonds et $D_n = \sqrt{X_n^2 + Y_n^2}$ sa distance à l'origine. Au départ de l'histoire elle est en O, donc : $X_0 = Y_0 = 0$.

1. Dans cette question uniquement, la grenouille reste sur la droite $(O, \vec{i})$ (donc $\forall n \geqslant 0, Y_n = 0$). Elle fait des bonds $x_i = X_i - X_{i-1}, (i \geqslant 1)$ successifs vers la droite (*i.e.* tels que $x_i = 1$) avec la probabilité constante $p \in [0,1]$ ou dans l'autre sens (*i.e.* $x_i = -1$) avec la probabilité $q = 1 - p$.

a) Calculer la probabilité qu'elle soit revenue à son point de départ après $2n$ bonds.

b) Déterminer la loi de la variable aléatoire D_n.

c) Déterminer deux réels a et b tels que $z_i = a x_i + b$ suive une loi de Bernoulli pour tout $i \geqslant 1$. On note $Z_n = \sum\limits_{i=1}^{n} z_i$. Exprimer X_n à l'aide de Z_n puis calculer les moments d'ordre 1 et 2 des variables aléatoires Z_n et X_n.

2. La grenouille s'enhardit : elle effectue des bonds dans une direction quelconque. On note θ_n l'angle que fait le $n^{\text{ème}}$ saut avec l'axe $(O\vec{i})$. On suppose que pour tout n, θ_n suit la uniforme sur $[0, 2\pi]$.

a) Exprimer X_n et Y_n en fonction des θ_i, $1 \leqslant i \leqslant n$ et calculer leurs moments d'ordre 1 et 2.

b) Montrer que : $\begin{cases} E(X_{n-1}Y_{n-1}^2\cos(\theta_n)) = E(X_{n-1}^2 Y_{n-1}\sin(\theta_n)) = 0 \\ E(Y_{n-1}\cos^2(\theta_n)\sin(\theta_n)) = E(X_{n-1}\cos(\theta_n)\sin^2(\theta_n)) = 0 \end{cases}$

c) En déduire $\mathrm{Cov}(X_n, Y_n)$ et $\mathrm{Cov}(X_n^2, Y_n^2)$. Les variables X_n et Y_n sont-elles indépendantes ?

Solution :

1. a) $[X_{2n} = 0]$ signifie qu'il y a eu n sauts à droite et n à gauche. Donc par application du schéma binomial :

$$P(X_{2n} = 0) = \binom{2n}{n} p^n q^n = \frac{2n!}{n!n!}\, p^n q^n$$

b) On a $X_n(\Omega) \subset [\![-n, n]\!]$, et $[X_n = k]$ est réalisé si et seulement si le nombre de sauts à droite (nsd) est égal au nombre de sauts à gauche (nsg) $+k$, et comme $nsg + nsd = n$ on en déduit : $nsg = \dfrac{n-k}{2}$ et $nsd = \dfrac{n+k}{2}$. Ces 2 quantités doivent être des entiers d'où :

$$P(X_n = k) = \begin{cases} \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}} & \text{si } k \text{ et } n \text{ de même parité} \\ 0 & \text{sinon} \end{cases}$$

Par conséquent $D_n(\Omega) \subset [\![0, n]\!]$ et :

- $P(D_n = 0) = P(X_n = 0) = \begin{cases} \binom{n}{\frac{n}{2}} p^{\frac{n}{2}} q^{\frac{n}{2}} & \text{si } n \text{ pair} \\ 0 & \text{si } n \text{ impair} \end{cases}$.

- Pour $0 < k \leqslant n$: $P(D_n = k) = P(X_n = k) + P(X_n = -k)$, soit par symétrie des coefficients binomiaux :

$$P(D_n = k) = \begin{cases} \binom{n}{\frac{n+k}{2}} \left(p^{\frac{n+k}{2}} q^{\frac{n-k}{2}} + p^{\frac{n-k}{2}} q^{\frac{n+k}{2}}\right) & \text{si } k \text{ et } n \text{ de même parité} \\ 0 & \text{sinon} \end{cases}$$

c) $ax_i + b$ prend les valeurs $b - a$ et $b + a$, on peut donc prendre $a = b = \dfrac{1}{2}$ et $z_i = ax_i + b$ suit alors la loi de Bernoulli de paramètre p.

On a $X_n = \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} (2z_i - 1) = 2Z_n - n$ avec $Z_n \hookrightarrow \mathcal{B}(n, p)$. Ainsi :

$$E(X_n) = 2E(Z_n) - n = 2np - n\,;\; V(X_n) = 4V(Z_n) = 4npq$$

2. a) On a $X_n = \sum\limits_{i=1}^{n} \cos(\theta_i)$ et $Y_n = \sum\limits_{i=1}^{n} \sin(\theta_i)$. Par le théorème de transfert :

$E(\cos(\theta_i)) = \displaystyle\int_0^{2\pi} \frac{1}{2\pi}\cos(\theta)d\theta = 0$ et de même $E(\sin(\theta_i)) = 0$. D'où :

$$\forall\, n \in \mathbb{N}, E(X_n) = E(Y_n) = 0$$

De même :

$V(\cos(\theta_i)) = E(\cos^2(\theta_i)) = \displaystyle\int_0^{2\pi} \frac{1}{2\pi}\cos^2(\theta)d\theta = \frac{1}{4\pi}\int_0^{2\pi}(1+\cos(2\theta))\,d\theta = \frac{1}{2}$;

d'où, par indépendance $V(X_n) = \dfrac{n}{2}$. De la même façon $V(Y_n) = \dfrac{n}{2}$

b) Les variables aléatoires $X_{n-1}Y_{n-1}^2$ et $X_{n-1}^2Y_{n-1}$ sont des fonctions des θ_i pour $i \in [\![1, n-1]\!]$, donc sont indépendantes de θ_n. Ainsi :

$$\begin{cases} E(X_{n-1}Y_{n-1}^2\cos(\theta_n)) = E(X_{n-1}Y_{n-1}^2)E(\cos(\theta_n)) = 0 \\ E(X_{n-1}^2Y_{n-1}\sin(\theta_n)) = 0 \end{cases}$$

On a : $E(\cos^2(\theta_n)\sin(\theta_n)) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}\cos^2\theta\sin\theta\,d\theta = -\frac{1}{6\pi}\big[\cos^3\theta\big]_0^{2\pi} = 0$, et de la même façon : $E(\cos(\theta_n)\sin^2(\theta_n)) = 0$. Donc encore par indépendance :

$$E(Y_{n-1}\cos^2(\theta_n)\sin(\theta_n)) = 0 \text{ et } E(X_{n-1}\cos(\theta_n)\sin^2(\theta_n)) = 0.$$

c) On a :

$\mathrm{Cov}(X_n, Y_n) = \displaystyle\sum_{1\le i,j\le n}\mathrm{Cov}(\cos(\theta_i), \sin(\theta_j)) = \sum_{i=1}^{n}\mathrm{Cov}(\cos(\theta_i), \sin(\theta_i))$

Or $\mathrm{Cov}(\cos\theta_i, \sin\theta_i) = E(\cos\theta_i\sin\theta_i) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}\cos\theta\sin\theta\,d\theta = 0$ et donc $\mathrm{Cov}(X_n, Y_n) = 0$.

On a : $E(X_n^2Y_n^2) = E\big((X_{n-1} + \cos(\theta_n))^2(Y_{n-1} + \sin(\theta_n))^2\big)$

On développe, on utilise la linéarité de l'espérance, l'indépendance de θ_n avec X_{n-1} et Y_{n-1}, et les résultats précédents. Il reste :

$$\begin{aligned} E(X_n^2Y_n^2) &= E(X_{n-1}^2Y_{n-1}^2) + \frac{n-1}{2} + E(\cos^2(\theta_n)\sin^2(\theta_n)) \\ &= E(X_{n-1}^2Y_{n-1}^2) + \frac{n-1}{2} + \frac{1}{8}, \text{ et en sommant :} \end{aligned}$$

$$E(X_n^2Y_n^2) = \frac{n(n-1)}{4} + \frac{n}{8} = \frac{n(2n-1)}{8}$$

$$\mathrm{Cov}(X_n^2, Y_n^2) = E(X_n^2Y_n^2) - E(X_n^2)E(Y_n^2) = -\frac{n}{8} \neq 0$$

Si les variables X_n et Y_n étaient indépendantes il en serait de même de X_n^2 et de Y_n^2, or celles-ci sont corrélées, donc X_n et Y_n ne sont pas indépendantes.

Exercice 3.12.

Soit X une variable aléatoire prenant un nombre fini de valeurs $x_1, x_2, \ldots, x_n$ avec les probabilités respectives $p_1, p_2, \ldots, p_n$.

On définit la fonction φ_X sur $\mathbb{R}^*$ par :

$$\forall\, t \in \mathbb{R}^*, \varphi_X(t) = \frac{1}{t}\ln(E(\mathrm{e}^{tX}))$$

où E désigne l'opérateur espérance.

1. Montrer que φ_X est ainsi bien définie sur $\mathbb{R}^*$ et prolongeable par continuité en 0, en posant $\varphi_X(0) = E(X)$.
On note encore φ_X la fonction ainsi prolongée.

2. Démontrer que φ_X est dérivable en 0 et calculer $\varphi'_X(0)$ à l'aide de la variance de X.

3. a) Montrer que pour tout $u \leqslant 0$, on a : $\mathrm{e}^u \leqslant 1 + u + \frac{1}{2}u^2$.

b) Montrer que si X ne prend que des valeurs négatives ou nulles, on a :

$$\forall\, t \geqslant 0, \varphi_X(t) \leqslant E(X) + \frac{t}{2}E(X^2)$$

4. a) Pour tout $i \in [\![1, n]\!]$, on note f_i la fonction $t \mapsto \mathrm{e}^{tx_i}$. Montrer que la famille $(f_1, \ldots, f_n)$ est libre.

b) En déduire que deux variables discrètes finies X et Y ont la même loi si et seulement si les fonctions φ_X et φ_Y sont égales.

5. Montrer que si X et Y sont des variables discrètes finies indépendantes, alors $\varphi_{X+Y} = \varphi_X + \varphi_Y$.

6. Montrer que φ_X est impaire si et seulement si X est une variable symétrique, c'est-à-dire telle que X et $-X$ ont la même loi.

Solution :

1. e^{tX} ne prend que des valeurs strictement positives (en nombre fini) donc son espérance existe et est strictement positive, ce qui prouve que φ_X est bien définie sur $\mathbb{R}^*$.
On a $E(\mathrm{e}^{tX}) = \sum\limits_{i=1}^{n} \mathrm{e}^{tx_i} p_i$. Or : $\mathrm{e}^{tx_i} = 1 + tx_i + o(t)$, donc

$$E(\mathrm{e}^{tX}) = \sum_{i=1}^{n} \mathrm{e}^{tx_i} p_i = \sum_{i=1}^{n} p_i + \sum_{i=1}^{n} tx_i p_i + o(t) = 1 + \sum_{i=1}^{n} tx_i p_i + o(t)$$

Ainsi :

$$\varphi_X(t) = \frac{1}{t}\ln(1 + \sum_{i=1}^{n} tx_i p_i + o(t)) \underset{(t\to 0)}{\sim} \frac{1}{t}\sum_{i=1}^{n} tx_i p_i$$

et $\varphi_X(t) \underset{(t\to 0)}{\longrightarrow} \sum\limits_{i=1}^{n} x_i p_i = E(X)$.

2. On poursuit les développements asymptotiques :

$$\begin{aligned} E(\mathrm{e}^{tX}) &= \sum_{i=1}^{n} \mathrm{e}^{tx_i} p_i = 1 + \sum_{i=1}^{n} tx_i p_i + \frac{1}{2}\sum_{i=1}^{n} t^2 x_i^2 p_i + o(t^2) \\ &= 1 + E(X)t + \frac{E(X^2)}{2}t^2 + o(t^2) \end{aligned}$$

et avec $\ln(1+u) = u - \dfrac{u^2}{2} + o(u^2)$, il vient :

$$\begin{aligned}\varphi_X(t) &= \frac{1}{t}\ln(1+E(X)t + \frac{E(X^2)}{2}t^2 + o(t^2)) \\ &= E(X) + \frac{E(X^2)-E(X)^2}{2}t + o(t)\end{aligned}$$

Ainsi φ_X est dérivable en 0, avec $\varphi_X'(0) = \frac{1}{2}V(X)$.

3. a) la fonction $u \mapsto \mathrm{e}^u - 1 - u - \dfrac{1}{2}u^2$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^-$ et $f'(u) = \mathrm{e}^u - 1 - u$, $f''(u) = \mathrm{e}^u - 1 \leqslant 0$. Avec $f'(0) = 0$, on en déduit que f' est positive sur $\mathbb{R}^-$ et comme $f(0) = 0$, f est bien négative sur $\mathbb{R}^-$.

b) On peut donc écrire dans ce cas :

$$E(\mathrm{e}^{tX}) = \sum \mathrm{e}^{tx_i}p_i \leqslant (1 + t\sum x_i + \frac{1}{2}t^2\sum x_i^2)p_i = 1 + tE(X) + \frac{1}{2}t^2E(X^2)$$

et donc :

$$\varphi_X(t) \leqslant \frac{1}{t}\ln(1 + tE(X) + \frac{1}{2}t^2E(X^2)) \leqslant E(X) + \frac{t}{2}E(X^2))$$

Car on sait que $\ln(1+u) \leqslant u$.

4. a) L'application $f \mapsto f'$ est un endomorphisme de l'espace des fonctions de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et l'application $t \mapsto \mathrm{e}^{tx_i}$ est propre pour la valeur propre x_i, d'où la liberté demandée.

b) $\star$ Si X et Y ont même loi, il est clair que $\varphi_X = \varphi_Y$.

$\star$ Réciproquement supposons que X et Y soient discrètes finies telles que $\varphi_X = \varphi_Y$. Quitte à ajouter des valeurs prises avec une probabilité nulle, on peut supposer que $X(\Omega) = Y(\Omega) = \{x_1, \dots, x_n\}$ et on pose $p_i = P(X = x_i), q_i = P(Y = x_i)$. On a alors $\forall\, t \neq 0$, $\sum\limits_{i=1}^n \mathrm{e}^{tx_i}p_i = \sum\limits_{i=1}^n \mathrm{e}^{tx_i}q_i$, résultat bien entendu valable pour $t = 0$. La liberté précédente donne $\forall\, i, p_i = q_i$: X et Y ont même loi.

5. Si X et Y sont indépendantes, il en est de même de e^{tX} et e^{tY} et :

$$E(\mathrm{e}^{t(X+Y)}) = E(\mathrm{e}^{tX}\mathrm{e}^{tY}) = E(\mathrm{e}^{tX})E(\mathrm{e}^{tY})$$

D'où, par passage au logarithme : $\varphi_{X+Y} = \varphi_X + \varphi_Y$.

6. φ_X impaire $\Longleftrightarrow \varphi_X(t) = -\varphi_X(-t)$, or le retour à la relation de définition donne : $-\varphi_X(-t) = \varphi_{-X}(t)$

Donc φ_X impaire $\Longleftrightarrow \varphi_X = \varphi_{-X} \Longleftrightarrow X$ et $-X$ ont même loi (résultat 4. b)).

Exercice 3.13.

On note $VP(\lambda, \theta)$ la loi de Pareto de paramètres $\lambda > 0, \theta > 0$ et 0 , c'est-à-dire la loi de densité f définie par : $f(x) = \dfrac{\lambda}{\theta}\Big(\dfrac{\theta}{x}\Big)^{\lambda+1}$ si $x > \theta$, et $f(x) = 0$ sinon.

Un phénomène économique suit une loi $VP(\lambda, \theta)$, θ étant connu et λ un paramètre que l'on veut estimer. On dispose pour cela d'un échantillon $(X_1, \ldots, X_n)$ de variables aléatoires indépendantes suivant cette loi.

On pose $T = \sum_{i=1}^{n} \ln\left(\frac{X_i}{\theta}\right)$ et $\widehat{\lambda} = \frac{n}{T}$.

1. Soit X une variable aléatoire suivant la loi $VP(\lambda, \theta)$ et $Y = \ln\left(\frac{X}{\theta}\right)$.

a) Déterminer la fonction de répartition F_X de X .

b) Déterminer, lorsqu'elles existent, l'espérance et la variance de X .

c) Montrer que Y suit une loi Γ dont on précisera les paramètres.

2. Quelle est la loi de T ? En donner une densité.

3. Calculer l'espérance et la variance de $\widehat{\lambda}$.

4. Déduire de $\widehat{\lambda}$ un estimateur $\widehat{\lambda_1}$ sans biais de λ. L'estimateur $\widehat{\lambda_1}$ est-il convergent ?

5. On admet que la suite de variables aléatoires $n \mapsto Z_n = \dfrac{\sqrt{n}(\widehat{\lambda} - \lambda)}{\lambda}$ converge en loi vers la loi normale centrée réduite $\mathcal{N}(0,1)$, et on rappelle que la fonction de répartition Φ de la loi normale centrée réduite vérifie $\Phi(1,96) \simeq 0{,}975$.

En utilisant la loi $\mathcal{N}(0,1)$ comme approximation de la loi de Z_n, donner, en fonction de n (supposé assez grand) et de la valeur observée λ_0 de $\widehat{\lambda}$, un intervalle de confiance à 95% de λ.

Solution :

1. a) Pour $x \leqslant \theta$, $F_X(x) = 0$ et pour $x > \theta$, $F_X(x) = 1 - \left(\frac{\theta}{x}\right)^{\lambda}$

b) La variable aléatoire X admet une espérance si et seulement si $\lambda > 1$ et dans ce cas : $E(X) = \dfrac{\lambda\theta}{\lambda - 1}$.
De même, la variable aléatoire X admet une variance si et seulement si $\lambda > 2$ et alors : $V(X) = \dfrac{\lambda\theta^2}{(\lambda-2)(\lambda-1)^2}$.

c) La variable aléatoire Y prend ses valeurs dans $\mathbb{R}_+^*$ et :

$$\forall\, x > 0, F_Y(x) = P(X \leqslant \theta.\mathrm{e}^x) = F_X(\theta.\mathrm{e}^x) = 1 - \mathrm{e}^{-\lambda x}$$

Donc Y suit la loi exponentielle de paramètre λ, c'est-à-dire la loi Γ de paramètres $\dfrac{1}{\lambda}$ et 1.

2. La variable aléatoire T est la somme de n variables aléatoires indépendantes de même loi $\Gamma(\frac{1}{\lambda}, 1)$, donc T suit la loi $\Gamma(\frac{1}{\lambda}, n)$. Une densité f_T de T est donnée par :

$$\forall t \leqslant 0, f_T(x) = 0, \forall t > 0, f_T(x) = \frac{\mathrm{e}^{-\lambda x} x^{n-1} \lambda^n}{(n-1)!}$$

3. $\star$ On a $\widehat{\lambda} = \frac{n}{T}$, donc sous réserve d'existence :

$$E(\widehat{\lambda}) = \int_{-\infty}^{+\infty} \frac{n}{x} f_T(x)\, dx = \frac{n\lambda}{n-1} \int_0^{+\infty} \frac{\mathrm{e}^{-\lambda x} x^{n-2} \lambda^{n-1}}{(n-2)!}\, dx$$

On reconnaît l'intégrale d'une densité d'une variable suivant la loi $\Gamma(\frac{1}{\lambda}, n-1)$, ce qui prouve que l'espérance existe et vaut :

$$E(\widehat{\lambda}) = \frac{n\lambda}{n-1}$$

$\star$ En procédant de la même façon on montre que $\widehat{\lambda}$ admet un moment d'ordre 2 valant $\frac{n^2\lambda^2}{(n-1)(n-2)}$, donc une variance valant : $V(\widehat{\lambda}) = \frac{n^2\lambda^2}{(n-1)^2(n-2)}$

4. $\widehat{\lambda}_1 = \frac{n-1}{n}\widehat{\lambda}$ vérifie $E(\widehat{\lambda}_1) = \lambda$ et $V(\widehat{\lambda}_1) = \frac{\lambda^2}{n-2}$. Cette variance est de limite nulle lorsque n tend vers l'infini, donc $\widehat{\lambda}_1$ est un estimateur sans biais et convergent de λ.

5. En approchant la loi de Z par la loi normale centrée réduite, on obtient :

$$P(-1,96 \leqslant \frac{\sqrt{n}(\widehat{\lambda} - \lambda)}{\lambda} \leqslant 1,96) = 0,95$$

C'est-à-dire : $P(\frac{\sqrt{n}}{\sqrt{n}+1,96}\widehat{\lambda} \leqslant \lambda \leqslant \frac{\sqrt{n}}{\sqrt{n}-1,96}\widehat{\lambda}) = 0,95$

On obtient donc comme intervalle de confiance pour λ, au niveau de confiance $0,95$:

$$[\frac{\sqrt{n}}{\sqrt{n}+1,96}\lambda_0, \frac{\sqrt{n}}{\sqrt{n}-1,96}\lambda_0]$$

Exercice 3.14.

Soit α et λ deux réels strictement positifs et f la fonction définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} \lambda\alpha x^{\alpha-1}\mathrm{e}^{-\lambda x^\alpha} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

1. Montrer que f est une fonction de densité.

On note X une variable aléatoire admettant f pour densité. On dit alors que X suit la loi $W(\alpha, \lambda)$)

2. Calculer la fonction de répartition F de X, ainsi que la fonction r définie sur $\mathbb{R}^+$ par $r(x) = \frac{f(x)}{1-F(x)}$.

3. Dans cette question, on suppose que $\alpha > 1$ et que X suit la loi $W(\alpha, \lambda)$.

a) Montrer que la fonction r est strictement croissante sur $\mathbb{R}^+$ et vérifie $r(0) = 0$

b) Montrer que la variable aléatoire $Y = r(X)$ suit la loi $W(\alpha', \lambda')$ avec $\dfrac{1}{\alpha} + \dfrac{1}{\alpha'} = 1$. Exprimer λ' en fonction de λ, α.

4. Réciproquement, on suppose que la fonction r vérifie les deux propriétés :

• la fonction est continue, strictement croissante sur $\mathbb{R}^+$ et vérifie $r(0) = 0$,

• la variable aléatoire $Y = r(X)$ suit la loi $W(\alpha', \lambda')$, avec $\alpha' > 1$.

Montrer que X suit la loi $W(\alpha, \lambda)$, le lien entre (α, λ) et (α', λ') étant donné dans la question précédente.

Solution :

1. La fonction f est continue sur $\mathbb{R}^{+*}$ et sur $\mathbb{R}^-$. Elle est positive sur $\mathbb{R}$ et

$$\int_{-\infty}^{+\infty} f(t)dt = \int_0^{+\infty} \lambda\alpha t^{\alpha-1} \mathrm{e}^{-\lambda x^\alpha} dt = \big[- \mathrm{e}^{-\lambda x^\alpha} \big]_0^\infty = 1$$

2. Pour $x \leqslant 0$, on a $F(x) = 0$ et pour $x > 0$,

$$F(x) = \int_0^x \lambda\alpha t^{\alpha-1} \mathrm{e}^{-\lambda x^\alpha} dt = \big[- \mathrm{e}^{-\lambda x^\alpha} \big]_0^x = 1 - \mathrm{e}^{-\lambda x^\alpha}$$

Et donc, pour $x > 0$: $r(x) = \dfrac{f(x)}{1 - F(x)} = \lambda\alpha x^{\alpha-1}$

3. a) On suppose que $\alpha > 1$, ce qui fait que la fonction f est continue sur $\mathbb{R}$. Par la question précédente, $r(x) = \lambda\alpha x^{\alpha-1}$ est strictement positive sur $\mathbb{R}^+$, vérifie $r(0) = 0$ et $r'(x) = \lambda\alpha(\alpha - 1)x^{\alpha-2} > 0$ sur $\mathbb{R}_+^*$.

b) Posons $Y = r(X)$. La fonction de répartition G de Y est nulle sur $\mathbb{R}^-$ et pour $x > 0$: $G(x) = P(Y \leqslant x) = P(r(X) \leqslant x) = P(X \leqslant r^{-1}(x)) = F(r^{-1}(x))$

puisque la fonction r est bijective strictement croissante sur $\mathbb{R}^+$.

On montre aisément que $r^{-1}(x) = \big(\dfrac{x}{\lambda\alpha}\big)^{1/(\alpha-1)}$. Ainsi :

$$G(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - \exp\big(- \lambda(\dfrac{x^{\alpha/(\alpha-1)}}{(\lambda\alpha)^{\alpha/(\alpha-1)}})\big) & \text{sinon} \end{cases}$$

On a donc $\alpha' = \dfrac{\alpha}{\alpha - 1}$ et $\lambda' = \dfrac{\lambda}{(\lambda\alpha)^{\alpha'}}$. On vérifie que l'on a bien : $\dfrac{1}{\alpha} + \dfrac{1}{\alpha'} = 1$.

4. Notons $R : x \mapsto \displaystyle\int_0^x r(t)dt$. Cette fonction est de classe C^1 sur $\mathbb{R}^+$ et $R'(x) = r(x)$. Ainsi :

$r(x) = \dfrac{F'(x)}{1-F(x)} \implies R'(x) = -\dfrac{d}{dx}(\ln(1-F(x))$ et $R(x) = \ln(1-F(x))+K$.

La condition initiale donne $K = 0$ et :

$$F(x) = 1 - \mathrm{e}^{-R(x)} \qquad (\star)$$

Par stricte croissance de r, pour $x > 0$:

$$1 - \mathrm{e}^{\lambda' x^{\alpha'}} = P(r(X) \leqslant x) = P(X \leqslant r^{-1}(x))$$

et donc :

$$F(x) = P(X \leqslant x) = 1 - \mathrm{e}^{\lambda' (r(x))^{\alpha'}} \qquad (\star\star)$$

En utilisant les relations $(\star)$ et $(\star\star)$, il vient :

$$R(x) = \lambda'(r(x))^{\alpha'} = \lambda'(R'(x))^{\alpha'}, \text{ soit } R'(x)R(x)^{-1/\alpha'} = \Big(\frac{1}{\lambda'}\Big)^{1/\alpha'}$$

En intégrant, il vient, comme $R(0) = 0$:

$$\frac{1}{1-\frac{1}{\alpha'}} R^{\frac{1}{1-1/\alpha'}} = \frac{1}{\lambda'^{1/\alpha'}} x$$

Comme $\dfrac{1}{\alpha} + \dfrac{1}{\alpha'} = 1$, cela est équivalent à : $R^{1/\alpha}(x) = \dfrac{1}{\alpha(\lambda')^{1/\alpha'}} x$ et

$$R(x) = \frac{1}{(\alpha \lambda'^{1/\alpha'})^{\alpha}} x^{\alpha}$$

On termine en utilisant la relation $(\star)$: $F(x) = 1 - \mathrm{e}^{-R(x)} = 1 - \exp(\lambda x^{\alpha})$.

Exercice 3.15.

Soit $(A_n)_n$ une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, P)$. On note $p_n = P(A_n)$. On note B l'événement $\bigcap\limits_{n\geq 1} \Big(\bigcup\limits_{k \geq n} A_k \Big)$.

On rappelle que cet événement est en fait :

$$B = \{\omega \in \Omega \mid \omega \text{ appartient à une infinité des } A_n\}.$$

1. On suppose que la série $\sum\limits_k P(A_k)$ converge. Montrer que $P(B) = 0$.

2. On suppose que les événements (A_n) sont indépendants et que la série $\sum\limits_n P(A_n)$ est divergente.

a) Montrer que l'événement $\overline{B}$ est égal à $\bigcup\limits_{n\geq 1} \Big(\bigcap\limits_{k \geq n} \overline{A_k} \Big)$, où $\overline{M}$ désigne l'événement contraire de l'événement M.

b) Exprimer $P\Big(\bigcap\limits_{k=n}^{m} \overline{A_k} \Big)$ en fonction des p_k.

c) Montrer que la série $\sum\limits_k \ln(1-p_k)$ est divergente.

d) En déduire que $P(B) = 1$.

3. Soit α un réel strictement positif et (X_n) une suite de variables aléatoires indépendantes telles que pour tout n, X_n suit la loi de Bernoulli de paramètre $\dfrac{1}{n^\alpha}$.

a) Montrer que $\lim\limits_{n\to+\infty} E(X_n) = 0$.

b) On suppose que $0 < \alpha < 1$. Montrer qu'avec une probabilité égale à 1, l'ensemble $\{n \mid X_n = 1\}$ contient une infinité d'éléments.

c) On suppose que $\alpha > 1$. Montrer qu'avec une probabilité égale à 1, l'ensemble $\{n \mid X_n = 1\}$ est fini.

Solution :

1. Pour tout entier m, on a : $\bigcap\limits_{n=1}^{m} \bigcup\limits_{k\geq n} A_k = \bigcup\limits_{k=m}^{\infty} A_k$.

La série $\sum P(A_n)$ est convergente. Donc $\lim\limits_{m\to+\infty} \sum\limits_{k=m}^{\infty} P(A_k) = 0$.
Or $P(A\cup B) = P(A) + P(B) - P(A\cap B) \leqslant P(A) + P(B)$, donne par une récurrence simple $P(A_m\cup A_{m+1}\cup \ldots \cup A_n) \leqslant P(A_m)+P(A_{m+1})+\cdots+P(A_n)$ et par passage à la limite par σ-additivité :

$$P\Big(\bigcup_{k=m}^{\infty} A_k\Big) \leqslant \sum_{k=m}^{\infty} P(A_k)$$

ce qui montre que $P(B) = \lim\limits_{m\to\infty} P\Big(\bigcap\limits_{n=1}^{m} \bigcup\limits_{k\geq n} A_k\Big) = 0$.

2. a) Les règles de de Morgan donnent immédiatement $\overline{\Big(\bigcap\limits_{n\geq 1} \bigcup\limits_{k\geq n} A_k\Big)} = \bigcup\limits_{n\geq 1} \bigcap\limits_{k\geq n} \overline{A_n}$.

b) Par indépendance des événements (A_n) et donc de leurs complémentaires :

$$P\Big(\bigcap_{k=n}^{m} \overline{A_k}\Big) = \prod_{k=n}^{m} P(\overline{A_k}) = \prod_{k=n}^{m} (1-p_k)$$

c) Deux cas :

• si (p_n) ne tend pas vers 0, alors $\ln(1-p_n)$ ne tend pas vers 0 et la série $\sum \ln(1-p_n)$ diverge grossièrement.

• si $\lim\limits_{n\to+\infty} p_n = 0$, alors, $-\ln(1-p_n) \sim p_n$, et par hypothèse de la question, la série $\sum p_n$ diverge. On conclut par le théorème d'équivalence des séries à termes positifs ou nuls.

Dans les deux cas, la série $\sum \ln(1-p_n)$ diverge vers $-\infty$.

d) Comme $\ln\big(\prod_{k=n}^{m}(1-p_k)\big)=\sum_{k=n}^{m}\ln(1-p_k)$, par continuité de la fonction exponentielle, il vient : $\lim_{m\to+\infty}\prod_{k=n}^{m}(1-p_k)=\prod_{k=n}^{\infty}(1-p_k)=0$. Ainsi, $P\big(\bigcap_{k\geq n}\overline{A_k}\big)=0$.

Comme la suite $\big(\bigcap_{k\geq n}\overline{A_k}\big)_n$ est décroissante, pour tout m, on a :

$$\bigcup_{n=1}^{m}\bigcap_{k\geq n}\overline{A_k}=\bigcap_{k\geq m}\overline{A_k}$$

et donc, $P\big(\bigcup_{k=1}^{\infty}\bigcap_{k\geq n}\overline{A_k}\big)=0$, ce qui entraîne que $P\big(\bigcap_{k=1}^{\infty}\bigcup_{k\geq n}A_k\big)=1$.

3. a) De manière évidente $E(X_n)=P(Z_n=1)=\dfrac{1}{n^{\alpha}}$ qui tend vers 0 lorsque n tend vers $+\infty$

Pour les deux questions suivantes, on pose : $A_n=[X_n=1]$. Les variables aléatoires (X_n) étant indépendantes, les événements (A_n) le sont aussi.

b) La série $\sum P(A_n)$ diverge. Par la question 2, $P(B)=1$, ce qui signifie que les événements $[X_n=1]$ sont réalisés infiniment souvent avec la probabilité 1.

c) La série $\sum P(A_n)$ converge. Par la question 1, $P(B)=0$, ce qui signifie que les événements $[X_n=1]$ sont réalisés infiniment souvent avec une probabilité de 0, ou qu'ils ne sont réalisés qu'un nombre fini de fois avec la probabilité 1.

Exercice 3.16.

1. On considère la fonction f définie sur $\mathbb{R}$ par $t\mapsto f(t)=\frac{1}{2}\exp(-|t|)$.

Montrer que f est la densité de probabilité d'une variable aléatoire X dont on déterminera la fonction de répartition, l'espérance et l'écart-type.

2. Soient X_1 et X_2 deux variables aléatoires indépendantes, de même loi que X. Déterminer une densité g de $Y=X_1+X_2$. Quelles sont les valeurs de l'espérance et de l'écart-type de Y ?

3. a) Donner une densité de la variable aléatoire X_1-X_2.

b) On effectue indépendamment l'une de l'autre, deux mesures de la masse d'un même objet. On suppose que l'erreur commise à chaque fois suit la loi de X.

Soit a réel positif. Quelle est la probabilité que les deux mesures effectuées ne diffèrent pas de plus de a ?

c) Pour quelles valeurs de a l'inégalité de Bienaymé-Chebychev donne-t-elle une information sur cette probabilité ?

Solution :

1) La fonction f est continue sur $\mathbb{R}$, positive.

On a : $\displaystyle\int_{-\infty}^{+\infty} f(t)\,dt = \frac{2}{2}\int_0^{+\infty} \mathrm{e}^{-t}\,dt = 1$, donc f est une densité.

$\star$ Si F est la fonction de répartition de X, on a :

Si $x \leqslant 0$, $F(x) = \displaystyle\frac{1}{2}\int_{-\infty}^{x} \mathrm{e}^{t}\,dt = \frac{1}{2}\,\mathrm{e}^{x}$

Si $x \geqslant 0$, par symétrie, $F(x) = 1 - F(-x) = 1 - \dfrac{1}{2}\,\mathrm{e}^{-x}$.

$\star$ La convergence (absolue) de l'intégrale définissant l'espérance est banale et, par symétrie, l'espérance est nulle.

$\star$ Enfin la variance est égale au moment d'ordre 2 et :

$$V(X) = \int_{-\infty}^{+\infty} \frac{1}{2}\,t^2\mathrm{e}^{-|t|}\,dt = \int_0^{+\infty} t^2\mathrm{e}^{-t}\,dt = \Gamma(3) = 2$$

2) Une densité g de Y est donnée, par convolution, par :

$$g(x) = \int_{-\infty}^{+\infty} \frac{1}{4}\mathrm{e}^{-|t|}\mathrm{e}^{-|x-t|}\,dt$$

Pour $x < 0$, $4g(x) = \displaystyle\int_{-\infty}^{x} \mathrm{e}^{2t-x}\,dt + \int_x^0 \mathrm{e}^{x}\,dt + \int_0^{+\infty} \mathrm{e}^{x-2t}\,dt$

D'où : $g(x) = \dfrac{1}{4}(1-x)\mathrm{e}^{x}$

Tandis que pour $x > 0$: $4g(x) = \displaystyle\int_{-\infty}^{0} \mathrm{e}^{2t-x}\,dt + \int_0^{x} \mathrm{e}^{-x} + \int_x^{+\infty} \mathrm{e}^{x-2t}\,dt$

D'où : $g(x) = \dfrac{1}{4}(x+1)\mathrm{e}^{-x}$.

On peut donc écrire :

$$\forall\, x \in \mathbb{R}, g(x) = \frac{1}{4}(1+|x|)\mathrm{e}^{-|x|}$$

Par linéarité de l'espérance : $E(Y) = 0$ et, par indépendance :

$V(Y) = V(X_1) + V(X_2) = 4$.

3) a) La variable $-X_2$ a même loi que X_2, donc, toujours par indépendance, $X_1 - X_2$ a même loi que $Y = X_1 + X_2$.

b) Soit X_1 l'erreur aléatoire commise lors de la première mesure et X_2 celle commise lors de la seconde. On cherche la probabilité de l'événement $|X_1 - X_2| \leqslant a$, c'est-à-dire $-a \leqslant X_1 - X_2 \leqslant a$. On peut donc écrire :

$$p = P(-a \leqslant X_1 - X_2 \leqslant a) = \int_{-a}^{a} g(x)\,dx = \frac{1}{2}\int_0^{a} (x+1)\mathrm{e}^{-x}\,dx$$

Une intégration par parties donne la solution et :

$$p = \frac{1}{2}\big(2 - (a+2)\mathrm{e}^{-a}\big)$$

c) L'inégalité de Bienaymé-Chebychev donne une indication dès que la largeur de l'intervalle considéré est égale au moins à deux fois l'écart-type de Y, c'est-à-dire pour $a \geqslant 2$.

Exercice 3.17.

On note φ la densité continue sur $\mathbb{R}$ de la loi normale centrée réduite et Φ la fonction de répartition correspondante. On dira qu'une variable aléatoire réelle positive X suit la loi $LN(\mu,\theta), \mu \in \mathbb{R}, \theta \in \mathbb{R}_+^*$ si sa fonction de répartition F_X est donnée par :

$$F_X(x) = \begin{cases} \Phi\big(\frac{1}{\theta}((\ln x) - \mu)\big) & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$$

1. On suppose que X suit la loi $LN(\mu,\theta)$, et on pose $Y = \ln X$. Déterminer la fonction de répartition F_Y de Y.

2. Soit X_1 et X_2 deux variables aléatoires indépendantes, respectivement de lois $LN(\mu_1,\theta_1)$, $LN(\mu_2,\theta_2)$. Déterminer la loi du produit X_1X_2.

3. Soit Z est une variable aléatoire de loi $LN(\mu,\theta)$. Montrer que pour tout $c, \alpha \in \mathbb{R}_+^*$, la variable aléatoire cZ^α est de loi $LN(\mu',\theta')$, où on précisera les valeurs de μ' et θ'.

4. Soit $p \in \,]0,1[$ et $q = 1-p$. On définit $X_1,\dots,X_n$ des variables aléatoires indépendantes de même loi définie par :

$$P(X_i = \mathrm{e}) = p, P(X_i = 1) = q = 1-p.$$

a) Quelle est la loi de $S_n = \sum\limits_{i=1}^{n}(\ln X_i)$?

b) Montrer que lorsque n tend vers $+\infty$, la fonction de répartition de la variable aléatoire $T_n = \mathrm{e}^{-p\sqrt{n}}\big(\prod\limits_{i=1}^{n} X_i\big)^{\frac{1}{\sqrt{n}}}$ converge en tout point de $\mathbb{R}_+^*$ vers une limite que l'on déterminera.

Solution :

1. On a : $\forall\, x \in \mathbb{R}, P(Y \leqslant x) = P(\ln X \leqslant x) = P(X \leqslant \mathrm{e}^x) = \Phi\big(\frac{1}{\theta}(x-\mu)\big)$.

Donc Y suit la loi $\mathcal{N}(\mu,\theta)$.

2. Si on pose $Y_i = \ln X_i$, on a $X_1X_2 = \mathrm{e}^{Y_1+Y_2}$. Par stabilité et indépendance, Y_1+Y_2 suit la loi $\mathcal{N}(\mu_1+\mu_2, \sqrt{\theta_1^2+\theta_2^2})$. Par suite X_1X_2 suit la loi $LN(\mu_1+\mu_2, \sqrt{\theta_1^2+\theta_2^2})$.

3. On a $\ln(cZ^\alpha) = \ln c + \alpha \ln Z = \ln c + \alpha T$.
Soit $t \in \mathbb{R}$, on a :

$$P(\alpha T + \ln c \leqslant t) = P\big(T \leqslant \frac{1}{\alpha}(t - \ln c)\big) = \Phi_T\big(\frac{1}{\alpha}(t - \ln c)\big)$$
$$= \Phi\big(\frac{1}{\theta}\big(\frac{1}{\alpha}(t - \ln c) - \mu\big)\big) = \Phi\big(\frac{1}{\alpha\theta}(t - \ln c - \mu\alpha)\big).$$

Donc cZ^α suit la loi $LN\big(\mu\alpha + \ln c, \frac{1}{\alpha\theta}\big)$.

4. a) $\ln X_i$ suit la loi $\mathcal{B}(p)$. Par suite S_n suit la loi $\mathcal{B}(n, p)$.

b) On a :

$$\forall t \in \mathbb{R}_+^*, P(T_n \leqslant t) = P\big(\mathrm{e}^{-p\sqrt{n}}(\prod_{i=1}^{n} X_i)^{1/\sqrt{n}} \leqslant t\big)$$
$$= P(-p\sqrt{n} + \frac{1}{\sqrt{n}}\, S_n \leqslant \ln t) = P(\frac{S_n - np}{\sqrt{n}} \leqslant \ln t)$$

Or $\big(\dfrac{S_n - np}{\sqrt{npq}}\big)_n$ converge en loi vers une variable suivant la loi normale centrée réduite, donc :

$$\lim_{n\to\infty} P(T_n \leqslant t) = \lim_{n\to\infty} P(\frac{S_n - np}{\sqrt{npq}} \leqslant \frac{\ln t}{\sqrt{pq}}) = \Phi(\frac{\ln t}{\sqrt{pq}})$$

Exercice 3.18.

Dans cet exercice, toutes les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ telle que $P(X \geqslant n) > 0$ pour tout $n \in \mathbb{N}$.

On appelle *taux de panne* de X la suite réelle $(x_n)_{n\in\mathbb{N}}$ des probabilités conditionnelles définies par :

$$\forall n \in \mathbb{N}, x_n = P_{(X\geq n)}\big(X = n\big).$$

1. a) Vérifier que $P(Y = n) = \dfrac{1}{n(n+1)}$ définit bien une loi de probabilité sur $\mathbb{N}^*$.
(On pourra déterminer deux réels a et b tels que, pour tout $n \in \mathbb{N}^*$, on ait : $\dfrac{1}{n(n+1)} = \dfrac{a}{n} + \dfrac{b}{n+1}$.)

b) Déterminer le taux de panne de la variable aléatoire Y.

2. Dans le cas général, montrer que :

$$\forall n \in \mathbb{N}^*, P(X \geqslant n) = \prod_{k=0}^{n-1}(1 - x_k) = (1 - x_0)(1 - x_1)\ldots(1 - x_{n-1})$$

puis exprimer $p_n = P(X = n)$ en fonction des x_k.

3. Déterminer les lois de variables à valeurs dans $\mathbb{N}^*$ ayant un taux de panne constant pour $n > 0$.

4. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$, de taux de panne $(x_n)_{n\in\mathbb{N}}$, et $(U_k)_{k\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi uniforme sur $[0,1]$.
Soit Z définie par :

$$\forall\,\omega\in\Omega, Z(\omega)=\begin{cases} 0 & \text{si } \forall\,k\in\mathbb{N}, U_k(\omega)>x_k \\ \min\big\{k\in\mathbb{N}\ /\ U_k(\omega)\leqslant x_k\big\} & \text{sinon}\end{cases}$$

Montrer que Z est une variable aléatoire et déterminer sa loi.

Solution :

1. a) $\forall\,n\in\mathbb{N}^*, p_n=\dfrac{1}{n(n+1)}=\dfrac{1}{n}-\dfrac{1}{n+1}\geqslant 0$ et $\sum\limits_{k=1}^{n}p_k=1-\dfrac{1}{n+1}\underset{n\to+\infty}{\longrightarrow}1.$
On a donc bien affaire à une loi de probabilité sur $\mathbb{N}^*$.

b) Le taux de panne (y_n) est donné par $y_0=0$ et comme pour tout $n\geqslant 1$:

$$P(Y\geqslant n)=\lim_{p\to+\infty}\sum_{k=n}^{p}\Big(\frac{1}{k}-\frac{1}{k+1}\Big)=\lim_{p\to+\infty}\Big(\frac{1}{n}-\frac{1}{p+1}\Big)=\frac{1}{n}$$

il vient :

$$y_n=\frac{\frac{1}{n(n+1)}}{\frac{1}{n}}=\frac{1}{n+1}$$

2. On a : $1=P_{(X\geq n)}(X\geqslant n)=P_{(X\geq n)}(X\geqslant n+1)+P_{(X\geq n)}(X=n)$, ce qui donne :

$$1-x_n=\frac{P[(X\geqslant n+1)\cap(X\geqslant n)]}{P(X\geqslant n)}=\frac{P(X\geqslant n+1)}{P(X\geqslant n)}$$

d'où :

$$\prod_{k=0}^{n-1}(1-x_k)=\frac{P(X\geqslant n)}{P(X\geqslant 0)}=P(X\geqslant n)$$

puis :

$$p_n=P(X\geqslant n)-P(X\geqslant n+1)=x_n\prod_{k=0}^{n-1}(1-x_k)$$

3. Condition nécessaire : si X est à valeurs dans $\mathbb{N}^*$, de taux de panne $(x_n)_{n\in\mathbb{N}^*}$ constant égal à x, alors $p_0=x_0=0$ et par la question précédente, $p_n=x(1-x)^{n-1}$ donc $X\hookrightarrow\mathcal{G}(x)$.

Condition suffisante : Si $X\hookrightarrow\mathcal{G}(p)$, on a $p_0=0$ et $P(X\geqslant 0)=1$, d'où $x_0=0$ puis, pour tout $n\in\mathbb{N}^*$:

$$P(X=n)=pq^{n-1}, P(X\geqslant n)=\sum_{k=n}^{+\infty}pq^{k-1}=\frac{pq^{n-1}}{1-q}=q^{n-1}>0$$

d'où :

$$x_n=\frac{P(X=n)}{P(X\geqslant n)}=p$$

4. Soit $A = \big\{\omega \in \Omega \ / \ \forall\, k \in \mathbb{N}, U_k(\omega) > x_k\big\} = \bigcap\limits_{k=0}^{+\infty} \big(U_k > x_k\big)$.

Par indépendance des U_k et $x_k \in [0,1]$, on a :

$P\big(\bigcap\limits_{k=0}^{n} (U_k > x_k)\big) = \prod\limits_{k=0}^{n} P(U_k > x_k) = \prod\limits_{k=0}^{n} (1 - x_k) = P(X \geqslant n+1) \underset{n\to+\infty}{\longrightarrow} 0$

d'où $P(A) = 0$, par limite monotone, et donc Z est définie presque partout.

On a $Z(\Omega) \subset \mathbb{N}$ et, par indépendance des U_k de loi uniforme sur $[0,1]$:

$$\begin{aligned} P(Z = n) &= P\big(\big(\bigcap_{k=0}^{n-1} (U_k > x_k)\big) \cap (U_n \leqslant x_n)\big) = P(U_n \leqslant x_n) \prod_{k=0}^{n-1} P(U_k > x_k) \\ &= x_n \prod_{k=0}^{n-1} (1 - x_k) = p_n \end{aligned}$$

Conclusion : Z suit la même loi que X.

Exercice 3.19.

1. Montrer que pour tout $x \in \mathbb{R}^+$, on a $1 - x \leqslant \mathrm{e}^{-x}$.

2. On dispose d'une urne vide au départ. Le premier jour, une personne met une boule numérotée 1 dans l'urne, la tire, note son numéro et la remet dans l'urne (!). Ensuite, à chaque nouvelle journée, elle ajoute une boule qui porte le numéro du jour considéré, elle tire alors une boule au hasard, note le numéro de cette boule et la remet dans l'urne. Le processus se poursuit indéfiniment ...

a) Soient $\ell \in \mathbb{N}^*$ et $E_1, E_2, \dots, E_\ell$ une famille de ℓ événements indépendants. Montrer que l'on a

$$P\big(\bigcap_{1\leq i\leq \ell} \overline{E_i}\big) \leqslant \mathrm{e}^{-\sum\limits_{i=1}^{\ell} P(E_i)}.$$

où $\overline{E}$ est l'événement contraire de l'événement E.

b) On note A_k l'événement « la boule numérotée 10 sort lors du $k^{\text{ème}}$ tirage. Que vaut la probabilité de A_k ?

c) Quelle est la probabilité que la boule 10 sorte au moins une fois à partir du $n^{\text{ème}}$ tirage, où n est un entier positif fixé ?

d) Quelle est la probabilité que la boule numérotée 10 sorte une infinité de fois ?

e) Calculer la probabilité que le 10 sorte une infinité de fois de suite.

3. On suppose cette fois que la personne remplit l'urne de sorte qu'il y ait dans l'urne n^2 boules, numérotées de 1 à n^2, le $n^{\text{ème}}$ jour (elle met donc une boule numérotée 1 le premier jour, trois boules numérotées $2, 3, 4$ le deuxième

jour, cinq boules le troisième, ...). Comme à la question précédente, elle tire alors une boule, note son numéro et la remet immédiatement dans l'urne.

a) Soient $\ell \in \mathbb{N}^*$ et $E_1, E_2, \dots, E_\ell$ une famille de ℓ événements . Montrer que l'on a : $P(\bigcup_{1\leqslant i\leqslant \ell} E_i) \leqslant \sum_{i=1}^{\ell} P(E_i)$.

b) Quelle est la probabilité que le nombre 10 sorte une infinité de fois ?

Solution :

1. Banal, par inégalité de convexité ou étude de la fonction associée.

2. a) Comme les événements considérés sont indépendants, en utilisant la question 1, il vient :

$$P(\bigcap_{1\leq i\leq \ell} \overline{E_i}) = \prod_1^{\ell} P(\overline{E_i}) = \prod_1^{\ell}(1-P(E_i)) \leqslant \prod_1^{\ell} \mathrm{e}^{-P(E_i)} = \mathrm{e}^{-\sum\limits_{i=1}^{\ell} P(E_i)}.$$

b) Il est clair que :

$$P(A_k) = 0 \text{ si } 1 \leqslant k < 10 \text{ et } P(A_k) = \frac{1}{k} \text{ si } k \geqslant 10.$$

c) Ceci correspond à la probabilité de l'événement $\bigcup_{k=n}^{+\infty} A_k$. Passons à l'événement contraire, on obtient :

$$P(\bigcap_{k\geq n} \overline{A_k}) = \lim_{\ell\to+\infty} P(\bigcap_{n\leq k\leq l} \overline{A_k}).$$

Or d'après la question 2 a), on a :

$$P(\bigcap_{n\leq k\leq \ell} \overline{A_k}) \leqslant \exp\big(-\sum_{k=n}^{l} P(A_k)\big) = \exp\big(-\sum_{k=\sup(n,10)}^{\ell} 1/k\big) \underset{\ell\to+\infty}{\longrightarrow} 0$$

D'où $P(\bigcap_{k\geq n} \overline{A_k}) = 0$ et par suite, $P(\bigcup_{k=n}^{+\infty} A_k) = 1$.

d) L'événement « le 10 sort une infinité de fois » s'écrit au moyen de l'intersection décroissante des événements précédents. C'est donc :

$$P(\bigcap_{n\geq 1} (\bigcup_{k=n}^{+\infty} A_k)) = \lim_{n\to+\infty} P(\bigcup_{k=n}^{+\infty} A_k) = 1$$

Le 10 sort donc presque sûrement une infinité de fois.

e) Ici l'événement considéré correspond à la réunion croissante : $\bigcup_{n\geq 1}\Big(\bigcap_{k=n}^{+\infty} A_k\Big)$ et on a :

$$P(\bigcup_{n\geq 1} (\bigcap_{k=n}^{+\infty} A_k)) = \lim_{n\to+\infty} P(\bigcap_{k=n}^{+\infty} A_k) \leqslant \lim_{n\longrightarrow+\infty} P(A_n) = 0.$$

La probabilité que le 10 sorte une infinité de fois de suite est donc nulle.

3. a) La formule de Poincaré $P(A\bigcup B) = P(A)+P(B) - P(A\bigcap B)$ implique $P(A\bigcup B) \leqslant P(A) + P(B)$. On obtient ensuite l'inégalité souhaitée par récurrence.

b) Dans ce cas, on a : $P(A_k) = 0$ si $1 \leqslant k < 10$ et $P(A_k) = \dfrac{1}{k^2}$ si $k \geqslant 10$.

Avec la majoration donnée en 3. a), on obtient en passant à la limite :

$$P\big(\bigcup_{k=n}^{+\infty} A_k\big) \leqslant \sum_{k=n}^{+\infty} \frac{1}{k^2},$$

d'où :

$$P\big(\bigcap_{n\geq 1} (\bigcup_{k=n}^{+\infty} A_k)\big) = \lim_{n\to+\infty} P\big(\bigcup_{k=n}^{+\infty} A_k\big) \leqslant \lim_{n\to+\infty} \sum_{k=n}^{+\infty} \frac{1}{k^2} = 0$$

puisque le reste d'une série convergente tend vers 0.
Cette fois la probabilité que le 10 sorte une infinité de fois est nulle.

Exercice 3.20.

Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé. On note $\mathcal{E}$ l'ensemble des variables aléatoires X à densité définies sur cet espace, à valeurs dans $[-1, 1]$ et centrées (*i.e.* telles que l'espérance $E(X)$ existe et vaut 0).

1. Soit $X \in \mathcal{E}$.

a) Montrer que, pour tout $\lambda \in \mathbb{R}$, la variable aléatoire $\mathrm{e}^{\lambda X}$ admet une espérance.

b) Montrer que, pour tout $a \in \mathbb{R}$ et tout $\lambda \geqslant 0$, on a :

$$P(X \geqslant a) \leqslant \mathrm{e}^{-\lambda a} E\big(\mathrm{e}^{\lambda X}\big)$$

c) L'ensemble $\mathcal{E}$ est-il un espace vectoriel ?

2. Soit $\lambda \geqslant 0$ et soit $X \in \mathcal{E}$.

a) Pour tout $n \in \mathbb{N}$, montrer que : $(2n)! \geqslant 2^n n!$.

b) En déduire que : $\frac{1}{2}\big(\mathrm{e}^{\lambda} + \mathrm{e}^{-\lambda}\big) \leqslant \mathrm{e}^{\frac{\lambda^2}{2}}$.

c) Montrer que, pour tout $x \in [-1, 1]$, on a : $\mathrm{e}^{\lambda x} \leqslant \frac{1}{2}\big(\mathrm{e}^{\lambda} + \mathrm{e}^{-\lambda}\big) + \frac{x}{2}\big(\mathrm{e}^{\lambda} - \mathrm{e}^{-\lambda}\big)$.

d) Déduire des questions précédentes que :

$$E\big(\mathrm{e}^{\lambda X}\big) \leqslant \mathrm{e}^{\frac{\lambda^2}{2}} \text{ et } E\big(\mathrm{e}^{-\lambda X}\big) \leqslant \mathrm{e}^{\frac{\lambda^2}{2}}.$$

3. Montrer que, pour tout $a \geqslant 0$ et tout $X \in \mathcal{E}$, on a : $P(|X| \geqslant a) \leqslant 2\,\mathrm{e}^{-\frac{a^2}{2}}$.

4. Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes qui appartiennent à $\mathcal{E}$. Montrer que :

$$\forall a \geqslant 0, \forall n \in \mathbb{N}^*, P\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n} X_i\Big| \geqslant a\Big) \leqslant 2\,\mathrm{e}^{-\frac{a^2}{2}}.$$

Solution :

1. a) Soit f une densité de X nulle en dehors de $[-1,1]$. Par le théorème de transfert il suffit que l'intégrale $\displaystyle\int_{\mathbb{R}} \mathrm{e}^{\lambda x} f(x)\,dx$ converge, ce qui est banalement le cas car la fonction à intégrer est nulle en dehors de $[-1,1]$.

b) L'inégalité est évidente pour $\lambda = 0$ car alors $\mathrm{e}^{-\lambda a}E\big(\mathrm{e}^{\lambda X}\big) = 1$.
Pour $\lambda > 0$, d'après l'inégalité de Markov appliquée à la variable positive $\mathrm{e}^{\lambda X}$, on a :

$$P(X \geqslant a) = P\big(\mathrm{e}^{\lambda X} \geqslant \mathrm{e}^{\lambda a}\big) \leqslant \frac{E(\mathrm{e}^{\lambda X})}{\mathrm{e}^{\lambda a}}.$$

c) Non : si X suit une loi uniforme sur $[-1,1]$, alors $X \in \mathcal{E}$ mais $2X \notin \mathcal{E}$, car cette variable aléatoire prend ses valeurs entre -2 et 2 ....

2. a) Pour $n \geqslant 1$, $\dfrac{(2n)!}{n!} = (n+1)(n+2)\dots(2n)$ et les n facteurs sont minorés par 2, donc $(2n)! \geqslant 2^n n!$ et le résultat est banal au rang 0.

b) On a par simplifications des sommations de séries :

$$\frac{1}{2}\big(\mathrm{e}^{\lambda} + \mathrm{e}^{-\lambda}\big) = \sum_{n=0}^{+\infty} \frac{\lambda^{2n}}{(2n)!} \leqslant \sum_{n=0}^{+\infty} \frac{\lambda^{2n}}{2^n n!} = \mathrm{e}^{\frac{\lambda^2}{2}}$$

c) Cette relation s'écrit : $\mathrm{e}^{\lambda x} \leqslant \dfrac{x+1}{2}\mathrm{e}^{\lambda} + \dfrac{1-x}{2}\mathrm{e}^{-\lambda}$, soit encore, en posant $t = \dfrac{x+1}{2} \in [0,1]$:

$$\mathrm{e}^{\lambda[t.1+(1-t).(-1)]} \leqslant t\mathrm{e}^{\lambda} + (1-t)\mathrm{e}^{-\lambda},$$

inégalité qui est vraie d'après la convexité de la fonction $u \mapsto \mathrm{e}^{\lambda u}$ sur $[-1,1]$.
[On peut aussi étudier la fonction $\varphi : x \mapsto \mathrm{e}^{\lambda x} - \frac{1}{2}\big(\mathrm{e}^{\lambda} + \mathrm{e}^{-\lambda}\big) - \frac{x}{2}\big(\mathrm{e}^{\lambda} - \mathrm{e}^{-\lambda}\big)$ *associée, dont la dérivée seconde est évidemment strictement positive sur* $[-1,1]$. *Comme* $\varphi(-1) = \varphi(1) = 0$, *il n'est même pas nécessaire d'étudier le signe de* φ' *...]*

d) D'après 2. b) et 2. c), on a :

$$E\big(\mathrm{e}^{\lambda X}\big) = \int_{-1}^{1} \mathrm{e}^{\lambda x} f(x)\,dx \leqslant \int_{-1}^{1} \big(\mathrm{e}^{\frac{\lambda^2}{2}} + \tfrac{x}{2}(\mathrm{e}^{\lambda} - \mathrm{e}^{-\lambda})\big) f(x)\,dx = \mathrm{e}^{\frac{\lambda^2}{2}} + 0$$

Comme $-X \in \mathcal{E}$ on peut remplacer X par $-X$ ce qui donne : $E\big(\mathrm{e}^{-\lambda X}\big) \leqslant \mathrm{e}^{\frac{\lambda^2}{2}}$.

3. On peut écrire, pour tout $\lambda \geqslant 0$:

$$\begin{aligned} P(|X| \geqslant a) &= P(X \geqslant a) + P(X \leqslant -a) = P(X \geqslant a) + P(-X \geqslant a) \\ &\leqslant \mathrm{e}^{-\lambda a}E(\mathrm{e}^{\lambda X}) + \mathrm{e}^{-\lambda a}E(\mathrm{e}^{-\lambda X}) \end{aligned}$$

$\leqslant 2\mathrm{e}^{-\lambda a}\mathrm{e}^{\frac{\lambda^2}{2}}$

et en prenant $\lambda = a$: $P(|X| \geqslant a) \leqslant \mathrm{e}^{-\frac{a^2}{2}}$

4. De la même façon, on remarque que $\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \in \mathcal{E}$. Donc, pour tout $\lambda \geqslant 0$, comme à la question précédente, on a :

$$\begin{aligned}
P\big(\frac{1}{\sqrt{n}}|\sum_{i=1}^{n} X_i| \geqslant a\big) &= P\big(\frac{1}{n}|\sum_{i=1}^{n} X_i| \geqslant \frac{a}{\sqrt{n}}\big)\\
&\leqslant \mathrm{e}^{-\lambda\frac{a}{\sqrt{n}}}[E(\exp(\frac{\lambda}{n}\sum_{i=1}^{n} X_i)) + E(\exp(-\frac{\lambda}{n}\sum_{i=1}^{n} X_i))]\\
&\leqslant \mathrm{e}^{-\lambda\frac{a}{\sqrt{n}}}\big[\prod_{i=1}^{n} E(\mathrm{e}^{\frac{\lambda}{n}X_i}) + \prod_{i=1}^{n} E(\mathrm{e}^{-\frac{\lambda}{n}X_i})\big] \text{ (par indépendance)}\\
&\leqslant 2\mathrm{e}^{-\lambda\frac{a}{\sqrt{n}}}\prod_{i=1}^{n}\mathrm{e}^{\frac{\lambda^2}{2n^2}} = 2\mathrm{e}^{-\lambda\frac{a}{\sqrt{n}}}\mathrm{e}^{\frac{\lambda^2}{2n}}
\end{aligned}$$

Pour conclure il suffit donc de trouver $\lambda \geqslant 0$ tel que $-\lambda\dfrac{a}{\sqrt{n}} + \dfrac{\lambda^2}{2n} \leqslant -\dfrac{a^2}{2}$, soit $\big(a - \dfrac{\lambda}{\sqrt{n}}\big)^2 \leqslant 0$. On doit prendre : $\lambda = a\sqrt{n}$.