3

PROBABILITÉS

Exercice 3.01.

Soit n un entier tel que $n \geqslant 2$. On dispose de n urnes. On y place aléatoirement des jetons, un par un, indépendamment les uns des autres.

Pour $j \geqslant 1$, on note X_j le nombre aléatoire d'urnes non vides après avoir placé les j premiers jetons.

1. a) Quelles valeurs peut prendre la variable aléatoire X_j ? Déterminer les lois de X_1 et X_2.

b) Pour $i \in [\![1, n]\!]$ et $j \geqslant 1$, déterminer les probabilités conditionnelles

$$P_{(X_j=i)}(X_{j+1} = k)$$

c) En déduire la loi de X_{j+1} en fonction de la loi de X_j.

2. On considère la suite $(Q_j)_{j\geq 1}$ de fonctions polynômes définie par :

$$\forall\, x \in \mathbb{R}, \forall\, j \in \mathbb{N}^*, Q_1(x) = x^{n-1} \text{ et } Q_{j+1}(x) = Q_j(x) + \frac{1-x}{n} Q_j'(x)$$

a) Montrer que pour tout $j \geqslant 1$:

$$Q_j(x) = \sum_{i=0}^{n-1} (x-1)^i (1 - \frac{i}{n})^{j-1} \binom{n-1}{i}$$

b) Pour tout entier $j \geqslant 1$, soit la fonction G_j définie par :

$$G_j : \mathbb{R} \to \mathbb{R}, x \mapsto \sum_{k=1}^{n} P(X_j = k) x^{n-k}$$

Pour un réel x fixé, vérifier que les suites $(G_j(x))_{j\geq 1}$ et $(Q_j(x))_{j\geq 1}$ sont égales et en déduire l'expression des $P(X_j = n)$.

c) En déduire que pour tout $j \in [\![1, n-1]\!]$: $\displaystyle\sum_{i=0}^{n-1} (-1)^i (1 - \frac{i}{n})^{j-1} \binom{n-1}{j} = 0$.

Solution :

1. a) Le nombre d'urnes non vides X_j vérifie : $1 \leqslant X_j \leqslant \min(j,n)$.

On a : $P(X_1 = 1) = 1$; $P(X_2 = 1) = \dfrac{1}{n}$ (on a une chance sur n pour que le deuxième jeton arrive dans la même urne que le premier), donc par passage au complémentaire $P(X_2 = 2) = \dfrac{n-1}{n}$.

b) Soit $j \geqslant 1$ et $1 \leqslant k \leqslant j+1$:
$P_{(X_j=k)}(X_{j+1} = k) = \dfrac{k}{n}$, $P_{(X_j=k-1)}(X_{j+1} = k) = \dfrac{n-k+1}{n}$,
sinon $P_{(X_j=i)}(X_{j+1} = k) = 0$.
Car en plaçant un jeton, le nombre d'urnes non vides ne peut que rester inchangé ou bien augmenter d'une unité (le résultat vaut même pour $k = 1$ et bien sûr pour $k = n$).

c) L'ensemble des $(X_j = i)$, pour $0 \leqslant i \leqslant \min(j,n)$ est un système complet d'événements, donc :

$$\begin{aligned} P(X_{j+1} = k) &= \sum_{i=0}^{\min(j,n)} P_{(X_j=i)}(X_{j+1} = k)P(X_j = i) \\ &= \frac{k}{n}P(X_j = k) + \frac{n-k+1}{n}P(X_j = k-1). \end{aligned}$$

2. a) pour $j = 1$, on a : $\displaystyle\sum_{i=0}^{n-1} (x-1)^i(1-\tfrac{i}{n})^{1-1}\binom{n-1}{i} = (1+(x-1))^{n-1} = x^{n-1}$.

Soit $j \geqslant 1$, on suppose que $Q_j(x) = \displaystyle\sum_{i=0}^{n-1} (x-1)^i(1-\tfrac{i}{n})^{j-1}\binom{n-1}{i}$.

Alors $Q_j'(x) = \displaystyle\sum_{i=1}^{n-1} i(x-1)^{i-1}(1-\frac{i}{n})^{j-1}\binom{n-1}{i}$ et

$$\begin{aligned} Q_{j+1}(x) &= Q_j(x) + \frac{1-x}{n}Q_j'(x) \\ &= 1 + \sum_{i=1}^{n-1} (x-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i} + i(x-1)^{i-1}(1-\tfrac{i}{n})^{j-1}\binom{n-1}{i} \\ &= 1 + \sum_{i=1}^{n-1} (x-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i}\left(1-\frac{i}{n}\right) \\ &= 1 + \sum_{i=1}^{n-1} (x-1)^i(1-\frac{i}{n})^{j}\binom{n-1}{i} = \sum_{i=0}^{n-1} (x-1)^i(1-\frac{i}{n})^{j}\binom{n-1}{i}. \end{aligned}$$

On conclut par le principe de récurrence.

b) $G_1(x) = \displaystyle\sum_{k=1}^{n} P(X_1 = k)x^{n-k} = P(X_1 = 1)x^{n-1} = x^{n-1} = Q_1(x)$.

$$\begin{aligned} G_j(x) + \frac{1-x}{n}G_j'(x) &= \sum_{k=1}^{n} P(X_j = k)[1-\frac{n-k}{n}]x^{n-k} \\ &\qquad + \sum_{k=1}^{n-1} \frac{n-k}{n}P(X_j = k)x^{n-(k+1)} \\ &= \frac{P(X_j = 1)}{n}x^{n-1} \end{aligned}$$

$$+\sum_{k=2}^{n}[P(X_j = k)\frac{k}{n} + P(X_j = k-1)\frac{n-k+1}{n}]x^{n-k}$$

Donc $G_j(x) + \frac{1-x}{n}G'_j(x) = \sum_{k=1}^{n} P(X_{j+1} = k)x^{n-k}$.

Les suites $(G_j(x))_{j\geq 1}$ et $(Q_j(x))_{j\geq 1}$ vérifient la même relation de récurrence d'ordre 1 et ont les mêmes premiers termes, donc par …récurrence, sont égales :

$G_j(x) = Q_j(x) = \sum_{i=0}^{n-1}(x-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i}$.

c) En particulier, pour $x = 0$:

$G_j(0) = P(X_j = n)$ et $Q_j(0) = \sum_{i=0}^{n-1}(-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i}$, d'où :

$$P(X_j = n) = \sum_{i=0}^{n-1}(-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i}$$

Or, pour $1 \leqslant j < n$, l'événement $(X_j = n)$ est impossible et donc $P(X_j = n) = 0$, d'où :

$$1 \leqslant j < n \implies \sum_{i=0}^{n-1}(-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i} = 0$$

Exercice 3.02.

On considère l'application F définie sur $\mathbb{R}$ par : $F(t) = \dfrac{e^t}{1+e^t}$.

1. Montrer que F est la fonction de répartition d'une variable aléatoire à densité dont on déterminera une densité notée f.

2. Soit X une variable aléatoire de densité f. Montrer que X admet des moments d'ordre n pour tout entier naturel n.

On pose : $I = \displaystyle\int_0^{+\infty} \frac{t}{1+e^t}\,dt$.

Calculer l'espérance $E(X)$ de X ainsi que sa variance $V(X)$ en fonction de I.

3. On considère une suite de variables aléatoires réelles $(X_n)_{n\geq 1}$ définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, mutuellement indépendantes et de densité f.

Soit $(\overline{X}_n)_{n\geq 1}$ la suite de variables aléatoires définie par :

$$\forall\, n \geqslant 1, \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

a) Montrer que la suite $(\overline{X}_n)_{n\geq 1}$ converge en probabilité vers 0, puis déterminer une suite de réels $(a_n)_{n\geq 1}$ telle que la suite $(a_n\overline{X}_n)_{n\geq 1}$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

b) On pose $S_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$.

Construire à partir de S_n^2 un estimateur sans biais de I. Montrer que cet estimateur est convergent.

4. Proposer en `Scilab` une simulation de la loi associée à f.

Solution :

1. La fonction F est bien définie sur $\mathbb{R}$ et est une fonction de classe C^1 sur $\mathbb{R}$. La limite de F en $-\infty$ est 0 et celle en $+\infty$ est 1. De plus F est une fonction croissante sur $\mathbb{R}$ puisque sa dérivée est positive sur $\mathbb{R}$.

$$\forall x \in \mathbb{R}, f(x) = F'(x) = \frac{\mathrm{e}^t}{(1+\mathrm{e}^t)^2}$$

F est donc la fonction de répartition d'une variable aléatoire à densité dont une densité est la fonction f.

2. Soit X une variable aléatoire de densité f. Pour tout entier naturel n, on étudie la nature de l'intégrale : $\displaystyle\int_{-\infty}^{+\infty} t^n \times \frac{\mathrm{e}^t}{(1+\mathrm{e}^t)^2} dt$

$\to$ En $-\infty$, la fonction $t^n \times \dfrac{\mathrm{e}^t}{(1+\mathrm{e}^t)^2}$ est équivalente à $t^n \mathrm{e}^t$ qui est négligeable devant $1/t^2$, on conclut à la convergence en $-\infty$.

$\to$ En $+\infty$, la fonction $t^n \times \dfrac{\mathrm{e}^t}{(1+\mathrm{e}^t)^2}$ est équivalente à $t^n \mathrm{e}^{-t}$ qui est négligeable devant $1/t^2$, on en conclut à la convergence en $+\infty$.
Par conséquent X admet un moment d'ordre n pour tout entier n.

En remarquant que f est une fonction paire, on en déduit que $E(X) = 0$.

D'autre part : $V(X) = E(X^2) = 2\displaystyle\int_0^{+\infty} t^2 \frac{\mathrm{e}^t}{(1+\mathrm{e}^t)^2} dt = 2\int_0^{+\infty} t^2 (F-1)'(t) dt.$

On réalise l'intégration par parties ainsi préparée, d'où :

$$V(X) = 2\int_0^{+\infty} 2t(1 - F(t))dt = 4I$$

3. a) Puisque $E(\overline{X_n}) = 0$ et $V(\overline{X_n}) = \dfrac{4I}{n}$, en utilisant l'inégalité de Bienaymé-Tchebicheff, on en déduit que $(\overline{X_n})_n$ converge en probabilité vers 0.

En utilisant le théorème limite central, on a $\big(\sqrt{n}\dfrac{\overline{X_n}}{2\sqrt{I}}\big)_n$ qui converge en loi vers la loi normale centrée réduite. Donc la suite cherchée est $a_n = \dfrac{\sqrt{n}}{2\sqrt{I}}$.

b) On a $E(S_n^2) = 4I$, donc un estimateur sans biais de I est $\dfrac{S_n^2}{4}$. Comme les variables X_i sont indépendantes, les variables X_i^2 le sont aussi donc : $V(\dfrac{S_n^2}{4}) = \dfrac{V(X^2)}{16n}$.

Cette expression tend vers 0 quand n tend vers $+\infty$, donc $\dfrac{S_n^2}{4}$ est un estimateur convergent de I.

4. Soit U une variable suivant la loi uniforme sur $]0,1[$.

Puisque F réalise une bijection de $\mathbb{R}$ dans $]0,1[$ alors $F^{-1}(U)$ suit la loi de X. Or :

$$\forall x \in]0,1[, F^{-1}(x) = \ln\left(\frac{x}{1-x}\right)$$

On propose donc :

```
x=rand()
y=log(x/(1-x))
```

Exercice 3.03.

Soit un plan muni d'un repère orthonormé. On considère une particule se déplaçant au hasard dans ce plan de la manière suivante : à l'instant 0, la particule se trouve à l'origine O ; ensuite, à chaque instant (représenté par un entier naturel), la particule se déplace d'un pas de longueur 1 de façon équiprobable dans une direction parmi Nord, Sud, Est et Ouest.

Toutes les variables aléatoires modélisant cette situation seront définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

On admet la formule de Stirling :

$$n! \underset{(n\to\infty)}{\sim} n^n \mathrm{e}^{-n}\sqrt{2\pi n}$$

et le résultat suivant : si une variable aléatoire T discrète à valeurs dans $\mathbb{N}^*$ admet une espérance, alors $E(T) = \sum\limits_{k=1}^{+\infty} P(T \geqslant k)$.

On note $Z_n = (X_n, Y_n)$ le vecteur aléatoire représentant la position de la particule (abscisse, ordonnée) à l'instant n.

1. Écrire un algorithme en `Scilab` permettant de simuler une réalisation de Z_1.

2. Pour tout entier $p \in \mathbb{N}$, déterminer la valeur de $P(Z_{2p+1} = (0,0))$.

3. a) Montrer que pour tout entier p, $\sum\limits_{k=0}^{p} \binom{p}{k}\binom{p}{p-k} = \binom{2p}{p}$.

b) Montrer que pour tout entier $p \in \mathbb{N}^*$,

$$P(Z_{2p} = (0,0)) = \frac{1}{4^{2p}}\binom{2p}{p}\sum_{k=0}^{p}\binom{p}{k}\binom{p}{p-k} = \frac{1}{4^{2p}}\binom{2p}{p}^2$$

c) Montrer que la série de terme général $P(Z_{2p} = (0,0))$ diverge.

4. Pour $p \in \mathbb{N}$, on note N_p la variable aléatoire valant 1 si $Z_p = (0,0)$ et 0 sinon.

a) Pour $p \in \mathbb{N}^*$, quelle est la loi de N_p ?

b) Déterminer $\lim\limits_{p\to+\infty} \sum\limits_{i=1}^{p} E(N_i)$.

c) Pour $k \in \mathbb{N}^*$, soit R_k l'événement « au cours du temps la particule est repassée au moins k fois par l'origine » Comparer pour $k \in \mathbb{N}^*$ et $p \in \mathbb{N}$, les événements R_k et $\left(\sum\limits_{i=0}^{p} N_i \geqslant k\right)$.

d) Montrer par l'absurde que la série de terme général $P(R_k)$ est divergente.

e) On admet que $\forall\, k \in \mathbb{N}^*, P(R_k) = P(R_1)^k$. Déterminer alors la valeur de $P(R_k)$ pour tout entier $k \geqslant 1$. Conclure.

Solution :

1.
```
u=zeros(1,2)
if rand()<1/2 //choix entre N-S ou E-O
   then if rand()<1/2 //choix du déplacement
    then u(2)=-1
else u(2)=1
 end
  else if rand()<1/2
  then u(1)=-1
 else u(1)=1
   end
  end
  disp(u)
```

2. Pour revenir en $(0,0)$, il faut autant de déplacements vers le Nord que vers le Sud et autant de déplacements vers l'Est et l'Ouest. Ceci n'est possible que si le nombre de déplacements est pair. Ainsi, pour tout entier $p \in \mathbb{N}$, $P(Z_{2p+1} = (0,0)) = 0$.

3. a) Pour choisir p éléments parmi $2p$, on peut choisir pour tout $k \in [\![0,p]\!]$, k éléments parmi un sous-ensemble à p éléments, puis choisir les $p-k$ autres éléments parmi les p éléments restants. En additionnant tous les cas, il vient : $\sum\limits_{k=0}^{p} \binom{p}{k}\binom{p}{p-k} = \binom{2p}{p}$.

b) La question précédente et la formule admise donnent : $P(Z_{2p} = (0,0)) = \dfrac{1}{4^{2p}}\dbinom{2p}{p}^2$. Or d'après la formule de Stirling, $\dbinom{2p}{p} = \dfrac{(2p)!}{(p!)^2} \sim \dfrac{\sqrt{4\pi p}(2p)^{2p}\mathrm{e}^{-2p}}{2\pi p p^{2p}\mathrm{e}^{-2p}}$, donc :

$$\frac{1}{4^{2p}}\binom{2p}{p}^2 \sim \frac{1}{\pi p}$$

La série de terme général $\dfrac{1}{\pi p}$ diverge, donc par critère d'équivalence pour les séries à termes positifs, la série de terme général $P(Z_{2p} = (0,0))$ diverge.

4. a) Pour $p \in \mathbb{N}$, $N_p \hookrightarrow \mathcal{B}(P(Z_p = (0,0))$.

b) On a donc : $\forall\, i \in [\![0,p]\!], E(N_i) = P(Z_i = (0,0))$, et $\sum\limits_{i=0}^{p} E(N_i) = \sum\limits_{i=0}^{p} P(Z_i = (0,0))$.

Or la série de terme général $P(Z_i = (0,0))$ diverge et est à termes positifs, donc $\lim\limits_{p\to\infty} \sum\limits_{i=0}^{p} E(N_i) = +\infty$.

c) Si $[(\sum_{i=0}^{p} N_i) \geqslant k]$ est réalisé, alors en p déplacements, il y a eu au moins k retours en $(0,0)$, donc sur l'ensemble des déplacements, il y aura au moins k retours en $(0,0)$, donc $[N \geqslant k]$ est réalisé. Ainsi, $[(\sum_{i=0}^{p} N_i) \geqslant k] \subset [N \geqslant k]$.

d) Par croissance de l'application probabilité, pour p fixé,

$$P([(\sum_{i=0}^{p} N_i) \geqslant k] \leqslant P([N \geqslant k])$$

Ainsi, si la série de terme général $P([N \geqslant k])$ était convergente, alors par le critère de comparaison pour les séries à termes positifs , pour p fixé, la série de terme général $P([(\sum_{i=0}^{p} N_i) \geqslant k]$ convergerait et :

$$\forall p \in \mathbb{N}, \sum_{k=1}^{+\infty} P([N \geqslant k]) \geqslant \sum_{k=1}^{+\infty} P([\sum_{i=0}^{p} N_i \geqslant k]$$

donc d'après le résultat admis, $\forall p \in \mathbb{N}, \sum_{k=1}^{+\infty} P([N \geqslant k]) \geqslant E(N_0 + \cdots + N_p)$.

Ceci est absurde car $\lim_{p\to+\infty} \sum_{i=0}^{p} E(N_i) = +\infty$. La série de terme général $P([N \geqslant k])$ est donc divergente.

e) D'après le résultat admis, la série de terme général $P(N \geqslant k)$ est une série géométrique divergente de raison $P(N \geqslant 1)$, donc comme $P(N \geqslant 1) \in [0,1]$, on en déduit que $P(N \geqslant 1) = 1$ et aussi $\forall\, k \in P(N \geqslant k) = 1$.

On est quasiment certain que la particule repassera par le point $(0,0)$ une infinité de fois.

Exercice 3.04.

1. Soit F la fonction définie sur $\mathbb{R}$ par : $\forall\, x \in \mathbb{R}, F(x) = \exp(-\mathrm{e}^{-x})$.

a) Justifier que F est une fonction de répartition.

b) Soit X une variable aléatoire de fonction de répartition F. Déterminer une densité f de X.

On suppose désormais que X est une variable aléatoire réelle de densité f, définie sur un espace probabiisé $(\Omega, \mathcal{A}, P)$ et que toutes les variables aléatoires citées sont définies sur ce même espace.

2. a) Soit $Z = \mathrm{e}^{-X}$. Justifier que Z est une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$ et déterminer sa loi.

b) Soient x et y deux réels strictement positifs. Établir une relation entre la probabilité conditionnelle $P_{(X \leq -\ln x)}\big(X \leqslant -\ln(x+y)\big)$ et $P\big(X \leqslant -\ln y\big)$.

3. Soit $\big(Y_i\big)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, P)$, mutuellement indépendantes et de même loi exponentielle de paramètre 1.

Soit d'autre part L une variable aléatoire de loi de Poisson de paramètre 1 indépendante des variables aléatoires de la suite $\big(Y_i\big)_{i\in\mathbb{N}^*}$.

On définit S par :

Si $L(\omega) = 0$, alors $S(\omega) = 0$.
Si $L(\omega) = k$, avec $k \in \mathbb{N}^*$, alors $S(\omega) = \max(Y_1(\omega), \dots, Y_k(\omega))$.

a) Soit k un entier naturel non nul.
Déterminer la loi de la variable aléatoire $S_k = \max\big(Y_1, \dots, Y_k\big)$.

b) Démontrer que pour tous réels a et b tels que $0 < a < b$, on a :
$$P\big(a \leqslant S \leqslant b\big) = P\big(a \leqslant X \leqslant b\big)$$

c) Calculer $P\big(S = 0\big)$.

Solution :

1. a) La fonction $F : x \mapsto \exp(-\mathrm{e}^{-x})$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, strictement croissante. De plus $\lim\limits_{x\to-\infty} F(x) = 0$ et $\lim\limits_{x\to+\infty} F(x) = 1$, donc F est la fonction de répartition d'une variable aléatoire X à densité.

b) Une densité de X est $f = F'$ avec : pour tout réel x, $f(x) = \mathrm{e}^{-x}\exp(-\mathrm{e}^{-x})$.

2. a) Soit z réel.
Si $z \leqslant 0$ alors $(Z \leqslant z) = \emptyset \in \mathcal{A}$, si $z > 0$, on a $(Z \leqslant z) = (X \geqslant -\ln z) \in \mathcal{A}$, car X est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.

La fonction de répartition F_Z de Z est définie par : $F(z) = \begin{cases} 0 & \text{si } z < 0 \\ 1 - \mathrm{e}^{-z} & \text{si } z \geqslant 0 \end{cases}$.

On en déduit que $Z = \mathrm{e}^{-X}$ suit la loi exponentielle de paramètre 1.

b) On observe que pour tout $t > 0$,
$$P(X \leqslant -\ln t) = P(-X \geqslant \ln t) = P(\mathrm{e}^{-X} \geqslant t) = P(Z \geqslant t) = \mathrm{e}^{-t}$$
En conséquence pour tous x et y, réels strictement positifs, on a :
$$P_{[X\leq -\ln x]}(X \leqslant -\ln(x+y)) = P_{[Z\geq x]}(Z \geqslant x+y) = P(Z \geqslant y)$$
car Z suit une loi exponentielle, donc « sans mémoire », on en déduit
$$P_{[X\leq -\ln x]}(X \leqslant -\ln(x+y)) = P(Z \geqslant y) = \mathrm{e}^{-y}$$
Par ailleurs, $P(X \leqslant -\ln y) = P(Z \geqslant y) = \mathrm{e}^{-y}$, donc
$$P_{[X\leq -\ln x]}(X \leqslant -\ln(x+y)) = P(X \leqslant -\ln y)$$

3. a) Pour tout entier naturel k non nul, notons φ_k la fonction de répartition de S_k. Les variables aléatoires de la suite $(Y_i)_{i\in\mathbb{N}^*}$ étant mutuellement indépendantes de loi commune $\mathcal{E}(1)$, on a :

- $\forall\, u < 0$, $P(S_k < 0) = 0 = \varphi_k(u)$
- $\forall\, u \geqslant 0$, $\varphi_k(u) = P(S_k \leqslant u) = P\big(\bigcap\limits_{i=1}^{k}(Y_i \leqslant u)\big) = \prod\limits_{i=1}^{k} P(Y_i \leqslant u) = \prod\limits_{i=1}^{k}(1 - \mathrm{e}^{-u})$
$= (1 - \mathrm{e}^{-u})^k$.

La fonction φ_k ainsi définie détermine la loi de S_k.

b) Le système complet d'événements $((L = k))_{k\in\mathbb{N}}$ permet d'écrire :
$$P\big(a \leqslant S \leqslant b\big) = P\Big(\bigcup_{k=0}^{+\infty}(a \leqslant S \leqslant b) \cap (L = k)\Big)$$

$$= P((a \leqslant S \leqslant b) \cap (L = 0)) + P(\bigcup_{k=1}^{+\infty} (a \leqslant S \leqslant b) \cap (L = k))$$
$$= \sum_{k=1}^{+\infty} P((a \leqslant S \leqslant b) \cap (L = k))$$

car pour $(L = 0)$, par convention $S = 0$, donc $(a \leqslant S \leqslant b) \cap (L = 0) = \emptyset$.
D'autre part, pour tout entier naturel k non nul, les variables S_k et L sont indépendantes donc

$$P((a \leqslant S \leqslant b) \cap (L = k)) = P((a \leqslant \sup(Y_1, ..., Y_k) \leqslant b) \cap (L = k))$$
$$= P(a \leqslant S_k \leqslant b).P(L = k)$$

mais $P(a \leqslant S_k \leqslant b) = \varphi_k(b) - \varphi_k(a) = (1 - \mathrm{e}^{-b})^k - (1 - \mathrm{e}^{-a})^k$. On en déduit :

$$P(a \leqslant S \leqslant b) = \sum_{k=1}^{+\infty} ((1 - \mathrm{e}^{-b})^k - (1 - \mathrm{e}^{-a})^k) P(L = k)$$
$$= \sum_{k=1}^{+\infty} ((1 - \mathrm{e}^{-b})^k \frac{\mathrm{e}^{-1}}{k!} - \sum_{k=1}^{+\infty} ((1 - \mathrm{e}^{-a})^k \frac{\mathrm{e}^{-1}}{k!}$$

On peut ajouter alors les termes qui correspondraient à la valeur $k = 0$ (ils se détruisent) et on reconnaît des sommes de séries exponentielles, soit :

$$P(a \leqslant S \leqslant b) = \mathrm{e}^{-1} \exp(1 - \mathrm{e}^{-b}) - \mathrm{e}^{-1} \exp(1 - e^{-a}) = \exp(-\mathrm{e}^{-b}) - \exp(-\mathrm{e}^{-a})$$
$$= F(b) - F(a).$$

Donc, pour a, b tels que $0 < a < b$, $P(a \leqslant S \leqslant b) = P(a \leqslant X \leqslant b)$.

c) On a :

$$P(S = 0) = P(\bigcup_{k=0}^{+\infty} (S = 0) \cap (L = k))$$
$$= P((S = 0) \cap (L = 0)) + \sum_{k=1}^{+\infty} P((S = 0) \cap (L = k))$$

Or si $(L = 0)$ est réalisé, alors $(S = 0)$, l'est aussi donc

$$P((S = 0) \cap (L = 0)) = P((L = 0)) = \mathrm{e}^{-1},$$

alors que , pour tout k non nul, $(S = 0) \cap (L = k)) \subset (S_k = 0)$ et $P(S_k = 0) = 0$, d'où :

$$P(S = 0) = \mathrm{e}^{-1}$$

Exercice 3.05.

Soit $a \in \mathbb{R}_+^*$. La roue d'une loterie est représentée par un disque de rayon 1, dont le centre O est pris pour origine d'un repère orthonormé. Cette roue est lancée dans le sens trigonométrique, l'angle (exprimé en radians) dont elle tourne avant de s'arrêter est une variable aléatoire, notée U. On suppose que U suit la loi exponentielle de paramètre a.
La roue porte une marque M, qui, au départ, est située au point de coordonnées $(1, 0)$ et qui, après l'arrêt de la roue, se trouve au point de coordonnées aléatoires $X = \cos U$, $Y = \sin U$.

1. Soient $I = \displaystyle\int_0^{+\infty} \mathrm{e}^{-au} \cos u \, du$, $J = \displaystyle\int_0^{+\infty} \mathrm{e}^{-au} \sin u \, du$.

a) Montrer que les intégrales I et J sont convergentes.

b) A l'aide d'intégrations par parties, que l'on justifiera, établir deux relations liant I et J. En déduire les valeurs de I et J.

c) Calculer les espérances des variables aléatoires X et Y.

2. Un joueur gagne à cette loterie si, à l'arrêt de la roue, l'ordonnée de M vérifie la relation : $Y \geqslant \dfrac{1}{2}$.

a) Calculer la probabilité, notée $p(a)$, que le joueur gagne.

b) Déterminer $\lim\limits_{a\to 0} p(a)$.

Solution :

1. a) Montrons que les intégrales I et J sont absolument convergentes. On effet, on peut écrire : $|\cos u.\mathrm{e}^{-au}| \leqslant \mathrm{e}^{-au}, |\sin u.\mathrm{e}^{-au}| \leqslant \mathrm{e}^{-au}$, et $\displaystyle\int_0^{+\infty} \mathrm{e}^{-au}\,du$ existe.

D'où les conclusions.

b) En utilisant $A > 0$ et des intégrations par parties sur $[0, A]$, intervalle où les fonctions sinus, cosinus et exponentielle sont de classe C^∞, il vient :

$$I(A) = \int_0^A \cos u.\mathrm{e}^{-au}\,du = \sin A.\mathrm{e}^{-aA} + a(1 - \cos A.\mathrm{e}^{-aA}) - a^2 I(A)$$

$$J(A) = \int_0^A \sin u.\mathrm{e}^{-au}\,du = -a\sin A.\mathrm{e}^{-aA} + 1 - \cos A.\mathrm{e}^{-aA} - a^2 J(A)$$

En prenant la limite lorsque A tend vers l'infini, les fonctions sin et cos étant bornées sur $\mathbb{R}$, il vient :

$$I = a - a^2 I, \qquad J = 1 - a^2 J$$

Donc

$$I = \frac{a}{1+a^2}, \qquad J = \frac{1}{1+a^2}$$

c) Par le théorème du transfert :

$$E(X) = E(\cos U) = \int_a^{+\infty} \cos u.a\mathrm{e}^{-au}\,du = aI = \frac{a^2}{1+a^2}$$

$$E(Y) = E(\sin U) = \int_a^{+\infty} \sin u.a\mathrm{e}^{-au}\,du = aJ = \frac{a}{1+a^2}$$

2. a) Le joueur gagne si et seulement si $\sin U \geqslant 1/2$ soit si et seulement si $U \in [\pi/6, 5\pi/6] \pmod{2\pi}$. Donc :

$$p(a) = \sum_{k=0}^{\infty} P\Big(\frac{\pi}{6} + 2k\pi \leqslant U \leqslant \frac{5\pi}{6} + 2k\pi\Big) = \sum_{k=0}^{\infty} \int_{\pi/6+2k\pi}^{5\pi/6+2k\pi} a.\mathrm{e}^{-at}\,dt$$

$$= \sum_{k=0}^{\infty} \big(\mathrm{e}^{(-\pi/6-2k\pi)a} - \mathrm{e}^{(-5\pi/6-2k\pi)a}\big) = \frac{\mathrm{e}^{-a\pi/6} - \mathrm{e}^{-a5\pi/6}}{1 - \mathrm{e}^{-2\pi a}}$$

$$= \mathrm{e}^{-a\pi/6}\,\frac{1-\mathrm{e}^{-2a\pi/3}}{1-\mathrm{e}^{-2\pi a}}$$

b) Lorsque $a \to 0$, comme $\mathrm{e}^{-ax} = 1 - ax + o(x)$ au voisinage de 0, il vient :

$$\lim_{a\to 0} p(a) = \frac{1}{3}$$

Exercice 3.06.

On considère une variable aléatoire réelle discrète X définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{N}^*$ dont la loi est donnée par : $\forall\, k \in \mathbb{N}^*,\ P(X = k) = p_k \geqslant 0$.
On suppose que X admet une espérance $E(X)$.
Soit $(X_k)_k$ une suite de variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, P)$ indépendantes et de même loi que X. On note $(S_k)_{k\geq 0}$ la suite des sommes partielles définies par $S_0 = 0,$ et pour $n \geqslant 1, S_n = \sum\limits_{k=1}^{n} X_k$.
On étudie dans cet exercice la variable aléatoire $N(a,b)$ représentant le nombre d'éléments de la suite $(S_n)_{n\in\mathbb{N}}$ qui appartiennent à l'intervalle $[a,b]$, Cette variable $N(a,b)$ est donc définie par :

$$\forall\, \omega \in \Omega, N(a,b)(\omega) = \mathrm{card}\{k \in \mathbb{N}\ /\ S_k(\omega) \in [a,b]\} = \sum_{k=0}^{\infty} \mathbf{1}_{\{S_k\in[a,b]\}}(\omega)$$

1. Justifier pour tous $\ell \geqslant 0$ et $n \in \mathbb{N}$,

$$\begin{aligned}\{N(0,\ell) = n+1\} &= \{S_n \leq \ell < S_{n+1}\},\\ \{S_n \leq \ell\} &= \{N(0,\ell) \geqslant n+1\},\\ \{S_n \geqslant \ell\} &\subset \{N(0,\ell) \leq n+1\}.\end{aligned}$$

2. Soit Y une variable aléatoire à valeurs dans $\mathbb{N}$ admettant une espérance. Montrer que l'on a :

$$E(Y) = \sum_{k=1}^{\infty} P(Y \geqslant k)$$

3. a) Montrer pour $n \in \mathbb{N}$ et $\ell \geqslant 0$, les inégalités : $P(S_n \leqslant \ell) \leqslant E(\exp(\ell - S_n))$, puis : $P(S_n \leqslant \ell) \leqslant \mathrm{e}^{\ell}[E(\exp(-X))]^n$ (on commencera par justifier l'existence de ces espérances).

b) En déduire que $P(S_n \leqslant \ell)$ tend vers 0 quand $n \to \infty$ et que :

$$E(N(0,\ell)) \leqslant \frac{\mathrm{e}^{\ell}}{1 - E(\exp(-X))}$$

Solution :

1. $\star$ L'égalité $(N(0,\ell) = n+1) = (S_n \leqslant \ell < S_{n+1})$ est une tautologie.

$\star$ Pour tout $\omega \in \Omega$, la suite $k \mapsto S_k(\omega)$ est strictement croissante de premier terme nul. Dire que l'on réalise $S_n(\omega) \leqslant \ell$, c'est dire que tous les nombres $S_0(\omega)$,

$S_1(\omega), \ldots, S_n(\omega)$ sont majorés par ℓ, donc que la durée du passage entre 0 et ℓ est au moins égale à $n+1$:

$$(S_n \leqslant \ell) = (N(0,\ell) \geqslant n+1)$$

$\star$ X ne prenant que des valeurs strictement positives, $(S_n \geq \ell) \subset (S_{n+1} > \ell)$ et donc on est sorti du domaine $[0,\ell]$ au plus tard au rang n (on commence au rang 0) et donc

$$(S_n \geqslant \ell) \subset (N(0,\ell) \leqslant n+1)$$

2. *Méthode 1* : en passant par la notion de famille sommable. La variable Y ayant une espérance, on a :

$$E(Y) = \sum_{i=0}^{\infty} iP(Y=i) = \sum_{i=1}^{\infty} iP(Y=i) = \sum_{i=1}^{\infty}\sum_{j=1}^{i} P(Y=i)$$

Cette famille étant une famille dénombrable à termes positifs ou nuls la sommation peut se faire par paquets quelconques et donc :

$$E(Y) = \sum_{j=1}^{\infty}\sum_{i=j}^{\infty} P(Y=i) = \sum_{j=1}^{\infty} P(Y \geqslant j)$$

Méthode 2 : à la main. Pour tout $N \in \mathbb{N}^*$, comme $(Y \geqslant k) = (Y = k) \cup (Y \geqslant k+1)$ et cette réunion étant disjointe :

$$\begin{aligned}
\sum_{k=0}^{N} kP(Y=k) &= \sum_{k=0}^{N} k(P(Y \geqslant k) - P(Y \geqslant k+1)) \\
&= \sum_{k=0}^{N} kP(Y \geqslant k) - \sum_{k=0}^{N} kP(Y \geqslant k+1) \\
&= \sum_{k=0}^{N} kP(Y \geqslant k) - \sum_{k=1}^{N-1} (k-1)P(Y \geqslant k) \\
&= \sum_{k=1}^{N-1} P(Y \geqslant k) + NP(Y \geqslant N)
\end{aligned}$$

Mais : $NP(Y \geqslant N) = \sum\limits_{k=N}^{\infty} NP(Y=k) \leqslant \sum\limits_{k=N}^{\infty} kP(Y=k) \underset{N\to\infty}{\longrightarrow} 0$ (reste d'une série supposée convergente). On peut donc passer à la limite lorsque N tend vers l'infini et on obtient bien le même résultat.

3. a) $\star$ $(S_n \leqslant \ell) = (\ell - S_n \geqslant 0) = (\mathrm{e}^{\ell - S_n} \geqslant 1)$, donc $P(S_n \leqslant \ell) = P(\mathrm{e}^{\ell - S_n} \geqslant 1)$.

Comme $\mathrm{e}^{\ell - S_n}$ est une variable aléatoire bornée, elle admet une espérance, et comme elle est à valeurs positives, l'inégalité de Markov donne : $P(\mathrm{e}^{\ell - S_n} \geqslant 1) \leqslant \dfrac{1}{1}E(\mathrm{e}^{\ell - S_n})$, donc :

$$P(S_n \leqslant \ell) \leqslant E(\mathrm{e}^{\ell - S_n})$$

$\star$ D'autre part $E(\mathrm{e}^{\ell - S_n}) = \mathrm{e}^{\ell}E(\mathrm{e}^{-X_1 - X_2 - \cdots - X_n}) = \mathrm{e}^{\ell}E(\mathrm{e}^{-X_1}\mathrm{e}^{-X_2}\ldots\mathrm{e}^{-X_n})$

L'indépendance des différentes variables aléatoires X_k donne l'indépendance des différentes variables aléatoires e^{-X_k}, donc l'espérance du produit est le produit des espérances, et comme en plus elles ont même loi :

$$E(\mathrm{e}^{\ell - S_n}) = \mathrm{e}^{\ell}E(\mathrm{e}^{-X_1}) \times \ldots \times E(\mathrm{e}^{-X_n}) = \mathrm{e}^{\ell}[E(\mathrm{e}^{-X})]^n$$

$$P(S_n \leqslant \ell) \leqslant \mathrm{e}^{\ell}[E(\mathrm{e}^{-X})]^n$$

La formule obtenue reste valable pour $n = 0$.

b) X est à valeurs strictement positives, donc e^{-X} est à valeurs dans l'ouvert $]0,1[$ et par transfert : $E(\mathrm{e}^{-X}) = \sum_n \mathrm{e}^{-x_n} p_n < \sum_n p_n = 1$ (une somme finie ou infinie d'inégalités dont l'une au moins est stricte est une inégalité stricte).

Ainsi $\lim_{n\to\infty} [E(\mathrm{e}^{-X})]^n = 0$ et :

$$\lim_{n\to\infty} P(S_n \leqslant \ell) = 0$$

Ensuite, $P(N(0,\ell) \geqslant n) = P(S_{n-1} \leqslant \ell) \underset{n\to\infty}{\longrightarrow} 0$ donc par continuité décroissante :

$$P(N(0,\ell) = \infty) = 0$$

Enfin,

$$E(N(0,\ell)) = \sum_{n=1}^{\infty} P(N(0,\ell) \geqslant n) \leqslant \sum_{n=1}^{\infty} \mathrm{e}^{\ell}[E(\mathrm{e}^{-X})]^{n-1} = \frac{\mathrm{e}^{\ell}}{1 - E(\mathrm{e}^{-X})}$$

Exercice 3.07.

Toutes les variables aléatoires de cet exercice sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et suivant chacune la loi géométrique $\mathcal{G}(\theta)$ de paramètre θ. On suppose $\theta \in]0,1[$ inconnu et on désire l'estimer.

Pour tout $n \in \mathbb{N}^*$, on note $S_n = \sum_{i=1}^{n} X_i$, $M_n = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{X_i}$ et $T_n = \frac{n}{S_n}$.

1. Rappeler les valeurs de l'espérance et de la variance de X_1.

2. Soit $\varphi : x \mapsto \ln(1-x)$ définie sur $[0,1[$.

a) En appliquant la formule de Taylor avec reste intégral, montrer que :

$$\forall\, x \in]0,1[, \ln(1-x) + \sum_{k=1}^{n} \frac{x^k}{k} = -\int_0^x \frac{1}{1-t}\Big(\frac{x-t}{1-t}\Big)^n dt$$

b) En déduire que pour $x \in]0,1[$, $\ln(1-x) = -\sum_{k=1}^{+\infty} \frac{x^k}{k}$.

3. Calculer $E(M_n)$ et montrer que $E(M_n) > \theta$.

4. Montrer que la loi de S_n est donnée par :

$$\text{pour tout entier } k \geqslant n : P(S_n = k) = \binom{k-1}{n-1}\theta^n(1-\theta)^{k-n}$$

5 a). Montrer que $E(T_n) = \sum_{k=n}^{+\infty} \frac{n}{k}\binom{k-1}{n-1}\theta^n(1-\theta)^{k-n}$.

b) Montrer que $E(T_n) > \theta$.

6. a) Montrer que pour tout $n-$uplet $(x_1, x_2, \cdots, x_n) \in (\mathbb{R}_+^*)^n$, on a :

$$\Big(\sum_{k=1}^{n} x_k\Big)\Big(\sum_{k=1}^{n} \frac{1}{x_k}\Big) \geqslant n^2$$

b) Comparer $E(M_n)$ et $E(T_n)$.

Solution :

1. On a $E(X_1) = \dfrac{1}{\theta}$ et $V(X_1) = \dfrac{1-\theta}{\theta^2}$.

2. a) Soit $n \in \mathbb{N}$ et $x \in [0,1[$; la fonction φ est de classe C^{n+1} sur $[0,1[$ et :

$$\forall\, k \in \mathbb{N}^*, \forall\, t \in [0,1[, \varphi^k(t) = -\frac{(k-1)!}{(1-t)^k}$$

En appliquant la formule de Taylor avec reste intégral à l'ordre n entre 0 et x, on obtient :

$$\ln(1-x) + \sum_{k=1}^{n} \frac{x^k}{k} = -\int_0^x \frac{(x-t)^n}{n!} \times \frac{n!}{(1-t)^{n+1}} dt = -\int_0^x \frac{1}{1-t}\Big(\frac{x-t}{1-t}\Big)^n dt$$

b) L'étude de la fonction $t \mapsto \dfrac{x-t}{1-t}$ montre qu'elle admet un maximum en 0 égal à x. On a donc :

$$\forall\, t \in [0,x], \frac{1}{1-t}\Big(\frac{x-t}{1-t}\Big)^n \leqslant \frac{x^n}{1-x}$$

Par croissance de l'intégration avec $0 \leqslant x$, il vient :

$$\Big|\ln(1-x) + \sum_{k=1}^{n} \frac{x^k}{k}\Big| \leqslant \int_0^x \frac{x^n}{1-x} dt = \frac{x^{n+1}}{1-x}$$

Par encadrement, la série de terme général $\dfrac{x^k}{k}$ converge et $\displaystyle\sum_{k=1}^{+\infty} \frac{x^k}{k} = -\ln(1-x)$.

3. Sous réserve de convergence et en utilisant le théorème de transfert :

$$E\Big(\frac{1}{X_1}\Big) = \sum_{k=1}^{+\infty} \frac{1}{k}\theta(1-\theta)^{k-1} = \sum_{k=1}^{+\infty} \frac{\theta}{1-\theta} \times \frac{(1-\theta)^k}{k}$$

Comme $1-\theta \in]0,1[$, alors d'après la question précédente $E\Big(\dfrac{1}{X_1}\Big)$ existe et vaut : $\dfrac{-\theta\ln\theta}{1-\theta}$. Par linéarité, on obtient :

$$E(M_n) = \frac{-\theta\ln\theta}{1-\theta}$$

4. Classique (temps d'attente du $k^{\text{ème}}$ succès) : il suffit de placer $k-1$ succès parmi les $n-1$ premiers essais et de passer aux probabilités de chaque événement élémentaire. . . .

5 a). T_n est une variable aléatoire réelle presque sûrement à valeurs dans $]0,1]$, donc elle est presque sûrement bornée et admet une espérance. D'après le théorème de transfert,

$$E(T_n) = \sum_{k=n}^{+\infty} \frac{n}{k} P(S_n = k) = \sum_{k=n}^{+\infty} \frac{n}{k}\binom{k-1}{n-1}\theta^n(1-\theta)^{k-n}$$

b) On remarque que pour tout entier k tel que $2 \leqslant n < k$, $\dfrac{n-1}{k-1} < \frac{n}{k}$, donc $\displaystyle\sum_{k=n}^{+\infty}\binom{k-2}{n-2}\theta^n(1-\theta)^{k-n} < E(T_n)$, et en effectuant le changement d'indice $j = k-1$, il vient :

$$\theta \sum_{j=n-1}^{+\infty} \binom{j-1}{n-2} \theta^{n-1}(1-\theta)^{j-(n-1)} < E(T_n)$$

Donc $\theta \sum\limits_{j=n-1}^{+\infty} P([S_{n-1} = j]) < E(T_n)$, d'où $E(T_n) > \theta$.

6. a) D'après l'inégalité de Cauchy-Schwarz pour le produit scalaire canonique de $\mathbb{R}^n$, pour tout $n-$uplet $(x_1, x_2, \ldots, x_n) \in (\mathbb{R}_+^*)^n$:

$$\big(\sum_{k=1}^{n} x_k\big)\big(\sum_{k=1}^{n} \frac{1}{x_k}\big) \geqslant \big(\sum_{k=1}^{n} \sqrt{x_k} \times \frac{1}{\sqrt{x_k}}\big)^2 = n^2$$

b) On a $M_n \geqslant T_n$, puis par croissance de l'espérance, $E(M_n) \geqslant E(T_n)$.

Exercice 3.08.

Toutes les variables aléatoires de cet exercice sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

Pour tout $n \in \mathbb{N}$, on note D_n le diamètre d'un tronc d'arbre à la fin de l'année numéro n et l'on suppose que sa croissance suit le modèle décrit ci-dessous :

▷ le diamètre initial D_0 est tel que $D_0 > 0$;

▷ pour tout $n \in \mathbb{N}$, $D_{n+1} = D_n + X_{n+1} D_n$, où X_{n+1} est une variable aléatoire représentant les divers facteurs extérieurs (maladies, climat, ...) ;

▷ on suppose que les variables aléatoires X_k sont à densité, indépendantes et de même loi à valeurs dans $[0, 1]$ et de densité f.

On pose, pour tout $n \in \mathbb{N}^*$, $Q_n = \big(\frac{D_n}{D_0}\big)^{1/n}$ et $m = E(\ln(1 + X_1))$, où E désigne l'espérance.

1. Soit $n \in \mathbb{N}^*$. Exprimer D_n en fonction de D_0 et de $X_1, \ldots, X_n$.

2. a) Soit $n \in \mathbb{N}^*$. Calculer $E(Q_n)$ en fonction de $E((1 + X_1)^{1/n})$.

b) Montrer l'existence d'une constante $C > 0$ telle que, pour tout $y \in [0, 1]$, $|e^y - 1 - y| \leqslant C y^2$.

c) En déduire que $E(Q_n) \underset{n \to \infty}{\longrightarrow} \exp(m)$.

3. a) Montrer l'existence d'une constante $L > 0$ telle que pour tout $(x, y) \in [0, 1]^2$, $|e^x - e^y| \leqslant L|x - y|$.

b) Soit $\varepsilon > 0$. Prouver l'existence d'une constante $M > 0$ telle que, pour tout $n \in \mathbb{N}^*$,

$$P(|Q_n - e^m| \geqslant \varepsilon) \leqslant \frac{M}{n}$$

c) Que peut-on en déduire pour la suite $(Q_n)_{n \geq 1}$?

Solution :

Notons que X_1 prend ses valeurs dans $[0, 1]$, donc $\ln(1 + X_1)$ prend ses valeurs dans $[0, \ln 2]$ et étant bornée elle admet des moments de tous ordres.

1. $D_n = D_0 \prod_{k=1}^{n} (1 + X_k)$ pour tout $n \in \mathbb{N}^*$.

2. a) Soit $n \in \mathbb{N}^*$. On a $E(Q_n) = E\big(\prod_{k=1}^{n} (1 + X_k)^{1/n}\big)$, et donc, par indépendance des $(1 + X_k)^{1/n}$ et équidistribution des X_k, $E(Q_n) = (E((1 + X_1)^{1/n}))^n$.

b) Soit $y \in [0, 1]$. Par exemple par la formule de Taylor avec reste intégral en 0 :

$$\mathrm{e}^y - 1 - y = \int_0^y (y - t)\mathrm{e}^t\, dt$$

Pour tout $t \in [0, y]$, on a $0 \leqslant (y-t)\mathrm{e}^t \leqslant \mathrm{e}(y-t)$ et donc, par inégalité triangulaire et croissance de l'intégration :

$$|\mathrm{e}^y - 1 - y| \leqslant \int_0^y \mathrm{e}(y - t)\, dt = \frac{\mathrm{e}}{2}\, y^2$$

la valeur absolue étant d'ailleurs inutile.

c) Soit n un entier naturel non nul. Par la question précédente,

$$\Big|(1 + X_1)^{1/n} - 1 - \frac{1}{n} \ln(1 + X_1)\Big| \leqslant \frac{C \ln^2(1 + X_1)}{n^2}$$

et donc, par croissance de l'espérance et inégalité triangulaire pour l'espérance :

$$\Big|E\left((1 + X_1)^{1/n}\right) - 1 - \frac{1}{n} E(\ln(1 + X_1))\Big| \leqslant \frac{\lambda}{n^2}$$

où $\lambda = C.E(\ln^2(1 + X_1))$. On en déduit que, à partir d'un certain rang,

$$\Big(1 + \frac{1}{n} E(\ln(1 + X_1)) - \frac{\lambda}{n^2}\Big)^n \leqslant E\left((1 + X_1)^{1/n}\right)^n \leqslant \Big(1 + \frac{1}{n} E(\ln(1 + X_1)) + \frac{\lambda}{n^2}\Big)^n$$

Par passage au logarithme et utilisation de l'équivalent classique $\ln(1 + u) \underset{(0)}{\sim} u$, on obtient :

$\lim\limits_{n\to\infty} E\left((1 + X_1)^{1/n}\right)^n = \mathrm{e}^m$, soit $\lim\limits_{n\to\infty} E(Q_n) = \mathrm{e}^m$.

3. a) La constante $L = \mathrm{e}$ convient par l'inégalité des accroissements finis.

b) Soient $n \in \mathbb{N}^*$ et $\varepsilon > 0$. On note $S_n = \frac{1}{n} \sum_{k=1}^{n} \ln(1 + X_k)$.

On a $Q_n = \mathrm{e}^{S_n}$ donc $(|Q_n - \mathrm{e}^m| \geqslant \varepsilon) \subset \Big(|S_n - m| \geqslant \frac{\varepsilon}{L}\Big)$. Comme $E(S_n) = m$, on déduit de l'inégalité de Tchebychev que

$$P(|Q_n - e^m| \geqslant \varepsilon) \leqslant P\left(|S_n - m| \geqslant \frac{\varepsilon}{L}\right) \leqslant \frac{V(S_n)L^2}{\varepsilon^2} = \frac{V(\ln(1 + X_1))L^2}{n\varepsilon^2}$$

par indépendance et équidistribution des variables X_k.

c) On en déduit que la suite $(Q_n)_{n\geq 1}$ converge en probabilité vers la variable certaine égale à e^m.

Exercice 3.09.

Soit α un réel strictement positif et $(Y_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$. On suppose de plus que pour tout i, Y_i suit la loi exponentielle de paramètre $i{\times}\alpha$.

Pour tout n de $\mathbb{N}^*$, on pose $Z_n = \sum\limits_{i=1}^{n} Y_i$ et on note g_n la densité de Z_n nulle sur $\mathbb{R}_-$ et continue sur $\mathbb{R}_+^*$.

1. a) Déterminer la fonction g_2.

b) Montrer que pour $n \geqslant 1$ et $x > 0$, on a : $g_n(x) = n\alpha \mathrm{e}^{-\alpha x}(1 - \mathrm{e}^{-\alpha x})^{n-1}$.

c) Calculer l'espérance de Z_n et en donner un équivalent simple lorsque n tend vers l'infini.

d) Calculer la variance de Z_n et montrer qu'elle admet une limite finie lorsque n tend vers l'infini.

2. Pour $n \in \mathbb{N}^*$, on pose $U_n = \dfrac{1}{n} Z_n$.

a) Déterminer la fonction de répartition H_n de U_n.

b) Montrer que la suite $(U_n)_n$ converge en loi et déterminer la loi limite.

c) Déterminer la limite quand n tend vers l'infini de $E(U_n)$ et $V(U_n)$.

Solution :

1. a) Les variables Y_1 et Y_2 sont indépendantes. On a $g_1(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \alpha \mathrm{e}^{-\alpha x} & \text{si } x > 0 \end{cases}$

Par convolution, on a :

$$g_2(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \displaystyle\int_{-\infty}^{+\infty} g_{Y_1}(t) g_{Y_2}(x-t)\, dt = \int_0^x 2\alpha^2 \mathrm{e}^{-\alpha t} \mathrm{e}^{-2\alpha(x-t)} dt & \text{si } x > 0 \end{cases}$$

Donc pour $x > 0$,

$$g_2(x) = 2\alpha^2 \mathrm{e}^{-2\alpha x} \int_0^x \mathrm{e}^{\alpha t}\, dt = 2\alpha \mathrm{e}^{-2\alpha x}(\mathrm{e}^{\alpha x} - 1) = 2\alpha \mathrm{e}^{-\alpha x}(1 - \mathrm{e}^{-\alpha x}).$$

b) Le cas $n = 1$ est trivial et le cas $n = 2$ vient d'être vu.

On suppose donc le résultat acquis pour un certain rang n et on passe au rang suivant, en supposant évidemment $x > 0$. Comme $Z_{n+1} = Z_n + Y_{n+1}$ et que Y_{n+1} est indépendante de Z_n (lemme des coalitions), il vient :

$$\begin{aligned} g_{n+1}(x) &= \int_0^x g_n(t) g_{Y_{n+1}}(x-t)\, dt \\ &= \int_0^x n\alpha \mathrm{e}^{-\alpha t}(1 - \mathrm{e}^{-\alpha t})^{n-1} \times (n+1)\alpha \mathrm{e}^{-(n+1)\alpha(x-y)}\, dt \\ &= (n+1)\alpha^2 \mathrm{e}^{-(n+1)\alpha x} \int_0^x n \mathrm{e}^{\alpha t}(\mathrm{e}^{\alpha t} - 1)^{n-1}\, dt \\ &= (n+1)\alpha \mathrm{e}^{-(n+1)\alpha x} \left[(\mathrm{e}^{\alpha t} - 1)^n\right]_0^x = (n+1)\alpha \mathrm{e}^{-\alpha x}(1 - \mathrm{e}^{-\alpha x})^n \end{aligned}$$

ce qui est le résultat attendu.

c) Par linéarité : $E(Z_n) = \sum\limits_{k=1}^{n} E(Y_k) = \sum\limits_{k=1}^{n} \dfrac{1}{k\alpha} = \dfrac{1}{\alpha} \sum\limits_{k=1}^{n} \dfrac{1}{k}$

Par comparaison série intégrale, on en déduit classiquement :

$$E(Z_n) \sim \frac{1}{\alpha}\ln n$$

d) Par indépendance : $V(Z_n) = \dfrac{1}{\alpha^2}\sum\limits_{k=1}^{n}\dfrac{1}{k^2}$ et par convergence connue de la série de Riemann d'indice 2, ceci a bien une limite quand n tend vers l'infini.

2. a) Pour $x > 0$, on a : $P(Z_n \leqslant x) = \displaystyle\int_0^x g_n(t)\,dt = \big[(1-\mathrm{e}^{-\alpha t})^n\big]_0^x = (1-\mathrm{e}^{-\alpha x})^n$.
D'où $H_n(x) = P(Z_n \leqslant nx) = (1-\mathrm{e}^{-n\alpha x})^n$.
Toujours pour $x > 0$, $\ln(H_n(x)) = n\ln(1-\mathrm{e}^{-n\alpha x}) \underset{(n\to\infty)}{\sim} -n\mathrm{e}^{-n\alpha x} \underset{n\to\infty}{\longrightarrow} 0$, donc $H_n(x)$ tend vers 1.

Ainsi : $\lim\limits_{n\to\infty} H_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 & \text{si } x > 0 \end{cases}$ et sur $\mathbb{R}^*$ (domaine de continuité de la fonction limite) la limite de H_n est la fonction de répartition d'une variable constante égale à 0. On a ainsi montré que la suite $(Z_n/n)_n$ converge en loi vers 0.

c) $E(U_n) \underset{(n\to\infty)}{\sim} \dfrac{\ln n}{n\alpha} \underset{n\to\infty}{\longrightarrow} 0$ et $V(U_n) = \dfrac{1}{n^2\alpha^2}\sum\limits_{k=1}^{n}\dfrac{1}{k^2} \underset{n\to\infty}{\longrightarrow} 0$.

Exercice 3.10.

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes et définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On suppose que, pour chaque $n \in \mathbb{N}^*$, la variable $Y_n = \dfrac{1+X_n}{2}$ suit la loi de Bernoulli de paramètre p_n (avec $p_n \in\,]0,1[$). Pour tout entier $n \geqslant 1$, on définit la variable aléatoire S_n par : $S_n = \sum\limits_{k=1}^{n} X_k$.

1. Quelle est la loi de X_n ?

2. Peut-on appliquer la loi faible des grands nombres à la suite de variables aléatoires $(X_n)_{n\geq 1}$?

3. Calculer l'espérance et la variance de la variable aléatoire S_n.

4. Montrer que $0 \leqslant p_k(1-p_k) \leqslant \dfrac{1}{4}$ pour tout $k \geqslant 1$. En déduire que la variance $V(S_n)$ est majorée par n.

5. Soit $\varepsilon > 0$, montrer que $P\big(\big|\dfrac{S_n}{n} - \dfrac{E(S_n)}{n}\big| \geqslant \varepsilon\big) \leqslant \dfrac{4}{n^2\varepsilon^2}\sum\limits_{k=1}^{n} p_k(1-p_k) \leqslant \dfrac{1}{n\varepsilon^2}$.
Que peut-on en conclure ?

6. Dans cette question, on suppose que, pour tout $n \in \mathbb{N}^*$, $p_n = 1 - \dfrac{1}{2^{n+2}}$.

Soit L un entier fixé. On introduit l'événement A_L qui est égal à l'ensemble des $\omega \in \Omega$ pour lesquels la suite $(S_n(\omega))$ contient une infinité de termes qui se trouvent dans le segment $[-L, L]$.

a) Etudier la suite $\big(\dfrac{E(S_n)}{n}\big)_{n\geq 1}$. En déduire l'existence d'un entier $N > 2L$ tel que pour tout $n \geqslant N$ on a : $\big|\dfrac{E(S_n)}{n} - 1\big| \leqslant \dfrac{1}{4}$.

b) Montrer que pour $n \geqslant N$ on a $\Big[\big|\frac{S_n}{n} - 1\big| \geqslant \frac{1}{2}\Big] \subseteq \Big[\big|\frac{S_n}{n} - \frac{E(S_n)}{n}\big| \geqslant \frac{1}{4}\Big]$.
En déduire que $P\big(\big|\frac{S_n}{n} - 1\big| \geqslant \frac{1}{2}\big) \leqslant \frac{16}{n^2}$.

c) Soit $n \geqslant N$, prouver que

$$\Big[|S_n| \leqslant L\Big] \subseteq \Big[\big|\frac{S_n}{n} - 1\big| \geqslant \frac{1}{2}\Big] \text{ et que } A_L \subseteq \bigcup_{k \geq n} \Big[|S_k| \leqslant L\Big].$$

d) Etablir que pour tout $n \geqslant N$, on a $P(A_L) \leqslant 16 \sum\limits_{k=n}^{+\infty} \frac{1}{k^2}$.

(On pourra utiliser sans démonstration le fait que $P(E_1 \cup \cdots \cup E_m) \leqslant \sum\limits_{k=1}^{m} P(E_k)$ pour tout m-uplet d'événements $(E_1, \ldots, E_m)$.) En déduire la valeur de $P(A_L)$.

Solution :

1. On a $X_n(\Omega) = \{-1, 1\}$ avec $P(X_n = -1) = 1 - p_n$ et $P(X_n = 1) = p_n$.

2. On ne peut pas appliquer la loi faible des grands nombres car les p_n ne sont pas supposés égaux.

3. On a $E(S_n) = 2 \sum\limits_{k=1}^{n} p_k - n$, et par indépendance :

$$V(S_n) = 4 \sum_{k=1}^{n} V(Y_n) = 4 \sum_{k=1}^{n} p_k(1 - p_k)$$

4. Par exemple, le maximum sur $[0, 1]$ de la fonction $t \mapsto t(1-t)$ est $\frac{1}{4}$ et on utilise 3.

5. Avec l'inégalité de Bienaymé-Tchebychev et la question précédente, il vient :

$$P\big(\big|\frac{S_n}{n} - \frac{E(S_n)}{n}\big| \geqslant \varepsilon\big) \leqslant \frac{V(S_n/n)}{\varepsilon^2} = \frac{4}{n^2\varepsilon^2} \sum_{k=1}^{n} p_k(1 - p_k) \leqslant \frac{1}{n\varepsilon^2}.$$

La suite de variables $\big(\frac{S_n}{n} - \frac{E(S_n)}{n}\big)_{n \geq 1}$ converge donc en probabilité vers 0.

6. a) On a : $\frac{E(S_n)}{n} = \frac{2}{n} \sum\limits_{k=1}^{n} \big(1 - \frac{1}{2^{k+2}}\big) - 1 = 1 - \frac{1}{2n}\big(1 - \frac{1}{2^n}\big) \underset{n \to \infty}{\longrightarrow} 1.$

L'existence de l'entier N s'obtient en « jouant » $1/4$ sur la limite et on peut toujours le choisir tel que $N > 2L$.

b) Si $\big[\big|\frac{S_n}{n} - 1\big| \geqslant \frac{1}{2}\big]$ et si $n \geqslant N$, avec l'inégalité triangulaire, il vient :

$$\frac{1}{2} \leqslant \big|\frac{S_n}{n} - 1\big| \leqslant \big|\frac{S_n}{n} - \frac{E(S_n)}{n}\big| + \frac{1}{4}$$

L'inclusion demandée en découle directement. Par suite, en utilisant aussi la première inégalité de 5. il vient :

$$\begin{aligned} P\big(\big|\frac{S_n}{n} - 1\big| \geqslant \frac{1}{2}\big) \leqslant P\big(\big|\frac{S_n}{n} - \frac{E(S_n)}{n}\big| \geqslant \frac{1}{4}\big) &\leqslant \frac{64}{n^2} \sum_{k=1}^{n} p_k(1 - p_k) \leqslant \frac{64}{n^2} \sum_{k=1}^{n} \frac{1}{2^{k+2}} \\ &\leqslant \frac{16}{n^2}. \end{aligned}$$

c) Si $n \geqslant N > \dfrac{L}{2}$, l'inclusion $\big[|S_n| \leqslant L\big] \subseteq \big[|\dfrac{S_n}{n} - 1| \geqslant \dfrac{1}{2}\big]$ s'obtient facilement en utilisant l'inégalité triangulaire ou par calculs directs sur les inégalités.
Maintenant, si $\omega \in A_L$, il existe une infinité de $S_k(\omega)$ qui se trouvent dans le segment $[-L, L]$, donc au moins un pour lequel $k \geqslant n$, il est alors clair que ω appartient au membre de droite de l'inclusion souhaitée.

d) En utilisant l'inégalité mentionnée, les limites monotones et les deux questions précédentes, il vient :

$$P(A_L) \leqslant P\big(\bigcup_{k \geq n} (|S_k| \leqslant L)\big) = \lim_{m \to +\infty} P\big(\bigcup_{n \leq k \leq m} (|S_k| \leqslant L)\big) \leqslant \lim_{m \to +\infty} 16 \sum_{k=n}^{m} \frac{1}{k^2}$$

Donc :

$$P(A_L) \leqslant 16 \sum_{k=n}^{+\infty} \frac{1}{k^2}.$$

Comme le reste d'une série convergente tend vers 0, on a nécessairement $P(A_L) = 0$. Ainsi presque sûrement, on sort de l'intervalle $[-L, L]$ à un moment donné pour ne jamais y revenir.

Exercice 3.11.

Dans tout l'exercice, d désigne un entier naturel tel que $d \geqslant 2$.

1. Soit x un réel strictement positif. Déterminer $\displaystyle\lim_{d \to +\infty} \frac{1}{d} \sum_{k=1}^{\lfloor x\sqrt{d} \rfloor} k$, où $\lfloor . \rfloor$ désigne la fonction « partie entière ».

2. a) Montrer que la fonction f définie sur $]-1, 1[\ \setminus \{0\}$ par $f(x) = \dfrac{\ln(1-x) + x}{x^2}$ admet un prolongement par continuité en 0.

b) Déterminer $\displaystyle\lim_{d \to +\infty} \sum_{k=1}^{\lfloor x\sqrt{d} \rfloor} \ln\big(1 - \frac{k}{d}\big)$.

Soit $(U_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ suivant toute la loi uniforme discrète sur $[\![1, d]\!]$.
Soit N_d la variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$ par, pour tout $\omega \in \Omega$:

$$N_d(\omega) = \inf\{i \geq 2 \ /\ U_i(\omega) \in \{U_1(\omega), U_2(\omega), \ldots, U_{i-1}(\omega)\}\}$$

3. a) Montrer que pour tout $n \in \mathbb{N}$ tel que $2 \leqslant n \leqslant d$, on a :

$$P(N_d > n) = \big(1 - \frac{1}{d}\big)\big(1 - \frac{2}{d}\big) \cdots \big(1 - \frac{n-1}{d}\big)$$

b) Déterminer la limite en loi de $\big(\dfrac{N_d}{\sqrt{d}}\big)$ lorsque d tend vers l'infini.

4. a) Montrer que $E(N_d) = \displaystyle\sum_{k=0}^{d} P(N_d > k)$.

b) En déduire que $E(N_d) = \displaystyle\int_0^{+\infty} \big(1 + \frac{t}{d}\big)^d \mathrm{e}^{-t} dt$.

Solution :

1. On sait que $\sum_{k=1}^{n} k = \dfrac{n(n+1)}{2}$. Donc

$$\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} k = \frac{(\lfloor x\sqrt{d}\rfloor)(\lfloor x\sqrt{d}\rfloor + 1)}{2} \sim \frac{(\lfloor x\sqrt{d}\rfloor)^2}{2} \sim \frac{dx^2}{2}$$

En effet, $\lfloor x\sqrt{d}\rfloor \leqslant x\sqrt{d} < \lfloor x\sqrt{d}\rfloor + 1$, ce qui entraîne $\lfloor x\sqrt{d}\rfloor \sim x\sqrt{d}$. Finalement :

$$\lim_{d\to+\infty} \frac{1}{d} \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} k = \frac{x^2}{2}$$

2. a) Le développement limité de $\ln(1-x)$ à l'ordre 2 montre qu'il faut poser $f(0) = -1/2$.

b) Comme $1 \leqslant k \leqslant x\sqrt{d}$, on a $\dfrac{1}{d} < \dfrac{k}{d} \leqslant \dfrac{x}{\sqrt{d}}$, proche de 0 lorsque d tend vers ∞.

En utilisant la question a), pour $\varepsilon > 0$, pour d assez grand et k tel que $1 \leqslant k \leqslant x\sqrt{d}$, on a :

$$\left| \ln\left(1 - \frac{k}{d}\right) + \frac{k}{d} + \frac{k^2}{2d^2} \right| \leqslant \varepsilon \frac{k^2}{2d^2}$$

Donc

$$\left| \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1 - \frac{k}{d}\right) + \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k}{d} + \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{2d^2} \right| \leqslant \varepsilon \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{2d^2}$$

En utilisant la méthode de la première question, avec $\sum_{k=1}^{n} k^2 = \dfrac{n(n+1)(2n+1)}{6} \sim \dfrac{n^3}{3}$, il vient :

$$\sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \frac{k^2}{2d^2} \sim \frac{x^3}{6\sqrt{d}}$$

soit :

$$\left| \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1 - \frac{k}{d}\right) + \frac{x^2}{2} \right| \leqslant \frac{C}{d^{1/2}} \text{ et } \lim_{d\to+\infty} \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor} \ln\left(1 - \frac{k}{d}\right) = -\frac{x^2}{2}$$

3. a) On a $(N_d(\omega) = i)$ lorsque les nombres $U_1(\omega), \ldots, U_{i-1}(\omega)$ sont distincts et que $U_i(\omega)$ est égal à l'un des $U_j(\omega)$ précédents.

Ainsi $U_1(\omega)$ est un entier quelconque de $[\![1, d]\!]$, $U_2(\omega)$ est un entier différent, $U_3(\omega)$ est un entier différent des deux premiers, etc. Donc, en utilisant la formule des probabilités composées

$$P(N_d > n) = \left(1 - \tfrac{1}{d}\right)\left(1 - \tfrac{2}{d}\right) \cdots \left(1 - \tfrac{n-1}{d}\right) = \frac{(d-1)(d-2)\ldots(d-n+1)}{d^{n-1}}$$

On peut alors faire apparaître des factorielles, ce qui servira pour la dernière égalité de ce corrigé ...

$$n \in [\![2, d]\!] \implies P(N_d > n) = \frac{d!}{d^n (d-n)!}$$

b) Soit $x > 0$. On choisit l'entier d assez grand de façon à avoir $x\sqrt{d} \leqslant d$. Les questions précédentes donnent :

$$\ln P(N_d > x\sqrt{d}) = \sum_{k=1}^{\lfloor x\sqrt{d}\rfloor -1} \ln\left(1-\frac{k}{d}\right) \underset{d\to+\infty}{\longrightarrow} -\frac{x^2}{2}$$

et par croissance de l'exponentielle $\lim\limits_{d\to+\infty} P\left(\frac{N_d}{\sqrt{d}} > x\right) = \mathrm{e}^{-x^2/2}$.

La suite $\left(\frac{N_d}{\sqrt{d}}\right)_d$ tend en loi vers une variable aléatoire de densité $f(x) = \mathbf{1}_{\mathbb{R}_+}(x)\times x\mathrm{e}^{-x^2/2}$.

4. a) On a $N_d(\Omega) = [\![2, d+1]\!]$ et

$$E(N_d) = \sum_{k=2}^{d+1} k(P(X>k-1)-P(X>k)) = \sum_{k=1}^{d}(k-1)(P(X>k)-\sum_{k=2}^{d+1} kP(X>k)$$

$$= \sum_{k=0}^{d} P(X>k).$$

b) L'intégrale proposée converge par croissance comparée entre une exponentielle et un polynôme. On a alors :

$$\int_0^{+\infty}\left(1+\frac{t}{d}\right)^d \mathrm{e}^{-t}dt = \sum_{k=0}^{d}\binom{d}{k}\frac{1}{d^k}\int_0^{+\infty} t^k\mathrm{e}^{-t}dt = \sum_{k=0}^{d}\binom{d}{k}\frac{k!}{d^k} = \sum_{k=0}^{d} P(N_d > k)$$

Exercice 3.12.

On note $x^- = \max(-x, 0)$ pour tout $x \in \mathbb{R}$.
De même si X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, on définit la variable aléatoire X^- par : $\forall\,\omega \in \Omega, X^-(\omega) = \max(-X(\omega), 0)$.
On pourra utiliser, sans démonstration, les résultats suivants :

$\star$ si $(Z_n)_{n\geq 1}$ est une suite de variables aléatoires convergeant en loi vers Z, alors, pour toute fonction continue et bornée f, on a $E(f(Z_n)) \underset{n\to\infty}{\longrightarrow} E(f(Z))$.

$\star$ si X et Y sont des variables aléatoires admettant des moments d'ordre 2, on a

$$E(XY) \leqslant \sqrt{E(X^2)E(Y^2)}$$

1. Soit X une variable aléatoire admettant une variance.

a) Montrer que $\forall\, a > 0, |X - \min(X, a)| \leqslant \mathbf{1}_{(X\geq a)}\times X$.

b) Montrer que : $\forall\, a > 0, E(|X - \min(X, a)|) \leqslant \sqrt{E(X^2)P(X \geqslant a)}$.

2. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, mutuellement indépendantes suivant toutes la loi de Poisson de paramètre 1. On pose pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} X_k \text{ et } Y_n = \frac{S_n - n}{\sqrt{n}}$$

a) Soit $n \in \mathbb{N}^*$. Quelle est la loi de S_n ?

b) En appliquant le théorème limite central, montrer que $(Y_n^-)_{n\geq 1}$ converge en loi vers une variable Y^-, où Y suit la loi $\mathcal{N}(0, 1)$.

c) Soient $n \in \mathbb{N}^*$ et $a > 0$. Calculer $E(Y_n^2)$ et en déduire que $P(Y_n^- \geqslant a) \leqslant \dfrac{1}{a^2}$.

d) Montrer que pour tout $a > 0$, $P(Y^- \geqslant a) \leqslant \dfrac{1}{a^2}$.

e) Déduire de la question 1. et du résultat admis en préambule, que $E(Y_n^-) \underset{n\to\infty}{\longrightarrow} E(Y^-)$.

3. Vérifier que, pour tout $n \in \mathbb{N}^*$, $E(Y_n^-) = \big(\frac{n}{2}\big)^n \times \dfrac{\sqrt{n}}{n!}$.

4. En déduire la formule de Stirling : $n! \underset{(\infty)}{\sim} \sqrt{2\pi n}\big(\frac{n}{\mathrm{e}}\big)^n$.

Solution :

1. a) Si $X(\omega) > a$, alors $|X(\omega) - \min(X(\omega), a)| = |X(\omega) - a| = X(\omega) - a \leqslant X(\omega)$ tandis que si $X(\omega) \leqslant a$, alors $|X(\omega) - \min(X(\omega), a)| = |X(\omega) - X(\omega)| = 0 \leqslant 0$.

On a donc bien toujours $|X(\omega) - \min(X(\omega), a)| \leqslant \mathbf{1}_{(X\geq a)}(\omega)\times X(\omega)$.

b) Ainsi, par croissance de l'espérance : $E(|X - \min(X, a)|) \leqslant E(\mathbf{1}_{(X\geq a)}X)$. Puis, par l'inégalité de Cauchy-Schwarz admise ici :

$$E(|X - \min(X, a)|) \leqslant \sqrt{P(X \geqslant a)E(X^2)}$$

($\mathbf{1}_{(X\geq a)}$ est une variable de Bernoulli, donc est égale à son carré et son espérance est son paramètre)

2. a) Comme les X_k sont indépendantes et de même loi $\mathcal{P}(1)$, la variable S_n suit la loi de Poisson de paramètre n.

b) Ainsi, $E(S_n) = n$ et $V(S_n) = n$. Par le théorème limite central, la suite $(Y_n)_{n\geq 1}$ converge en loi vers une variable Y suivant la loi $\mathcal{N}(0, 1)$.

On en déduit que, pour tout $x \in \mathbb{R}_+$,

$$P(Y_n^- \leqslant x) = P(Y_n \geqslant -x) \underset{n\to\infty}{\longrightarrow} P(Y \geqslant -x) = P(Y^- \leqslant x)$$

Pour $x \in \mathbb{R}_-$, on a :

$$P(Y_n^- \leqslant x) = 0 = P(Y^- \leqslant x)$$

On en déduit que $(Y_n^-)_{n\geq 1}$ converge en loi vers Y^-.

c) Soit $n \in \mathbb{N}^*$. On a $E(Y_n^2) = \dfrac{V(S_n)}{n} = 1$. Soit $a > 0$.

Comme $|Y_n| \geqslant Y_n^-$, on a $(Y_n^- \geqslant a) \subset (|Y_n| \geqslant a) = (Y_n^2 \geqslant a^2)$ et l'on déduit de l'inégalité de Markov que

$$P(Y_n^- \geqslant a) \leqslant P(Y_n^2 \geqslant a^2) \leqslant \frac{E(Y_n^2)}{a^2} = \frac{1}{a^2}$$

d) Il suffit de passer à la limite dans l'inégalité précédente en utilisant la convergence en loi de $(Y_n^-)_{n\geq 1}$ vers Y^-.

e) Soit $\varepsilon > 0$. On note $f : x \mapsto \min(x, 1/\varepsilon)$. Par les questions 1., 2.c. et 2.d., on a

$$E(|Y_n^- - f(Y_n^-)|) \leqslant \frac{1}{1/\varepsilon} = \varepsilon \text{ et } E(|Y^- - f(Y^-)|) \leqslant \frac{1}{1/\varepsilon} = \varepsilon$$

car $(Y_n^-)^2 \leqslant Y_n^2$ et donc $E\big((Y_n^-)^2\big) \leqslant E(Y_n^2)$ et, de même, $E((Y^-)^2) \leqslant E(Y^2) = 1$.

On déduit donc de l'inégalité triangulaire que :

$$\begin{aligned}|E(Y_n^-)-E(Y^-)| &\leqslant E(|Y_n^- - f(Y_n^-)|)+E(|f(Y_n^-)-f(Y^-)|)+E(|f(Y^-)-Y^-|)\\ &\leqslant 2\varepsilon + E(|f(Y_n^-)-f(Y^-)|)\end{aligned}$$

Comme f est continue et bornée sur $\mathbb{R}_+$, on a $E(f(Y_n^-)) \underset{n\to\infty}{\longrightarrow} E(f(Y^-))$ et il existe donc un rang à partir duquel $|E(Y_n^-)-E(Y^-)| \leqslant 3\varepsilon$. Ainsi :

$$E(Y_n^-) \underset{n\to\infty}{\longrightarrow} E(Y^-)$$

3. Soit $n \in \mathbb{N}^*$. Comme S_n suit la loi de Poisson de paramètre n, on a par télescopage :

$$\begin{aligned}E(Y_n^-) &= \sum_{k=0}^{n} \frac{n-k}{\sqrt{n}} P(S_n = k) = \sum_{k=0}^{n} \frac{n-k}{\sqrt{n}} \mathrm{e}^{-n} \frac{n^k}{k!} = \frac{\mathrm{e}^{-n}}{\sqrt{n}} \sum_{k=0}^{n} \Big(\frac{n^{k+1}}{k!} - \frac{kn^k}{k!}\Big)\\ &= \Big(\frac{n}{\mathrm{e}}\Big)^n \frac{\sqrt{n}}{n!}.\end{aligned}$$

4. Comme $E(Y^-) = \dfrac{1}{\sqrt{2\pi}}$, on déduit des questions précédentes que :

$$n! \underset{(n\to\infty)}{\sim} \sqrt{2\pi n}\Big(\frac{n}{\mathrm{e}}\Big)^n$$

Exercice 3.13.

Toutes les variables aléatoires de cet exercice sont définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

Soit p un réel de $]0,1[$, $q = 1-p$ et λ, μ deux réels strictement positifs. On pose :

$$f_{p,\lambda,\mu}(x) = \begin{cases} p\mu\,\mathrm{e}^{\mu x} & \text{si } x < 0 \\ q\lambda\,\mathrm{e}^{-\lambda x} & \text{si } x \geqslant 0 \end{cases}$$

1. Montrer que $f_{p,\lambda,\mu}$ est une densité de probabilité.

Soit X une variable aléatoire de densité $f_{p,\lambda,\mu}$.

2. Dans cette question, on suppose que $p = 1/2$ et $\lambda = \mu = 1$. On note $f = f_{1/2,1,1}$ et F la fonction de répartition de la variable aléatoire X ainsi associée. Pour tout entier $n \geqslant 1$, on pose

$$H_n(x) = \int_{-\infty}^{x} f(t)(1 + t.\mathrm{e}^{-n|t|})dt$$

a) Montrer que H_n est une fonction de répartition.

b) On note Y_n une variable aléatoire de fonction de répartition H_n. Montrer qu'il existe une constante C telle que pour tout réel x :

$$|H_n(x) - F(x)| \leqslant \frac{C}{n} F(x)$$

c) En déduire que la suite $(Y_n)_{n\geq 1}$ converge en loi et préciser la limite en loi de la suite $(Y_n)_n$.

3. On revient au cas général.

Déterminer (s'il existe) le plus petit réel s_0 tel que pour tout $t > s \geqslant s_0$, on a :

$$P_{(X>s)}(X > t) = P(X > t - s)$$

Solution :

1. La fonction $f_{p,\lambda,\mu}$ est positive et continue sauf en 0. De plus, la convergence des intégrales est claire, ce qui permet de mettre directement les bornes infinies :
$$\int_{-\infty}^{+\infty} f_{p,\lambda,\mu}(x)dx = p\int_{-\infty}^{0} \mu e^{\mu x}dx + q\int_{0}^{+\infty} \lambda e^{-\lambda x}dx = p+q=1$$
2. On a ici $f(t) = \dfrac{1}{2}\,e^{-|t|}$.

a) La fonction $\varphi : t \to e^{-|t|}(1+t.e^{-n|t|})$ est continue sur $\mathbb{R}$ et au voisinage de $-\infty$, on a $\varphi(t) \sim e^{-|t|}$ dont l'intégrale converge. De plus
• $\lim\limits_{x\to-\infty} H_n(x) = 0$ comme reste d'intégrale convergente ;
• la fonction H_n est positive de classe $\mathcal{C}^1$ et croissante car φ est continue et $\varphi(t) \geqslant 0$ (étude rapide de $t \mapsto 1 + t\,e^{nt}$ sur $\mathbb{R}_-$) ;
• on a $\lim\limits_{x\to+\infty} H_n(x) = 1$. En effet :
$$\frac{1}{2}\int_{-\infty}^{+\infty} e^{-|t|}(1+t.e^{-n|t|})dt = 1 + \frac{1}{2}\int_{-\infty}^{+\infty} t.e^{-(n+1)|t|}dt = 1+0=1$$
car la fonction $t \to te^{-(n+1)|t|}$ est impaire sur $\mathbb{R}$.
On conclut.

b) On écrit :
$$\begin{aligned}|H_n(x)-F(x)| &= \frac{1}{2}\left|\int_{-\infty}^{x} e^{-|t|}(1+t.e^{-n|t|})dt - \int_{-\infty}^{x} e^{-|t|}dt\right| \\ &\leqslant \frac{1}{2}\int_{-\infty}^{x} e^{-|t|} \times |t|e^{-n|t|}dt\end{aligned}$$
On étudie la fonction $g : t \to |t|e^{-n|t|}$ sur $\mathbb{R}$, donc $t \mapsto te^{-nt}$ sur $\mathbb{R}_+$.
L'étude des variations de g montre que g atteint son maximum en $\pm\dfrac{1}{n}$, maximum qui vaut $\dfrac{1}{ne}$. Ainsi
$$|H_n(x) - F(x)| \leqslant \frac{1}{2ne}F(x)$$
c) La suite $(Y_n)_n$ converge en loi vers la loi de X.

3. On remarque que pour $t > s$, $P_{(X>s)}(X>t) = \dfrac{P(X>t)}{P(X>s)}$.
• Pour $s \geqslant 0$ et $t > s$, le calcul donne $P(X>t) = \displaystyle\int_{t}^{+\infty} q\lambda\,e^{-\lambda x}dx = q.e^{-\lambda t}$.
Ainsi $P(X>s) = q.e^{-\lambda s}$, $P(X > t-s) = q.e^{-\lambda t}e^{\lambda s}$ et
$P(X>t) = P(X>t-s)P(X>s)$.
• Pour $s < 0$ et $t > s$, le calcul donne
$$P(X>s) = \int_{s}^{0} p\mu\,e^{\mu x}dx + \int_{0}^{+\infty} q\lambda\,e^{-\lambda x}dx = q - p.e^{\mu s}$$
$$P(X>t) = \begin{cases} q.e^{-\lambda t} & \text{si } t>0 \\ q-p.e^{\mu t} & \text{si } t<0 \end{cases}, \quad P(X>t-s) = q.e^{-\lambda t}e^{\lambda s}$$

On vérifie enfin que l'on n'a pas $P(X > t) = P(X > t - s)P(X > s)$.
Le réel s_0 demandé est donc $s_0 = 0$.

Exercice 3.14.

On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On suppose que pour chaque n, X_n suit la loi de Bernoulli de paramètre $p \in]0,1[$.

Pour $n \geqslant 2$, on désigne par $\mathcal{P}_n(2)$ l'ensemble des parties à deux éléments de l'ensemble $\{1, \dots, n\}$.

Si $I = \{i, j\} \in \mathcal{P}_n(2)$, on pose $Y_I = X_i X_j$.

Pour $n \geqslant 2$, on définit les variables aléatoires S_n et V_n par :

$$S_n = \sum_{k=1}^{n} X_k \text{ et } V_n = \sum_{I\in\mathcal{P}_n(2)} Y_I$$

1. a) Quelle est la loi suivie par S_n ? Donner son espérance et sa variance.

b) Que peut-on dire de la suite de variables aléatoires $\left(\dfrac{S_n}{n}\right)_{n\geq 2}$?

c) Calculer l'espérance de la variable aléatoire $S_n V_n$.

d) Justifier l'égalité $S_n^2 = S_n + 2V_n$.

e) En déduire l'espérance de la variable aléatoire S_n^3.

2. a) Soit $I \in \mathcal{P}_n(2)$. Quelle est la loi de Y_I ?

b) Soient $(I, J) \in \mathcal{P}_n(2)^2$. Déterminer selon les cas les valeurs que peut prendre la covariance des variables aléatoires Y_I et Y_J.

c) Déterminer l'espérance et la variance de V_n.

d) On pose $W_n = \dfrac{V_n}{n^2}$. Montrer que W_n est un estimateur asymptotiquement sans biais d'une quantité que l'on précisera.

e) Calculer le risque quadratique de W_n. Qu'en déduit-on ?

Solution :

1. a) S_n suit la loi binomiale de paramètres n et p.
On a donc $E(S_n) = np$ et $V(S_n) = np(1-p)$.

b) La loi faible des grands nombres nous dit que la suite de variables aléatoires $\left(\dfrac{S_n}{n}\right)_{n\geq 2}$ converge en probabilité vers la variable certaine égale à p.

c) On a $E(S_n V_n) = \displaystyle\sum_{I\in\mathcal{P}_n(2)} \Big[\sum_{k\notin I} E(X_k Y_I) + 2E(Y_I)\Big] = \dfrac{n(n-1)p^2}{2}[(n-2)p+2]$.

d) Il suffit de développer le carré $(X_1 + \cdots + X_n)^2$ en remarquant que $X_k^2 = X_k$.

e) Avec l'égalité précédente et c), on voit que

$E(S_n^3) = E(S_n^2) + 2E(S_n V_n) = E(S_n) + 2E(V_n) + 2E(S_n V_n)$

$$= np + n(n-1)p^2\left[(n-2)p+3\right]$$

(On pouvait aussi faire intervenir $V(S_n^2)$ après la première étape).

2. a) Soit $I = \{i,j\} \in \mathcal{P}_n(2)$. La variable Y_I prend les valeurs 0 et 1 et $P(Y_I = 1) = p^2$ puisque X_i et X_j sont indépendantes. Il s'agit donc d'une variable de Bernoulli de paramètre p^2.

b) Soient $(I,J) \in \mathcal{P}_n(2)^2$.

$\rightarrow$ Si $I \cap J = \emptyset$, le lemme des coalitions nous dit que Y_I et Y_J sont indépendantes, et par suite $\mathrm{Cov}(Y_I, Y_J) = 0$.

$\rightarrow$ Si $I \cap J = \{k\}$, alors on a $I = \{k,i\}$ et $I = \{k,j\}$ où i, j et k sont deux à deux distincts. En tenant compte de l'indépendance des X_l il vient :

$$\mathrm{Cov}(Y_I, Y_J) = E(X_kX_iX_j) - E(X_kX_i)E(X_kX_j) = p^3 - p^4$$

Lorsque $I = J$, on a $\mathrm{Cov}(Y_I, Y_J) = V(Y_I) = p^2(1-p^2)$.

c) On a déjà vu que $E(V_n) = \dfrac{n(n-1)p^2}{2}$. La bilinéarité de la covariance nous donne :

$$V(V_n) = \mathrm{Cov}(V_n, V_n) = \sum_{I \in \mathcal{P}_n(2)} V(Y_I) + \sum_{(I,J)\in\mathcal{P}_n(2)^2, I\neq J} \mathrm{Cov}(Y_I, Y_J)$$

En tenant compte de la question précédente, on obtient alors

$$\begin{aligned} V(V_n) &= \frac{n(n-1)}{2}p^2(1-p^2) + \sum_{(I,J)\in\mathcal{P}_n(2)^2, \mathrm{card}(I\cap J)=1} \mathrm{Cov}(Y_I, Y_J) \\ &= \frac{n(n-1)}{2}p^2(1-p^2) + \frac{n(n-1)(n-2)}{2}(p^3-p^4) \\ &= \frac{n(n-1)p^2(1-p)\left[1+(n-1)p\right]}{2} \end{aligned}$$

d) On a $E(W_n) = \dfrac{n(n-1)p^2}{2n^2} \underset{n\to\infty}{\longrightarrow} \dfrac{p^2}{2}$. Par suite, W_n est un estimateur asymptotiquement sans biais de $\dfrac{p^2}{2}$.

e) Il vient, en notant b le biais, qui vaut donc $-\dfrac{p^2}{2n}$:

$$r(W_n) = b(W_n)^2 + V(W_n) = b(W_n)^2 + \frac{1}{n^4}V(V_n)$$

$$r(W_n) = \frac{1}{n^2}\times\frac{p^4}{4} + \frac{1}{n^4}\times\frac{n(n-1)p^2(1-p)\left[1+(n-1)p\right]}{2} \underset{n\to\infty}{\longrightarrow} 0$$

Il en résulte que W_n est un estimateur asymptotiquement sans biais et convergent de $\dfrac{p^2}{2}$.

Exercice 3.15.

Dans cet exercice, toutes les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$. Soit $(X_n)_n, (Y_n)_n$ deux suites de variables aléatoires et X, Y deux variables aléatoires.

On suppose que $(X_n)_n$ converge en probabilité vers X et que $(Y_n)_n$ converge en probabilité vers Y.

1. Vérifier que $X_nY_n - XY = (X_n - X)Y + (Y_n - Y)X + (X_n - X)(Y_n - Y)$.

Soit $\varepsilon > 0$. On pose :

$A = (|X_nY_n - XY| > \varepsilon)$,
$A_1 = (|X_n - X||Y| > \varepsilon/3), A_2 = (|Y_n - Y||X| > \varepsilon/3)$,
$A_3 = (|X_n - X||Y_n - Y| > \varepsilon/3)$.

2. Montrer que $P(A) \leqslant P(A_1) + P(A_2) + P(A_3)$.

3. Soit $t > 0$.
a) Montrer que $P(A_1) \leqslant P\big(|X_n - X| > \frac{\varepsilon}{3t}\big) + P(|Y| > t)$.

b) Soit $\delta > 0$. Montrer qu'il existe un réel $A > 0$ tel que pour tout $t > A$, $P(|Y| > t) < \delta/2$.

c) Montrer qu'il existe un entier N_0 tel que pour tout $n \geqslant N_0$, $P\big(|X_n - X| > \frac{\varepsilon}{3t}\big) < \frac{\delta}{2}$.

d) En déduire que $\lim\limits_{n\to+\infty} P(A_1) = 0$.

On montrerait de même et on admet ici que $\lim\limits_{n\to+\infty} P(A_2) = 0$.

4. Montrer que $P(A_3) \leqslant P\big(|X_n - X| > \sqrt{\frac{\varepsilon}{3}}\big) + P\big(|Y_n - Y| > \sqrt{\frac{\varepsilon}{3}}\big)$.
En déduire que $\lim\limits_{n\to+\infty} P(A_3) = 0$.

5. Montrer que la suite (X_nY_n) converge en probabilité vers XY.

Solution :

1. Il suffit de développer l'expression proposée.

2. On passe aux complémentaires :
$\overline{A_1}\cap\overline{A_2}\cap\overline{A_3} = [|X_n-X||Y| \leq \varepsilon/3]\cap[|Y_n-Y||X| \leqslant \varepsilon/3]\cap[|X_n-X||Y_n-Y| \leq \varepsilon/3]$
En passant alors aux valeurs absolues dans le résultat de la question 1. l'inégalité triangulaire donne :

$$\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \subseteq \overline{A}$$

Donc $A \subseteq [A_1 \cup A_2 \cup A_3]$ et $P(A) \leqslant P(A_1 \cup A_2 \cup A_3) \leqslant P(A_1) + P(A_2) + P(A_3)$.

3. a) La famille $(B, \overline{B})$ avec $B = [|Y| \leqslant t]$ et $\overline{B} = [|Y] > t]$ forme un système complet d'événements. Donc

$$\begin{aligned} P(A_1) &= P(A_1 \cap B) + P(A_1 \cap \overline{B}) \\ &\leqslant P([|X_n - X| > \varepsilon/(3|Y|)] \cap [|Y| \leqslant t]) + P(|Y| > t) \\ &\leqslant P([|X_n - X| > \varepsilon/(3t)]) + P(|Y| > t) \end{aligned}$$

b) On notant F la fonction de répartition de $|Y|$, on a $P(|Y| > t) = 1 - F(t) \to 0$ quand t tend vers $+\infty$. Ainsi

$$\forall\, \delta > 0, \exists A \text{ tel que } t \geqslant A \implies P(|Y| > t) < \delta$$

c) Comme la suite (X_n) converge en probabilité vers X, il existe N tel que si $n \geqslant N$:

$$P([|X_n - X| > \varepsilon/(3t)]) < \delta$$

d) Par les questions précédentes, il existe N tel que si $n \geqslant N$, $P(A_1) \leqslant 2\delta$. Ainsi :

$$\lim_{n\to+\infty} P(A_1) = 0$$

4. Si $|X_n - X| \leqslant \sqrt{\varepsilon/3}$ et $|Y_n - Y| \leqslant \sqrt{\varepsilon/3}$, alors $|X_n - X||Y_n - Y| \leqslant \varepsilon/3$. Donc

$$P(A_3) \leqslant P(|X_n - X| > \sqrt{\varepsilon/3}) + P(|Y_n - Y| > \sqrt{\varepsilon/3})$$

Comme (X_n) converge en probabilité vers X et (Y_n) converge en probabilité vers Y, on a $\lim\limits_{n\to+\infty} P(A_3) = 0$.

5. En regroupant les questions précédentes, on a $\lim\limits_{n\to+\infty} P(A) = 0$ et $(X_nY_n)_n$ converge en probabilité vers XY.

Exercice 3.16.

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, mutuellement indépendantes et de même loi uniforme sur $\{-1\ , +1\}$.

Pour tout entier naturel n non nul on pose : $S_n = X_1 + \cdots + X_n$ et $u_n = E(|S_n|)$, où $E(|S_n|)$ désigne l'espérance de la valeur absolue de S_n.

On admettra le résultat suivant : $n! \underset{+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$ (formule de Stirling).

1. Soit $n \in \mathbb{N}^*$. Calculer $P(S_{2n-1} = 0)$ et $P(S_{2n} = 0)$.

2. Déterminer un équivalent de $P(S_{2n} = 0)$ quand n tend vers $+\infty$.

3. a) Pour $k \in \mathbb{Z}$, simplifier l'expression $|k - 1| + |k + 1|$.
   b) Démontrer que pour tout entier naturel n non nul, $u_{n+1} = u_n + P(S_n = 0)$.

4. Pour tout entier naturel non nul n, on pose $t_n = \sum\limits_{k=1}^{n} \dfrac{1}{\sqrt{k}}$. Prouver que $t_n \underset{(+\infty)}{\sim} 2\sqrt{n}$.

5. Soit deux suites $\big(\alpha_n\big)_{n\geq 1}$ et $\big(\beta_n\big)_{n\geq 1}$ de réels strictement positifs telles que $\alpha_n \underset{(+\infty)}{\sim} \beta_n$.

Montrer que si la série de terme général β_n diverge, alors : $\sum\limits_{k=1}^{n} \alpha_k \underset{(+\infty)}{\sim} \sum\limits_{k=1}^{n} \beta_k$.

6. Montrer que $E(|S_n|) \underset{(+\infty)}{\sim} \sqrt{\dfrac{2n}{\pi}}$.

Solution :

1. Réaliser $(S_m = 0)$ c'est obtenir en cours de route autant de fois la valeur 1 que la valeur -1, donc : $P(S_{2n-1} = 0) = 0$ et en choisissant les n endroits parmi $2n$ où on obtient la valeur 1 :

$$P(S_{2n} = 0) = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} = \frac{(2n)!}{n!n!2^{2n}}$$

2. La formule de Stirling donne alors après simplifications : $P(S_{2n} = 0) \sim \dfrac{1}{\sqrt{\pi n}}$.

3. a) $|x-1| + |x+1| = \begin{cases} 2x & \text{si } x \geqslant 1 \\ -2x & \text{si } x \leq -1 \\ 2 & \text{si } x = 0 \end{cases}$

b) Par indépendance mutuelle des variables X_i et le lemme des coalitions, on déduit que pour tout entier n non nul, S_n et X_{n+1} sont indépendantes.
D'autre part, S_n prend des valeurs entières relatives dans $[\![-n, n]\!]$, la variable $|S_{n+1}|$ est bornée et admet une espérance qui s'exprime grâce à la formule du transfert par :

$$\begin{aligned} E\big(|S_{n+1}|\big) &= E\big(|S_n + X_{n+1}|\big) = \sum_{-n \leq k \leq n} \sum_{i \in \{-1,1\}} |k+i| P(S_n = k).P(X_{n+1} = i) \\ &= \frac{1}{2} \sum_{-n \leq k \leq n} \big(|k-1| + |k+1|\big) P(S_n = k) \\ &= \frac{1}{2} \sum_{-n \leq k \leq -1} (-2k) P(S_n = k) + P(S_n = 0) + \frac{1}{2} \sum_{1 \leq k \leq n} (2k) P(S_n = k) \\ &= P(S_n = 0) + \sum_{-n \leq k \leq n} |k| P(S_n = k) = P(S_n = 0) + E(|S_n|) \end{aligned}$$

On en déduit que pour tout entier n non nul :

$$u_{n+1} = u_n + P(S_n = 0)$$

4. Pour tout entier n, $n \geqslant 2$ et tout entier $k \in [\![1, n-1]\!]$: $\dfrac{1}{\sqrt{k+1}} \leqslant \displaystyle\int_k^{k+1} \frac{1}{\sqrt{x}} dx \leqslant \dfrac{1}{\sqrt{k}}$. Puis par addition : $\displaystyle\sum_{k=1}^{n-1} \frac{1}{\sqrt{k+1}} \leqslant \int_1^n \frac{1}{\sqrt{x}} dx \leqslant \sum_{k=1}^{n-1} \frac{1}{\sqrt{k}}$. D'où

$$t_n - 1 \leqslant 2\sqrt{n} - 2 \leqslant t_n - \frac{1}{\sqrt{n}} \leqslant t_n,$$

donc $2\sqrt{n} - 2 \leqslant t_n \leqslant 2\sqrt{n} - 1$. On déduit par encadrement et limite que $t_n \underset{+\infty}{\sim} 2\sqrt{n}$.

5. Soit $\varepsilon > 0$ et $n_0 \in \mathbb{N}$ tel que $\forall\, n \geqslant n_0, \alpha_n(1-\varepsilon) \leq \beta_n \leq \alpha_n(1+\varepsilon)$, alors pour $n \geqslant n_0$

$$\sum_{k=1}^{n_0-1} \alpha_k + (1-\varepsilon) \sum_{k=n_0}^{n} \alpha_k \leq \sum_{k=1}^{n} \beta_k \leq \sum_{k=1}^{n_0-1} \alpha_k + (1+\varepsilon) \sum_{k=n_0}^{n} \alpha_k$$

Soit en rectifiant les débuts des sommations :

$$K_1 + (1-\varepsilon) \sum_{k=1}^{n} \alpha_k \leq \sum_{k=1}^{n} \beta_k \leqslant K + (1+\varepsilon) \sum_{k=1}^{n} \alpha_k$$

où K_1 et K ne dépendent pas de n. Il suffit alors de diviser par $\sum\limits_{k=1}^{n} \alpha_k > 0$ et pour n assez grand le quotient est compris entre $1 - 2\varepsilon$ et $1 + 2\varepsilon$. Comme ε est quelconque on a le résultat voulu.

6. Avec $u_1 = 0$, il vient : $u_{2n+1} = \sum\limits_{k=1}^{2n} P(S_k = 0) = \sum\limits_{k=1}^{n} P(S_{2k} = 0)$ et $u_{2n+2} = u_{2n+1}$.
Or $P(S_{2n} = 0) \underset{+\infty}{\sim} \dfrac{1}{\sqrt{\pi n}}$ donc la série de terme général $P(S_{2k} = 0)$ est divergente, à termes strictement positifs. On déduit des questions 4. et 5. que :

$$\sum_{k=1}^{n} P(S_{2k} = 0) \underset{+\infty}{\sim} \sum_{k=1}^{n} \frac{1}{\sqrt{\pi k}} \underset{+\infty}{\sim} \frac{2\sqrt{n}}{\sqrt{\pi}} \text{ soit}$$

$$u_{2n+1} \underset{+\infty}{\sim} \frac{2\sqrt{n}}{\sqrt{\pi}}$$

De même $u_{2(n+1)} = u_{2n+1} \underset{+\infty}{\sim} \dfrac{2\sqrt{n}}{\sqrt{\pi}} \underset{+\infty}{\sim} \dfrac{2\sqrt{n+1}}{\sqrt{\pi}}$, ce qui conduit à $u_{2n} \underset{+\infty}{\sim} \dfrac{2\sqrt{n}}{\sqrt{\pi}}$.
Ce que l'on peut regrouper en :

$$E(|S_n|) \underset{+\infty}{\sim} \sqrt{\frac{2n}{\pi}}$$

Exercice 3.17.

Soit X_1, X_2 et Z trois variables aléatoires mutuellement indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ telles que X_1 suit la loi exponentielle de paramètre $\lambda_1 > 0$, X_2 suit la loi exponentielle de paramètre $\lambda_2 > 0$ et Z suit la loi de Bernoulli de paramètre p avec $p \in]0, 1[$.
On pose :

$$X = ZX_1 + (1 - Z)X_2$$

1. Montrer que X est une variable aléatoire.

2. Exprimer la fonction de répartition de X notée F_X à l'aide des fonctions de répartition de X_1 et de X_2 soit F_{X_1} et F_{X_2} et du paramètre p.

3. En déduire que X est une variable aléatoire à densité. Exprimer une densité de X en fonction de λ_1, λ_2 et p.

4. Exprimer l'espérance et la variance de X en fonction de λ_1, λ_2 et p.

5. On suppose dans cette question que $\lambda_1 = 1$ et que $\lambda_2 = 1/2$. On désire estimer le paramètre p supposé inconnu.

On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes de même loi que X. On pose :

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Proposer un estimateur sans biais et convergent de p.

Solution :

1. X est une application de Ω dans $\mathbb{R}$. De plus soit x un réel, on a

$$[X \leqslant x] = ([Z = 1] \cap [X_1 \leqslant x]) \cup (([Z = 0] \cap [X_2 \leqslant x])$$

Or X_1, X_2 et Z sont des variables aléatoires donc $[Z = 1]$, $[Z = 0]$, $[X_1 \leqslant x]$ et $[X_2 \leqslant]$ sont des événements ainsi par union et intersection, $[X \leqslant]$ est un événement. Ceci prouve que X est une variable aléatoire.

2. On a :
$\forall x \in \mathbb{R},\ \ F_X(x) = P(X \leqslant x) = P(([Z = 1] \cap [X_1 \leqslant x]) \cup (([Z = 0] \cap [X_2 \leqslant x]))$,
il s'agit d'une union d'événements disjoints et par indépendance des variables aléatoires, on obtient :

$$\forall x \in \mathbb{R},\ \ F_X(x) = pF_{X_1}(x) + (1 - p)F_{X_2}(x)$$

3. Comme X_1 et X_2 sont des variables à densité, F_{X_1} et F_{X_2} sont continues sur $\mathbb{R}$ et de classe C^1 sauf éventuellement en un nombre fini de points. La combinaison linéaire de ces deux fonctions conserve ces propriétés donc F_X est continue sur $\mathbb{R}$ et de classe C^1 sauf éventuellement en un nombre fini de points. Par conséquent X est une variable à densité.

Par dérivation une densité de X est f_X définie par :

$$\forall x \in \mathbb{R}, f_X(x) = (p\lambda_1 \mathrm{e}^{-\lambda_1 x} + (1 - p)\lambda_2 \mathrm{e}^{-\lambda_2 x})\mathbf{1}_{\mathbb{R}^+}$$

4. On reconnaît les intégrales qui apparaissent dans les calculs à effectuer et :

$$E(X) = pE(X_1) + (1 - p)E(X_2) = \frac{p}{\lambda_1} + \frac{1 - p}{\lambda_2}$$

puis :

$$\begin{aligned} V(X) &= pE(X_1^2) + (1 - p)E(X_2^2) - (\frac{p}{\lambda_1} + \frac{1 - p}{\lambda_2})^2 \\ &= p\frac{2}{\lambda_1^2} + (1 - p)\frac{2}{\lambda_2^2} - (\frac{p}{\lambda_1} + \frac{1 - p}{\lambda_2})^2 \\ &= \frac{p(2 - p)}{\lambda_1^2} + \frac{(1 - p)(1 + p)}{\lambda_2^2} - \frac{2p(1 - p)}{\lambda_1\lambda_2} \end{aligned}$$

5. On suppose dans cette question que $\lambda_1 = 1$ et que $\lambda_2 = 1/2$. On a donc $E(X) = 2 - p$. On en déduit que $2 - \overline{X_n}$ est un estimateur sans biais de p. De plus $V(X) = -p^2 - 2p + 4$. On en déduit que la variance de $2 - \overline{X_n}$ tend vers 0 quand n tend vers $+\infty$.

Donc $2 - \overline{X_n}$ est un estimateur sans biais et convergent de p.

Exercice 3.18.

Toutes les variables aléatoires de cet exercice sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

Un mobile se déplace par sauts d'une unité sur les points d'un axe $(O, \overrightarrow{i})$ dont les abscisses sont des entiers naturels. Un saut vers la droite (pour lequel l'abscisse augmente d'une unité) se fait avec la probabilité $p \in]0, 1[\setminus \{1/2\}$, tandis qu'un saut vers la gauche se fait avec la probabilité $q = 1-p$. Les différents sauts effectués sont indépendants les uns des autres.

Soit a un entier supérieur ou égal à 3 fixé et n un entier tel que $0 \leqslant n \leqslant a$. Le mobile démarre du point d'abscisse n et son voyage se termine lorsqu'il se trouve en l'origine ou au point A d'abscisse a (si au départ il est en un de ces deux points, le voyage s'achève donc immédiatement).

On note D_n la durée de ce voyage (c'est-à-dire le nombre de sauts effectués jusqu'à l'arrêt) et on admet que D_n possède une espérance que l'on note $E(D_n)$.

On note S l'événement «le premier saut est un saut vers la droite».

1. Ecrire un script `Scilab` permettant de simuler la variable D_n, pour une valeur de n entrée par l'utilisateur.

2. Montrer que lorsque $0 < n < a$, l'ensemble $D_n(\Omega)$ n'est pas borné.

3. a) Montrer que : $\forall\, k \in \mathbb{N}^*, P_S(D_n = k) = P(D_{n+1} = k-1)$.

b) En déduire que pour $n \in \rrbracket 0, a-1 \rrbracket$, $E(D_n/S) = E(D_{n+1}) + 1$.

c) Exprimer de même $E(D_n/\overline{S})$, où $\overline{S}$ est l'événement contraire de S.

4. a) Que valent $E(D_0)$ et $E(D_a)$?

b) Pour $n \in \llbracket 0, a \rrbracket$, on pose $u_n = E(D_n)$. Montrer qu'il existe des coefficients réels α, β, γ, que l'on précisera, tels que la liste $(u_n)_{n\in\llbracket 0,a \rrbracket}$ vérifie la relation de récurrence :

$$\forall\, n \in \llbracket 1, a-1 \rrbracket, u_{n+1} = \alpha u_n + \beta u_{n-1} + \gamma \quad (*)$$

c) Montrer qu'il existe une suite de la forme $n \mapsto \delta n$ vérifiant la relation $(*)$.

d) En déduire la valeur de $u_n = E(D_n)$ en fonction de a, n, p et q.

Solution :

1.
```
n = input('la valeur de n est :')
x = n ; D = 0
while x>0 & x<a
  if rand()<p then x = x+1 else x = x-1
  end
D=D+1
end
disp(D)
```

2. Il y a au moins un déplacement et comme $a \geqslant 3$, il se peut que le voyage ne s'achève pas au premier déplacement. On peut alors revenir au point de départ et recommencer indéfiniment. Ne pas chercher à donner exactement $D_n(\Omega)$, car il faut déjà s'approcher du but et en plus il peut y avoir une obligation de parité ...

3. a) Sachant que S est réalisé, le mobile va en $n+1$ et pour réaliser $(D_n = k)$ il lui reste $k-1$ sauts à faire depuis ce point. Ainsi $\forall\, k \in \mathbb{N}^*, P_S(D_n = k) = P(D_{n+1} = k-1)$, le résultat restant vrai même pour $k = 1$.

b) Alors, la convergence des sommes écrites étant admise par l'énoncé et le fait de commencer au rang 0 ou au rang 1 ne changeant rien :
$E(D_n/S) = \sum k\times P_S(D_n = k) = \sum k\times P(D_{n+1} = k-1) = \sum kP(D_{n+1}+1 = k)$, soit :

$$E(D_n/S) = E(D_{n+1}+1) = E(D_{n+1}) + 1$$

c) Mutatis mutandis :

$$E(D_n/\overline{S}) = E(D_{n-1}+1) = E(D_{n-1}) + 1$$

4. a) $E(D_0) = E(D_n) = 0$ (on ne se met pas en marche !)

b) Par la formule de l'espérance totale :

$$\begin{aligned} E(D_n) &= P(S)E(D_n/S) + P(\overline{S})E(D_n/\overline{S}) = pE(D_n/S) + qE(D_n/\overline{S}) \\ &= p(E(D_{n+1})+1) + q(E(D_{n-1})+1) \end{aligned}$$

Soit pour tout n de $[\![1, a-1]\!]$:

$$E(D_n) = pE(D_{n+1}) + qE(D_{n-1}) + 1$$

On note $u_n = E(D_n)$ et on met les termes aux places demandées ...

c) Avec $v_n = \delta n$, cette séquence vérifie $(*)$ si :

$$\forall\, n, \delta n = \delta(p(n+1) + q(n-1)) + 1 = \delta n + \delta(p-q) + 1$$

et ceci est vrai pour tout n si on choisit $\delta = \dfrac{1}{q-p}$.

d) On a : $\begin{cases} u_n = pu_{n+1} + qu_{n-1} + 1 \\ v_n = pv_{n+1} + qv_{n-1} + 1 \end{cases}$ (pour le bon choix de δ), donc par différence la séquence (w_n) définie par $w_n = u_n - v_n = u_n - \dfrac{n}{q-p}$ vérifie la récurrence linéaire sur deux rangs :

$$w_n = pw_{n+1} + qw_{n-1}$$

L'équation caractéristique est $px^2 - x + q = 0$ de racine évidente 1 et donc l'autre vaut $\dfrac{q}{p}$ et w_n est de la forme $w : n \mapsto \lambda_1 + \lambda_2\big(\frac{q}{p}\big)^n$, puis :

$$u_n = w_n + v_n = \lambda_1 + \lambda_2\Big(\frac{q}{p}\Big)^n + \frac{n}{q-p}$$

avec $u_0 = 0$ et $u_a = 0$, il vient $\lambda_1 + \lambda_2 = 0$ et $\lambda_1 + \lambda_2\big(\frac{q}{p}\big)^a + \dfrac{a}{q-p} = 0$, soit tous calculs faits :

$$u_n = E(D_n) = \frac{n}{q-p} + \Big(1 - \Big(\frac{q}{p}\Big)^n\Big)\times\frac{a}{(q-p)(1-(\frac{q}{p})^a)}$$

Exercice 3.19.

Préliminaire :

Calculer pour a et b deux réels tels que $0 \leqslant a < b$ l'intégrale $I(x) = \displaystyle\int_a^b \min(x,t)\,dt$ selon les valeurs du réel x.

Dans la suite de l'exercice, on se donne une variable aléatoire réelle X définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et on pose :

$$\forall\,\omega \in \Omega, Y(\omega) = \int_a^b \min(X(\omega), t)\,dt$$

1. Exprimer Y en fonction de X, $\mathbf{1}_{\{X<a\}}$, $\mathbf{1}_{\{a\leq X\leq b\}}$, et $\mathbf{1}_{\{X>b\}}$ d'une part et en fonction de $\mathbf{1}_{\{X\leq a\}}$, $\mathbf{1}_{\{a<X<b\}}$, et $\mathbf{1}_{\{X\geq b\}}$ d'autre part.

2. En déduire que Y est aussi une variable aléatoire.

3. On suppose ici que X est positive, discrète et admet une espérance.
On note $Z = X.\mathbf{1}_{\{X\leq a\}}$, $U = X^2.\mathbf{1}_{\{a<X<b\}}$ et $V = X.\mathbf{1}_{\{a<X<b\}}$.

Montrer que Y admet une espérance $E(Y)$ et l'exprimer en fonction de a, b, X, $E(Z)$, $E(U)$, $E(V)$.

4. On suppose ici que $a = 0$, $b = 1$, et que X suit la loi géométrique de paramètre $p \in\,]0, 1[$. Préciser Y et son espérance.

5. On suppose ici que $a = 0$, $b = 1$, et que X est à valeurs dans $\mathbb{N}$.
Préciser la loi de Y et montrer que Y possède une espérance que l'on calculera en fonction de X.

6. On suppose ici que X suit la loi uniforme sur $[a, b]$. Exprimer Y en fonction de X et préciser son espérance, en fonction de a et b.

Solution :

Préliminaire :

$$I(x) = \int_a^b \min(x,t)\,dt = \begin{cases} (b-a)x & \text{si } x \leqslant a \\ \dfrac{x^2-a^2}{2} + (b-x)x & \text{si } a \leqslant x \leqslant b \\ \dfrac{b^2-a^2}{2} & \text{si } x \geqslant b \end{cases}$$

1. Par définition des variables indicatrices qui valent 1 quand elles jouent un rôle et 0 sinon :

$$Y = (b-a)X\times\mathbf{1}_{(X\leq a)} + \big(\frac{X^2-a^2}{2} + (b-X)X\big)\times\mathbf{1}_{(a<X<b)} + \frac{b^2-a^2}{2}\times\mathbf{1}_{(X\geq b)}$$

ou, comme les formules du préliminaire se recollent aux bords :

$$Y = (b-a)X\times\mathbf{1}_{(X<a)} + \big(\frac{X^2-a^2}{2} + (b-X)X\big)\times\mathbf{1}_{(a\leq X\leq b)} + \frac{b^2-a^2}{2}\times\mathbf{1}_{(X>b)}$$

2. La somme et le produit de deux variables aléatoires sont des variables aléatoires, les constantes et les fonctions indicatrices d'événements comme $(X \leqslant a)$, $(a < X < b)$ et $(X \geqslant b)$ sont aussi des variables aléatoires.
Bref, comme somme de trois variables aléatoires, Y en est aussi une.

3. La variable $Z = X\times\mathbf{1}_{(X\leq a)}$ est discrète, comprise entre 0 et a, elle admet donc une espérance.

$U = X^2\times\mathbf{1}_{(a<X<b)}$ est discrète, bornée entre 0 et b^2 et donc admet aussi une espérance.

$V = X\times\mathbf{1}_{(a<X<b)}$ est discrète, comprise entre 0 et b, admet aussi une espérance.

Enfin, $\dfrac{b^2-a^2}{2}\times\mathbf{1}_{(X\geq b)}$ est discrète, positive et inférieure ou égale à $\dfrac{b^2}{2}$, donc admet aussi une espérance.

Par combinaison linéaire, Y admet bien une espérance qui vaut :

$$E(Y)=(b-a)E(Z)+bE(V)-\frac{1}{2}E(U)-\frac{a^2}{2}P(a<X<b)+\frac{b^2-a^2}{2}P(X\geqslant b)$$

$$E(Y)=(b-a)E(Z)+bE(V)-\frac{1}{2}E(U)+\frac{b^2}{2}P(X\geqslant b)-\frac{a^2}{2}P(X>a)$$

4. En appliquant les résultats précédents, on obtient :

$$Y=X\times\mathbf{1}_{(X\leq 0)}+\big(\frac{X^2}{2}+(1-X)X\big)\times\mathbf{1}_{(0<X<1)}+\frac{1}{2}\times\mathbf{1}_{(X\geq 1)}=\frac{1}{2}$$

et son espérance vaut banalement $\dfrac{1}{2}$.

5. $Y=X\times\mathbf{1}_{(X=0)}+\frac{1}{2}(\mathbf{1}-\mathbf{1}_{(X=0)})=\frac{1}{2}\big(\mathbf{1}-\mathbf{1}_{(X=0)}\big)$ et son espérance est

$$E(Y)=\frac{1}{2}(1-P(X=0))$$

Note : on retrouve le résultat de la question précédente.

6. $Y=\big(\dfrac{X^2-a^2}{2}+(b-X)X\big)\times\mathbf{1}_{(a\leq X\leq b)}=\big(\dfrac{X^2-a^2}{2}+(b-X)X\big)=bX-\dfrac{X^2+a^2}{2}$

Son espérance est

$$E(Y)=bE(X)-\frac{E(X^2)+a^2}{2}=b\frac{a+b}{2}-\frac{V(X)+(E(X))^2+a^2}{2}$$

Soit, après calculs :

$$E(Y)=\frac{a^2+b^2+ab}{3}$$

Exercice 3.20.

Soit n un entier naturel tel que $n\geqslant 2$. On dispose d'un paquet de n cartes $C_1,C_2,\dots,C_n$ que l'on distribue intégralement, les unes après les autres entre n joueurs $J_1,J_2,\dots,J_n$ selon le protocole suivant :
$\rightarrow$ la première carte C_1 est donnée à J_1 ;
$\rightarrow$ la deuxième carte C_2 est distribuée de façon équiprobable entre J_1 et J_2 ;
$\rightarrow$ la troisième carte C_3 est distribuée de façon équiprobable entre J_1,J_2 et J_3 ;
$\rightarrow$ et ainsi de suite, jusqu'à la dernière carte C_n qui est donc distribuée de façon équiprobable entre les joueurs $J_1,J_2,\dots,J_n$.

On suppose l'expérience modélisée sur un espace probabilisé $(\Omega,\mathcal{B},P)$.

On note X_n la variable aléatoire égale au nombre de joueurs qui n'ont reçu aucune carte à la fin de la distribution.

1. Déterminer $X_n(\Omega)$ et calculer $P(X_n=0)$ et $P(X_n=n-1)$.

2. Pour tout i de $[\![1,n]\!]$, on note B_i la variable aléatoire qui vaut 1 si J_i n'a reçu aucune carte à la fin de la distribution et qui vaut 0 sinon.

Déterminer la loi de B_i. Exprimer la variable aléatoire X_n en fonction des variables aléatoires B_i et en déduire l'espérance de X_n.

3. En faisant le moins de calculs possibles, donner la loi de X_4.

4. a) Montrer que pour i et j dans $[\![1,n]\!]$ tels que $i<j$, on a :

$$P((B_i=1)\cap(B_j=1))=\frac{(i-1)(j-2)}{n(n-1)}$$

En déduire la covariance des variables aléatoires B_i et B_j.

b) Montrer que $V(X_n)=\dfrac{n+1}{12}$.

Solution :

1. $X_n[\Omega]=[\![0,n-1]\!]$, car on peut donner une carte à chacun ou toutes les cartes à J_1, ou toute situation intermédiaire.

$P(X_n=0)=P(X_n=n-1)=\dfrac{1}{n!}$ (à chaque fois une seule façon de faire).

2. Pour $i\geqslant 2$, réaliser $(B_i=1)$ c'est donner les cartes $C_i,C_{i+1},\dots,C_n$ à d'autres (pour les $i-1$ premières cartes il n'est pas dans la course !), en suivant le nombre de joueurs en lice à chaque fois et par indépendance :

$$P(B_i=1)=\frac{i-1}{i}\times\frac{i}{i+1}\times\cdots\times\frac{n-1}{n}=\frac{i-1}{n}$$

valable aussi pour $i=1$.

B_i étant une variable de Bernoulli, on a $E(B_i)=\dfrac{i-1}{n}$ et comme $X_n=\sum\limits_{i=1}^{n}B_i$, il vient $E(X_n)=\dfrac{n-1}{2}$ (vérification élémentaire pour $n=1$ et $n=2$!)

3. X_4 prend les valeurs 0 et 3 avec à chaque fois la probabilité $\dfrac{1}{24}$.

Donc $P(X_4=1)+P(X_4=2)=\dfrac{22}{24}$ et comme

$P(X=1)+2P(X=2)+3P(X=3)=\dfrac{3}{2}$, on en tire en résolvant ce système affine de deux équations : $P(X=1)=P(X=2)=\dfrac{11}{24}$.

4. a) Réaliser $(B_i=1)\cap(B_j=1)$, c'est donner les cartes $C_i,C_{i+1}\dots,C_{j-1}$ à quelqu'un d'autre que J_i (J_j n'est pas dans la course) et donner les cartes $C_j,\dots,C_n$ ni à J_i ni à J_j. En procédant comme en 2., on a donc :

$$P((B_i=1)\cap(B_j=1))=\Big(\prod_{k=1}^{j-1}\frac{k-1}{k}\Big)\Big(\prod_{k=j}^{n}\frac{k-2}{k}\Big)=\frac{i-1}{j-1}\times\frac{(j-1)(j-2)}{n(n-1)}$$

Ainsi $P((B_i=1)\cap(B_j=1))=\dfrac{(i-1)(j-2)}{n(n-1)}$.

$\mathrm{Cov}(B_i,B_j)=E(B_iB_j)-E(B_i)E(B_j)$ et comme il n'y a ici que des variables aléatoires de Bernoulli :

$$\mathrm{Cov}(B_i,B_j)=\frac{(i-1)(j-2)}{n(n-1)}-\frac{(i-1)(j-1)}{n^2}=-\frac{(i-1)(n-j+1)}{n^2(n-1)}$$

b) On a

$$V(B_i)=\frac{i-1}{n}\Big(1-\frac{i-1}{n}\Big)=\frac{(i-1)(n-i+1)}{n^2}$$

donc en développant :

$$V(X_n)=\sum_i \frac{(i-1)(n-i+1)}{n^2}-2\sum_{i<j}\frac{(i-1)(n-j+1)}{n^2(n-1)}$$

La première somme est $\dfrac{1}{n^2}(n\sum\limits_{k=0}^{n-1}k-\sum\limits_{k=0}^{n-1}k^2)$ et vaut $\dfrac{n-1}{2}-\dfrac{(n-1)(2n-1)}{6n}=\dfrac{n^2-1}{6n}$

La deuxième est : $\sum\limits_{j=2}^{n}\sum\limits_{i=1}^{j-1}\dfrac{(i-1)(n-j+1)}{n^2(n-1)}$ donc est $\sum\limits_{j=2}^{n}\dfrac{(j-2)(j-1)(n-j+1)}{2n^2(n-1)}$.

On peut ajouter le terme correspondant à $j=1$ et connaissant la somme des premiers entiers, des premiers carrés et des premiers cubes, cette deuxième somme vaut $\dfrac{n^2-n-2}{12n}$

Finalement $V(X_n)=\dfrac{n+1}{12}$.

Exercice 3.21.

Soit $\sigma>0$. Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$ qui suit la loi normale centrée, d'écart-type σ.

1. Montrer que la variable aléatoire $U=\dfrac{X^2}{2\sigma^2}$ suit une loi γ dont on précisera le paramètre.

Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes définies sur $(\Omega,\mathcal{A},P)$ de même loi que X.

2. a) Pour tout $n\in\mathbb{N}^*$, déterminer la loi de $S_n=\dfrac{1}{2\sigma^2}\sum\limits_{i=1}^{n}X_i^2$.

b) Pour tout $n\in\mathbb{N}^*$, montrer que la variable aléatoire $Y_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2$ est un estimateur sans biais de σ^2.

3. a) Montrer l'existence et calculer l'espérance $E(\sqrt{Y_n})$ de $\sqrt{Y_n}$ en fonction de n et σ.

b) En déduire que $T_n=\dfrac{\Gamma(n/2)}{\Gamma((n+1)/2)}\sqrt{\dfrac{nY_n}{2}}$ est un estimateur sans biais de σ.

c) Montrer l'existence et calculer la variance $V(T_n)$ de T_n en fonction de n, σ.

d) Montrer que $(T_n)_n$ est un estimateur convergent de σ.

(*on admettra que, pour tout $x>0$, on a :* $\Gamma(x+n)\underset{(n\to\infty)}{\sim}n^x(n-1)!$)

Solution :

1. On a $P(U\leqslant t)=0$ si $t<0$, et, pour tout $t\geqslant 0$, on a :
$P(U\leqslant\leqslant t)=P(X^2\leqslant 2\sigma^2 t)=F_X(\sigma\sqrt{2t})-F_X(-\sigma\sqrt{2t})$, où F_X désigne la fonction de répartition de X. La fonction F_X est continue sur $\mathbb{R}$ et C^1 sauf en un nombre fini de points. Il en est donc de même pour $t\mapsto P(U\leqslant t)$ sur $\mathbb{R}^*$ ainsi qu'en 0

$(F_X(\sigma\sqrt{2t}) - F_X(-\sigma\sqrt{2t}) \to F_X(0) - F_X(0) = 0)$. Donc U est à densité et une densité f_U de U s'obtient en dérivant, soit $f_U(t) = 0$ si $t \leqslant 0$ et, pour $t \in \mathbb{R}_+^*$:

$$f_U(t) = \frac{\sigma}{\sqrt{2t}} f_X(\sigma\sqrt{2t}) + \frac{\sigma}{\sqrt{2t}} f_X(-\sigma\sqrt{2t}) = \frac{1}{\sqrt{\pi}} t^{-\frac{1}{2}} \mathrm{e}^{-t} = \frac{1}{\Gamma(1/2)} t^{\frac{1}{2}-1} \mathrm{e}^{-t}$$

On reconnaît que $U \hookrightarrow \gamma(1/2)$.

2. a) D'après les résultats de stabilité du cours, comme les variables aléatoires $\dfrac{X_i^2}{2\sigma^2}$ sont indépendantes et de même loi $\gamma(1/2)$, on sait qu'alors leur somme S_n suit la loi $\gamma(n/2)$.

b) On en déduit que $E(S_n) = n/2$ et comme $Y_n = \dfrac{2\sigma^2}{n} S_n$, on a donc : $E(Y_n) = \sigma^2$, soit :

$$Y_n \text{ est un estimateur sans biais de } \sigma^2.$$

3. a) Comme $\sqrt{Y_n} = \sigma\sqrt{\dfrac{2}{n}}\sqrt{S_n}$, on a $E(\sqrt{Y_n}) = \sigma\sqrt{\dfrac{2}{n}} E\left(\sqrt{S_n}\right)$. Or le théorème de transfert donne :

$E(\sqrt{S_n}) = \dfrac{1}{\Gamma(n/2)} \displaystyle\int_0^{+\infty} x^{\frac{1}{2}}\, x^{\frac{n}{2}-1} \mathrm{e}^{-x} dx = \dfrac{1}{\Gamma(n/2)} \Gamma((n+1)/2),$

où le calcul montre la convergence (absolue). Donc :

$$E(\sqrt{Y_n}) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \sigma \sqrt{\frac{2}{n}}$$

b) Ainsi $T_n = \dfrac{\Gamma(n/2)}{\Gamma((n+1)/2)} \sqrt{\dfrac{nY_n}{2}}$ est un estimateur sans biais de σ.

c) Soit $a_n = \dfrac{\Gamma(n/2)}{\Gamma((n+1)/2)}$, on a : $V(T_n) = \dfrac{n}{2} a_n^2 V(\sqrt{Y_n}) = \dfrac{n}{2} a_n^2 \big[E(Y_n) - [E(\sqrt{Y_n})]^2\big]$.

Or $E(Y_n) = \sigma^2$ et $[E(\sqrt{Y_n})]^2 = \dfrac{2\sigma^2}{n} \times \dfrac{1}{a_n^2}$, d'où : $V(T_n) = \sigma^2 \big(\dfrac{na_n^2}{2} - 1\big)$.

d) L'indication donne : $\Gamma\big(n + \frac{1}{2}\big) \sim \sqrt{n} \times (n-1)!$ et on sait que pour tout k de $\mathbb{N}^*$, on a : $\Gamma(k) = (k-1)!$.

Il y a donc deux cas :

- Si n est pair, $n = 2p$, alors

$$a_{2p} = \frac{\Gamma(p)}{\Gamma(p+1/2)} \sim \frac{1}{\sqrt{p}} \text{ soit } a_n \sim \sqrt{\frac{2}{n}}$$

- Si n est impair $n = 2p+1$,

$$a_{2p+1} = \frac{\Gamma(p+1/2)}{\Gamma(p+1)} = \frac{\Gamma(p+1/2)}{p\Gamma(p)} \sim \frac{\sqrt{p}}{p} \sim \frac{1}{\sqrt{p}} = \sqrt{\frac{2}{n-1}} \sim \sqrt{\frac{2}{n}}$$

Dans tous les cas on a : $a_n \sim \sqrt{\dfrac{2}{n}}$ et $\lim\limits_{n\to\infty} \dfrac{na_n^2}{2} = 1$, d'où $\lim\limits_{n\to\infty} V(T_n) = 0$ et donc, T_n est un estimateur convergent de σ.

Exercice 3.22.

On suppose que toutes les variables aléatoires de cet exercice sont définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

Soient $(X_n)_n$ une suite de variables aléatoires qui converge en probabilité vers X, $(c_n)_n$ une suite de réels qui converge vers un réel c et $\varepsilon > 0$.

1. Montrer que :
$$(|c_nX_n - cX| \geqslant \varepsilon) \subset (|c_n|\times|X_n - X| \geqslant \frac{\varepsilon}{2}) \cup (|c_n - c||X| \geqslant \frac{\varepsilon}{2})$$

2. Montrer que : $\lim\limits_{n\to+\infty} P\big(|X_n - X| \geqslant \dfrac{\varepsilon}{|c| + \varepsilon}\big) = 0$.

3. Montrer que : $\lim\limits_{n\to+\infty} P\big(|c_n - c||X| \geqslant \varepsilon\big) = 0$.

4. En déduire que $(c_nX_n)_n$ converge en probabilité vers cX.

5. **Application.** Soit $(X_n)_n$ une suite de variables aléatoires indépendantes, de même loi, possédant un moment d'ordre 4 et d'espérance μ. Montrer que la suite de variables aléatoires $\Big(\dfrac{1}{\binom{n}{2}} \sum\limits_{1\leq i<j\leq n} X_iX_j\Big)_n$ converge en probabilité vers μ^2.

Solution :

1. On remarque que si $(|c_n|(X_n - X) < \varepsilon/2)$ et $(|c_n - c|X < \varepsilon/2)$ sont réalisés, alors d'après l'inégalité triangulaire, on réalise $(|c_n(X_n - X) + (c_n - c)X| < \varepsilon)$.

On passe ensuite au complémentaire.

2. Il s'agit de la définition de la convergence en probabilité.

3. Comme (c_n) converge vers c, alors $\bigcap\limits_{n\in\mathbb{N}} (|c_n - c||X| \geqslant \varepsilon) = \emptyset$. Ainsi, d'après le théorème de la limite monotone, $\lim\limits_{n\to+\infty} P(|c_n - c||X| \geqslant \varepsilon) = 0$.

4. Soit $\varepsilon > 0$. Comme (c_n) converge vers c, il existe un rang n_0 à partir duquel $|c_n - c| < \varepsilon$, et donc $|c_n| < |c| + \varepsilon$. Ainsi, d'après la question 1,
$$\begin{aligned} P(|c_nX_n - cX| \geqslant \varepsilon) &\leqslant P(|c_n||X_n - X| \geqslant \varepsilon/2) + P(|c_n - c||X| \geqslant \varepsilon/2) \\ &\geqslant P((|c| + \varepsilon)|X| \geqslant \varepsilon/2) + P(|c_n - c||X| \geqslant \varepsilon/2) \end{aligned}$$
D'après la question précédente, cette quantité converge bien vers 0.

5. On remarque que
$$\underbrace{\binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} X_iX_j}_{Y_n} = \underbrace{\frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2}_{Z_n} - \underbrace{\frac{1}{n(n-1)}\sum_{i=1}^n X_i^2}_{W_n}$$
D'une part, comme la suite $(X_i^2)_n$ est une suite de variables aléatoires indépendantes de même loi possédant un moment d'ordre 2, d'après la loi faible des grands nombres, $\dfrac{1}{n}\Big(\sum\limits_{i=1}^n X_i^2\Big)_n$ converge vers $E(X_1^2)$. Ainsi, d'après la question précédente,

$\left(\frac{1}{n(n-1)}\sum_{i=1}^{n} X_i^2\right)_n$ converge en probabilité vers 0.

D'autre part, comme toujours d'après la loi faible des grands nombres $\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)_n$ converge en probabilité vers μ, et comme la fonction $x \mapsto x^2$ est continue, $\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right)_n$ converge en probabilité vers μ^2. Enfin, comme $\frac{n}{n-1}$ converge vers 1, le premier terme converge en probabilités vers μ^2.

Ainsi, en reprenant le raisonnement de la question 1,

$$P(|Y_n - \mu^2| \geqslant \varepsilon) \leqslant P(|Z_n - \mu^2| \geqslant \varepsilon/2) + P(|W_n| \geqslant \varepsilon/2) \underset{n\to\infty}{\longrightarrow} 0.$$

Exercice 3.23.

On admet que $\int_0^{+\infty} \frac{\sin t}{t}\, dt$ converge et que sa valeur est $\frac{\pi}{2}$.

1. Montrer que pour tout n entier naturel, $u_n = \int_0^{+\infty} \frac{1-\cos^n(t)}{t^2} dt$ existe et que $u_1 = \frac{\pi}{2}$.

2. a) Montrer que pour tout s réel, $\int_0^{+\infty} \frac{1-\cos(st)}{t^2}\, dt$ converge et que

$$\int_0^{+\infty} \frac{1-\cos(st)}{t^2}\, dt = \frac{\pi}{2}|s|$$

b) En déduire la valeur de u_2.

Dans la suite, on considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, définies sur $(\Omega, \mathcal{A}, P)$, à valeurs dans $\{-1, 1\}$ et telles que, pour tout $k \in \mathbb{N}^*$,

$$P(X_k = 1) = P(X_k = -1) = \frac{1}{2}$$

Pour tout $n \in \mathbb{N}^*$, on pose $S_n = X_1 + \cdots + X_n$.

3. Déterminer l'espérance et la variance de S_n.

4. a) Calculer pour tout réel t, $E(\sin(X_1 t))$, puis $E(\cos(X_1 t))$.

b) Montrer par récurrence sur n, que pour tout entier $n \in \mathbb{N}^*$ et tout réel t, on a : $E(\cos(S_n t)) = (\cos t)^n$.

5. Montrer que pour tout $n \in \mathbb{N}^*$, $E(|S_n|) = \frac{2}{\pi} u_n$.

Solution :

1. Le cas $n = 0$ étant banal, nous supposerons $n \geqslant 1$. La fonction $t \mapsto \frac{1-\cos^n(t)}{t^2}$ est continue sur $\mathbb{R}^{+*}$.

• Au voisinage de 0, un développement limité donne $\frac{1-\cos^n(t)}{t^2} \sim \frac{nt^2}{2t^2} = \frac{n}{2}$. L'intégrale est donc faussement impropre en 0.

• Au voisinage de $+\infty$, on a $0 \leqslant \dfrac{1-\cos^n(t)}{t^2} \leqslant \dfrac{2}{t^2}$ dont l'intégrale est convergente sur $[1,+\infty[$.

Une intégration par parties donne, en se plaçant d'abord sur un segment de $\mathbb{R}_+^*$, puis en passant aux limites :

$$\int_0^{+\infty} \frac{\sin t}{t} dt = \Big[\frac{1-\cos t}{t}\Big]_{\to 0}^{\to +\infty} + \int_0^{+\infty} \frac{1-\cos t}{t^2} dt$$

Ainsi $\displaystyle\int_0^{+\infty} \frac{\sin t}{t} dt = \int_0^{+\infty} \frac{1-\cos t}{t^2} dt = u_1$.

2. a) Si $s=0$, on a $\forall\, t \in \mathbb{R}_+^*$, $\dfrac{1-\cos(st)}{t^2} = 0$ donc $|s| = 0 = \dfrac{2}{\pi}\displaystyle\int_0^{\infty} \frac{1-\cos(0\times t)}{t^2}\, dt$.

Si $s>0$, le changement de variable $t = su$ est de classe C^1 strictement croissant de $\mathbb{R}_+^*$ sur $\mathbb{R}_+^*$ et on a donc

$$\frac{\pi}{2} = \int_0^{+\infty} \frac{1-\cos(t)}{t^2}\, dt = \int_0^{+\infty} \frac{1-\cos(su)}{s^2u^2}\, s\, du = \frac{1}{s}\int_0^{+\infty} \frac{1-\cos(su)}{u^2}\, du$$

Par parité du cosinus, on a donc $\forall\, s \in \mathbb{R},\ |s| = \dfrac{2}{\pi}\displaystyle\int_0^{+\infty} \frac{1-\cos(su)}{u^2}\, du$.

b) Ainsi $\displaystyle\int_0^{+\infty} \frac{1-\cos(2u)}{u^2}\, du = \pi$ et comme $1-\cos(2u) = 2(1-\cos^2 u)$, cela s'écrit :

$$\pi = 2u_2$$

3. Pour tout $k \in \mathbb{N}^*$, $E(X_k) = 0$ et $V(X_k) = 1$. Par linéarité de l'espérance et indépendance pour la variance, il vient donc $E(S_n) = 0$ et $V(S_n) = n$. (Notons que S_n prend toutes les valeurs entières de $-n$ à n qui ont même parité que n)

4. a) Par le théorème de transfert :

$$E(\sin(X_1t)) = \tfrac{1}{2}\,(\sin(-t) + \sin(t)) = 0$$

et

$$E(\cos(X_1t)) = \tfrac{1}{2}\,(\cos(-t) + \cos(t)) = \cos(t)$$

b) Pour $n=2$, comme $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$, il vient, par indépendance et linéarité :

$$\begin{aligned} E(\cos(X_1+X_2)t) &= E(\cos(X_1t)\cos(X_2t)) - E(\sin(X_1t)\sin(X_2t)) \\ &= E(\cos(X_1t))E(\cos(X_2t)) - 0 = \cos^2(t) \end{aligned}$$

La récurrence suit avec le même schéma de démonstration :

$$E(\cos((X_1+\cdots+X_n)t)) = (\cos t)^n$$

5. On a grâce à 2. a) :

$$E(|S_n|) = \sum_{s\in S_n(\Omega)} |s| P(S_n=s) = \frac{2}{\pi}\sum_{s\in S_n(\Omega)} P(S_n=s)\int_0^{+\infty} \frac{1-\cos(su)}{u^2}\, du$$

et la somme étant finie :

$$E(|S_n|) = \frac{2}{\pi}\int_0^{+\infty} \sum_{s\in S_n(\Omega)} P(S_n = s)\frac{1-\cos(su)}{u^2}\,du$$
$$= \frac{2}{\pi}\int_0^{\infty} \frac{1}{u^2}\big[1 - \sum_{s\in S_n(\Omega)} P(S_n = s)\cos(su)\big]\,du$$
$$= \frac{2}{\pi}\int_0^{+\infty} \frac{1}{u^2}\big[1 - E(\cos(S_n u))\big]\,du = \frac{2}{\pi}\int_0^{+\infty}\frac{1}{u^2}\big[1-(\cos(u))^n\big]\,du = \frac{2}{\pi}\,u_n$$

Exercice 3.24.

Soit n un entier de $\mathbb{N}^*$. Soient $X_1,\dots,X_n$ des variables aléatoires indépendantes, définies sur un même espace probabilisé $(\Omega,\mathcal{A},P)$, chacune d'espérance nulle et de variance notée $V(X_i)$ $(i \in [\![1,n]\!])$. On pose pour tout $n \geqslant 1$

$$S_n = \sum_{i=1}^{n} X_i$$

L'objet de cet exercice est de démontrer l'inégalité suivante :

$$\text{pour tout réel } t > 0\ ,\quad P(\max_{1\leq i\leq n}|S_i| \geqslant t) \leqslant \frac{1}{t^2}\sum_{i=1}^{n} V(X_i) \qquad (\star)$$

1. a) Calculer pour $k \in [\![1,n]\!]$, $E(S_n - S_k)$.

   b) Calculer $V(S_n)$ en fonction des $V(X_i)$

2. Soit $A = \big(\max_{1\leq k\leq n}|S_k| \geqslant t\big)$ et :

$$\text{pour tout } k \in [\![2,n]\!],\ A_k = \Big(\bigcap_{j=1}^{k-1}(|S_j| < t)\Big)\cap(|S_k|\geqslant t) \text{ et } A_1 = (|S_1| \geqslant t)$$

Montrer que les événements $(A_k)_{1\leq k\leq n}$ forment une partition de A.

3. a) Montrer que $E(S_k^2\times\mathbf{1}_{A_k}) \geqslant t^2P(A_k)$.

   b) Montrer que $E(S_n^2) \geqslant \sum_{k=1}^{n} E(S_n^2\times\mathbf{1}_{A_k})$.

   c) Montrer que, pour tout $k \in [\![1,n]\!]$,
   $E(S_n^2\times\mathbf{1}_{A_k}) \geqslant E(S_k^2\times\mathbf{1}_{A_k}) + 2E(S_n - S_k)E(S_k\times\mathbf{1}_{A_k})$
(utiliser : $S_n^2 = (S_k + (S_n - S_k))^2$).

   d) En déduire que $E(S_n^2) \geqslant t^2P(A)$. Conclure.

4. On suppose dans cette question que chaque X_i suit la loi uniforme sur $[-\sqrt{i},+\sqrt{i}]$. Que donne la majoration $(\star)$?

Solution :

1. a) Comme $S_n - S_k = \sum_{j=k+1}^{n} X_j$, il vient $E(S_n - S_k) = 0$, car pour tout j, $E(X_j) = 0$.

b) Par indépendance, $V(S_n) = \sum_{i=1}^{n} V(X_i)$.

2. $\star$ Les événements A_k et A_j sont disjoints pour $k \neq j$. En effet on peut supposer que $k < j$ et si $\omega \in A_k \cap A_j$, on a $|S_k(\omega)| > t$ et $|S_j(\omega)| \leqslant t$.

$\star$ On a $A = \bigcup_{k=1}^{n} A_k$, car :

- soit $\omega \in A$. Il existe $k \in [\![1, n]\!]$ tel que $|S_k(\omega)| \geqslant t$. Soit k_0 le premier indice k pour lequel $|S_k(\omega)| \geqslant t$. Par définition de k_0, on a $\omega \in A_{k_0}$, donc $\omega \in \displaystyle\bigcup_{k=1}^{n} A_k$.
- soit $\omega \in \displaystyle\bigcup_{k=1}^{n} A_k$. Il existe j tel que $\omega \in A_j$. Donc $|S_j(\omega)| \geqslant t$ ce qui entraîne que $\displaystyle\max_{1 \leq k \leq n} |S_k(\omega)| \geqslant t$.

3. a) On a $S_k^2 \mathbf{1}_{A_k}(\omega) = \begin{cases} 0 & \text{si } \omega \notin A_k \\ S_k^2(\omega) \geqslant t^2 & \text{si } \omega \in A_k \end{cases}$.

Par croissance de l'espérance $E(S_k^2 \mathbf{1}_{A_k}) \geqslant E(t^2 \mathbf{1}_{A_k}) = t^2 P(A_k)$.

b) Comme les événements $(A_k)_{1 \leq k \leq n}$ forment une partition de A, on a $\mathbf{1}_A = \displaystyle\sum_{k=1}^{n} \mathbf{1}_{A_k}$, car si U et V sont des événements incompatibles , $\mathbf{1}_{U \cup V} = \mathbf{1}_U + \mathbf{1}_V$. Comme $S_n^2 \geqslant S_n^2 \mathbf{1}_A$, il vient :

$$E(S_n^2) \geqslant E(S_n^2 \mathbf{1}_A) = \sum_{k=1}^{n} E(S_n^2 \mathbf{1}_{A_k})$$

c) Comme $S_n^2 = (S_k + (S_n - S_k))^2 = S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2$, il vient : $S_n^2 \mathbf{1}_{A_k} = S_k^2 \mathbf{1}_{A_k} + 2S_k(S_n - S_k)\mathbf{1}_{A_k} + (S_n - S_k)^2 \mathbf{1}_{A_k}$ et par croissance et positivité de l'espérance :

$$E(S_n^2 \mathbf{1}_{A_k}) \geqslant E(S_k^2 \mathbf{1}_{A_k}) + 2E((S_n - S_k) S_k \mathbf{1}_{A_k})$$

Les variables aléatoires $S_n - S_k = \displaystyle\sum_{j=k+1}^{n} X_j$ et $S_k = \displaystyle\sum_{j=1}^{k} X_j$ étant indépendantes, par le lemme des coalitions, il vient :

$$\begin{aligned} E(S_n^2 \mathbf{1}_{A_k}) &\geqslant E(S_k^2 1_{A_k}) + 2E((S_n - S_k) S_k \mathbf{1}_{A_k}) \\ &\geqslant E(S_k^2 \mathbf{1}_{A_k}) + 2E(S_n - S_k) E(S_k \mathbf{1}_{A_k}) = E(S_k^2 \mathbf{1}_{A_k}) \end{aligned}$$

d) En utilisant les questions précédentes :

$$E(S_n^2) \geqslant \sum_{k=1}^{n} E(S_n^2 \mathbf{1}_{A_k}) \geqslant \sum_{k=1}^{n} E(S_k^2 \mathbf{1}_{A_k}) \geqslant t^2 \sum_{k=1}^{n} P(A_k) = t^2 P(A)$$

Et comme $E(X_i) = 0$, il vient $P(A) \leqslant \dfrac{1}{t^2} E(S_n)^2 = \dfrac{1}{t^2} \displaystyle\sum_{i=1}^{n} V(X_i)$.

4. Dans ce cas particulier, $E(X_i) = 0$ et $V(X_i) = \dfrac{i}{3}$. Donc $\displaystyle\sum_{i=1}^{n} V(X_i) = \dfrac{n(n+1)}{6}$ et :

$$P(\max_{1 \leq i \leq n} |S_i| \geqslant t) \leqslant \frac{n(n+1)}{6t^2}$$

Exercice 3.25.

Soit n un entier tel que $n \geqslant 2$. Une urne contient initialement n boules numérotées de 1 à n. On vide l'urne en extrayant toutes les boules une par une au hasard et sans remise. Pour tout $i \in [\![1, n]\!]$, on note X_i la variable aléatoire qui vaut 1 si la boule obtenue au i-ème tirage porte le numéro i et qui vaut 0 dans le cas contraire.

1. a) Quelle est la loi de X_i ?

b) En déduire l'espérance de la variable aléatoire N qui donne le nombre de fois où, lors du tirage, le rang du tirage et le numéro de la boule obtenue sont égaux.

Pour tout $k \in [\![1, n]\!]$, on dira que le résultat du k-ième tirage est un « sommet » lorsque la boule obtenue à ce tirage porte un numéro strictement supérieur à tous les numéros obtenus jusqu'alors (en particulier, lorsque $k = 1$, la première boule est toujours un sommet).

2. Combien existe-t-il de façons de vider l'urne pour lesquelles il n'y a qu'un seul sommet ?
Combien existe-t-il de façons de vider l'urne pour lesquelles il y a n sommets ?

3. Montrer la relation suivante : pour tout $p, q \in \mathbb{N}$, $\sum\limits_{k=0}^{q} \binom{p+k}{p} = \binom{p+q+1}{p+1}$.

4 a) Soient $k \in [\![2, n]\!]$ et $j \in [\![k, n]\!]$ fixés. Combien existe-t-il de façons de vider l'urne, pour lesquelles, à la fois la k-ième boule obtenue porte le numéro j et le k-ième tirage constitue un sommet ?

b) Combien existe-t-il de façons de vider l'urne pour lesquelles le k-ième tirage est un sommet ? En déduire la probabilité que le k-ième tirage soit un sommet.

5. Soit R le nombre aléatoire de sommets obtenus lorsque l'on vide l'urne. Déterminer, sous forme d'une somme, l'espérance de R.

Solution :

1. a) On modélise l'expérience aléatoire par l'univers Ω qui est l'ensemble des permutations des n boules (tirages de toutes les boules sans remise) muni de la probabilité uniforme ; alors $\text{card}(\Omega) = |\Omega| = n!$. L'événement A_i « la boule obtenue au i-ème tirage porte le numéro i » est formé des permutations qui tirent cette boule i au i-ème tirage et les $n-1$ autres boules aux $n-1$ autres rangs. Ainsi : $P(A_i) = \dfrac{|A_i|}{|\Omega|} = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$.

On reconnaît que X_i suit la loi de Bernoulli de paramètre $\dfrac{1}{n}$.

b) Comme $N = \sum\limits_{i=1}^{n} X_i$ et $E(X_i) = \dfrac{1}{n}$, par linéarité de l'espérance, on a :

$$E(N) = n \times \frac{1}{n} = 1$$

2. Le premier tirage étant considéré comme un sommet, s'il n'y a eu qu'un seul sommet, 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 sommet.

S'il y a n sommets, 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.

3. La relation à démontrer est claire pour $q = 0$ et passer d'un rang q au rang $q+1$ résulte de la formule de Pascal.

4. a) Si le k-ième tirage est un sommet et porte le numéro j, 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 $\binom{j-1}{k-1}(k-1)!$ façons (arrangements) de tirer les boules sorties lors des $k-1$ premiers tirages parmi les numéros compris entre 1 et $j-1$; puis une seule façon de placer la boule j en place k ; et la fin du tirage est une permutation quelconque des $n-k$ boules restantes. D'après le lemme des bergers, le nombre recherché est :

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

b) Il reste à sommer sur les valeurs de j accessibles et le nombre cherché est :

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

(d'après le résultat de la question 3.) La probabilité recherchée est donc :

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

Remarque. On peut retrouver ce résultat directement avec le raisonnement suivant : dire que le k-ième tirage est un sommet, 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 $\dfrac{1}{k}$.

5. Pour tout $k \in [\![1, n]\!]$, soit Y_k la variable aléatoire qui prend la valeur 1 si le résultat du k-ième tirage est un sommet et 0 sinon. D'après la question précédente, la variable aléatoire Y_k suit la loi de Bernoulli de paramètre $\dfrac{1}{k}$.

Comme $R = \sum_{k=1}^{n} Y_k$, par linéarité de l'espérance, on a :

$$E(R) = \sum_{k=1}^{n} E(Y_k) = \sum_{k=1}^{n} \frac{1}{k}$$