

PROBABILITÉS

Exercice 3.01.

Soit $(X_i)_{i \in \mathbb{N}^*}$ une famille de variables aléatoires à valeurs dans \mathbb{N}^* , indépendantes, de même loi et définies sur le même espace probabilisé (Ω, \mathcal{A}, P) . Pour tout entier $n \in \mathbb{N}^*$, on définit la variable aléatoire Y_n par :

$$Y_n = \min_{1 \leq i \leq n} X_i$$

1. Exprimer la fonction de répartition de Y_n à l'aide de la fonction $k \mapsto P(X_1 \geq k)$.
2. **Dans cette question seulement** on suppose que les X_i suivent la loi géométrique de paramètre $p \in]0, 1[$; on note $q = 1 - p$.
 - a) Calculer l'espérance de Y_n .
 - b) Étudier la convergence en loi de la suite $(Y_n)_{n \in \mathbb{N}^*}$.
3. Montrer que la série de terme général $P(X_1 \geq k)$ converge si et seulement si X_1 admet une espérance et qu'alors :

$$E(X_1) = \sum_{k=1}^{\infty} P(X_1 \geq k)$$

4. En déduire que si X_1 admet une espérance, alors Y_n admet une espérance et comparer ces deux espérances.
5. Soit N une variable aléatoire à valeurs dans \mathbb{N}^* indépendante des X_i , et soit Y la variable aléatoire définie par :

$$\forall \omega \in \Omega, Y(\omega) = \inf_{1 \leq i \leq N(\omega)} X_i(\omega)$$

Montrer que si X_1 admet une espérance, alors Y admet une espérance et que l'on a $E(Y) \leq E(X_1)$.

Solution :

1. La variable Y_n est à valeurs dans \mathbb{N}^* et, pour tout $k \in \mathbb{N}^*$ (les X_i sont indépendantes et de même loi que X_1) on a :

$$\begin{aligned} P(Y_n \leq k) &= 1 - P(X_1 > k, \dots, X_n > k) = 1 - P(X_1 > k) \cdots P(X_n > k) \\ &= 1 - P(X_1 > k)^n = 1 - P(X_1 \geq k+1)^n \end{aligned}$$

On en déduit que :

$$P(Y_n = k) = P(Y_n \leq k) - P(Y_n \leq k-1) = P(X_1 \geq k)^n - P(X_1 \geq k+1)^n.$$

2. a) Si X_1 suit la loi géométrique de paramètre p , on a $P(X_1 \geq k) = q^{k-1}$ donc :

$$P(Y_n = k) = q^{n(k-1)} - q^{nk} = (1 - q^n)(q^n)^{k-1}$$

on reconnaît la loi géométrique de paramètre $p' = 1 - q^n$.

$$\text{Donc } E(Y_n) = \frac{1}{1 - q^n}.$$

b) Comme $q \in]0, 1[$, et $P(Y_n = k) = (1 - q^n)(q^n)^{k-1}$, on a :

$$\forall k \in \mathbb{N}^*, \lim_{n \rightarrow \infty} P(Y_n = k) = \begin{cases} 0 & \text{si } k > 1 \\ 1 & \text{si } k = 1 \end{cases}$$

On reconnaît donc que (Y_n) converge en loi vers la variable constante égale à 1. (On peut raisonner sur les probabilités ponctuelles plutôt que sur les fonctions de répartition car toutes les variables aléatoires sont à valeurs dans \mathbb{N}^* .)

3. On remarque que :

$$P(X_1 \geq k) = \sum_{j=k}^{+\infty} P(X_1 = j) \text{ et que } jP(X_1 = j) = \sum_{k=1}^j P(X_1 = j).$$

Par sommation par paquets (tout est positif), on a donc l'équivalence des convergences et :

$$\begin{aligned} \sum_{k=1}^{\infty} P(X \geq k) &= \sum_{k=1}^{\infty} \sum_{j \leq k} P(X_1 = j) = \sum_{j=1}^{\infty} \sum_{k=1}^j P(X_1 = j) = \sum_{j=1}^{\infty} jP(X_1 = j) \\ &= E(X_1) \end{aligned}$$

4. D'après les calculs de la question 1, on a :

$$0 \leq P(Y_n \geq k) = 1 - P(Y_n \leq k-1) = P(X_1 > k-1) = P(X_1 \geq k) \leq P(X_1 \geq k)$$

car $P(X_1 \geq k) \in]0, 1[$. Par théorème de comparaison sur les séries à termes positifs, comme la série $\sum_{k \geq 0} P(X_1 \geq k)$ converge, il en est de même de la série $\sum_{k \geq 0} P(Y_n \geq k)$.

D'après la question précédente, la variable Y_n (à valeurs dans \mathbb{N}) a pour espérance $\sum_{k=0}^{\infty} P(Y_n \geq k)$

et celle-ci est inférieure à $\sum_{k=0}^{\infty} P(X_1 \geq k)$, soit $E(Y_n) \leq E(X_1)$.

(cette dernière inégalité provient aussi directement de $Y_n \leq X_1$)

5. C'est un théorème (admis) du cours ; ici on le démontre dans un cas particulier.

Comme $E(Y|N = n) = E(Y_n)$, on a

$$0 \leq E(Y|N = n)P(N = n) \leq E(X_1)P(N = n),$$

donc par comparaison à la série de terme général $E(X_1)P(N = n)$ qui est convergente, la série $\sum E(Y|N = n)P(N = n)$ converge.

Comme Y est une variable aléatoire positive, la formule de l'espérance totale s'applique, et on a :

$$E(Y) = \sum_{n=0}^{\infty} E(Y|N = n)P(N = n) \leq \sum_{n=0}^{\infty} E(X_1)P(N = n) = E(X_1)$$

Exercice 3.02.

Toutes les variables aléatoires considérées sont définies sur un même espace probabilisé (Ω, \mathcal{A}, P) donné. Pour tout ensemble fini E , on note $\text{card}(E)$ son cardinal (nombre de ses éléments).

Soit $\lambda > 0$. Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes qui suivent toutes la loi uniforme sur le segment $[0, 1]$. Soit Z une variable aléatoire indépendante des X_n et qui suit une loi de Poisson de paramètre λ .

1. Soit $a \in]0, 1]$. On définit les variables aléatoires N et N_n , pour $n \geq 1$, par :

$$\begin{aligned} \forall \omega \in \Omega, N_n(\omega) &= \text{card}\{i \in \llbracket 0, n-1 \rrbracket / X_i(\omega) \in [0, a]\} \\ \forall \omega \in \Omega, N(\omega) &= \text{card}\{i \in \llbracket 0, Z(\omega)-1 \rrbracket / X_i(\omega) \in [0, a]\} \end{aligned}$$

La variable aléatoire N_0 est la variable certaine égale à 0.

- a) Pour tout $n \geq 1$, déterminer la loi, l'espérance et la variance de N_n .
- b) Déterminer la loi, l'espérance et la variance de N .

c) Pour tout $\lambda > 0$, on pose $F_n(\lambda) = \int_0^\lambda \frac{u^n}{n!} e^{-u} du$.

Écrire la formule de Taylor avec reste intégral pour la fonction $x \mapsto e^{\lambda x}$ sur l'intervalle $[0, 1]$. En déduire que $P(Z \geq n+1) = F_n(\lambda)$ ainsi qu'une expression de $P(N \geq n+1)$ à l'aide de F_n .

2. On définit la suite de variables aléatoires $(T_n)_{n \geq 0}$ en posant, pour tout événement $\omega \in \Omega$, que les $Z(\omega)$ premiers termes $T_0(\omega), T_1(\omega), \dots, T_{Z(\omega)-1}(\omega)$ sont les nombres $X_0(\omega), X_1(\omega), \dots, X_{Z(\omega)-1}(\omega)$ rangés dans l'ordre croissant, et pour tout $n \geq Z(\omega)$, $T_n(\omega) = 2$.

Déterminer la fonction de répartition de T_n .

La variable aléatoire T_n est-elle discrète ? Est-elle à densité ?

3. Étudier la convergence en loi de la suite de variables aléatoires $(T_n)_{n \in \mathbb{N}}$.

Solution :

1. a) Les n événements $(0 \leq X_i \leq a)_{0 \leq i \leq n-1}$ sont indépendants de même probabilité a .

Or N_n compte le nombre de ces événements qui sont réalisés, donc N_n suit la loi binomiale de paramètres n et a , d'espérance na et de variance $na(1-a)$.

Le résultat reste vrai même pour $n = 0$.

b) Pour tout $k \in N(\Omega) = \mathbb{N}$, la formule des probabilités totales donne :

$$\begin{aligned} P(N = k) &= \sum_{n=0}^{\infty} P_{(Z=n)}(N = k)P(Z = n) = \sum_{n=k}^{\infty} P(N_n = k)P(Z = n) \\ &= \sum_{n=k}^{\infty} \binom{n}{k} a^k (1-a)^{n-k} \times \frac{\lambda^n}{n!} e^{-\lambda} = \sum_{j=0}^{\infty} \binom{j+k}{k} a^k (1-a)^j \times \frac{\lambda^{j+k}}{(j+k)!} e^{-\lambda} \\ &= \frac{(a\lambda)^k}{k!} e^{-\lambda} \sum_{j=0}^{\infty} \frac{[(1-a)\lambda]^j}{j!} = \frac{(a\lambda)^k}{k!} e^{-a\lambda} \end{aligned}$$

Donc N suit la loi de Poisson de paramètre λa et $E(N) = V(N) = \lambda a$.

c) La formule de Taylor avec reste intégral est :

$$f(b) = \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t) dt$$

Pour la fonction $x \mapsto e^{\lambda x}$ sur l'intervalle $[0, 1]$, cela donne :

$$e^\lambda = \sum_{k=0}^n \frac{\lambda^k}{k!} + \int_0^1 \frac{(1-t)^n}{n!} \lambda^{n+1} e^{\lambda t} dt$$

En divisant par e^λ et en faisant le changement de variable $u = (1-t)\lambda$ on obtient :

$$1 = P(Z \leq n) + \int_0^\lambda \frac{u^n}{n!} e^{-u} du, \text{ soit } P(Z \geq n+1) = F_n(\lambda)$$

Comme ceci est vrai pour n'importe quelle loi de Poisson, on a : $P(N \geq n+1) = F_n(\lambda a)$.

2. D'après leur définition les T_n sont à valeurs dans $[0, 1] \cup \{2\}$ et on a :

→ si $a \in]0, 1]$, $(T_n \leq a) = (N < n) = (N \geq n+1)$.

→ si $1 < a < 2$, $(T_n \leq a) = (T_n \leq 1)$

Donc :

$$P(T_n \leq x) = \begin{cases} 0 & \text{si } x \leq 0 \\ F_n(\lambda x) & \text{si } x \in]0, 1] \\ F_n(\lambda) & \text{si } x \in]1, 2[\\ 1 & \text{si } x \geq 2 \end{cases}$$

La variable T_n n'est pas à densité car sa fonction de répartition est discontinue en 2 (puisque $F_n(\lambda) \neq 1$) ; elle n'est pas discrète car sa fonction de répartition n'est pas en escalier.

3. On a : $F_n(\lambda) = P(Z \geq n+1) \xrightarrow{n \rightarrow \infty} 0$. Donc :

$$\lim_{n \rightarrow \infty} P(T_n \leq x) = \begin{cases} 0 & \text{si } x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

i.e. $(T_n)_n$ converge en loi vers la variable constante égale à 2.

Exercice 3.03.

Soit un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$, où Ω est un ensemble fini.

On note \mathcal{F} l'ensemble des variables aléatoires réelles définies sur Ω (qui sont toutes les applications de Ω dans \mathbb{R}). On rappelle que \mathcal{F} est un \mathbb{R} -espace vectoriel.

Si A est une partie de Ω , on note $\mathbf{1}_A$ la fonction caractéristique de A , c'est-à-dire l'application définie par :

$$\forall \omega \in \Omega, \mathbf{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}$$

1. Soit $A \subset \Omega$. Calculer l'espérance $E(\mathbf{1}_A)$ de $\mathbf{1}_A$.
2. Montrer que l'application $\varphi : (X, Y) \mapsto E(XY)$ définie sur $\mathcal{F} \times \mathcal{F}$ est un produit scalaire sur \mathcal{F} si et seulement si pour tout $\omega \in \Omega$, on a : $P(\{\omega\}) > 0$.

Dans la suite de l'exercice, on supposera que P vérifie cette propriété et \mathcal{F} sera muni de ce produit scalaire.

3. Soient $X, Y \in \mathcal{F}$ deux variables aléatoires non constantes.

On note G le sous-espace vectoriel de \mathcal{F} engendré par X et la variable aléatoire constante égale à 1, c'est-à-dire que $G = \text{Vect}(X, \mathbf{1}_\Omega)$.

a) Montrer l'existence et l'unicité de $(a_0, b_0) \in \mathbb{R}^2$ tel que $Y - a_0X - b_0$ est orthogonal à tout élément de G .

b) En déduire l'expression de la projection orthogonale de Y sur G .

On la notera $p_G(Y)$.

c) Comparer $E(p_G(Y))$ et $E(Y)$.

d) On suppose que $X = \mathbf{1}_A$, où A est une partie de Ω non vide et distincte de Ω .

Montrer que pour tout $B \subset \Omega$:

$$p_G(\mathbf{1}_B) = P_A(B) \times \mathbf{1}_A + P_{\bar{A}}(B) \times \mathbf{1}_{\bar{A}},$$

où $P_V(U)$ désigne la probabilité conditionnelle de l'événement U , sachant que l'événement V est réalisé.

Solution :

1. $E(\mathbf{1}_A) = 1.P(A) + 0.P(\bar{A}) = P(A)$.
2. • L'application φ est bilinéaire (par linéarité de l'espérance), symétrique (par commutativité du produit dans \mathbb{R} , donc dans \mathcal{F}), et positive (car si $X \in \mathcal{F}$, $\varphi(X, X) = E(X^2) \geq 0$).

• \rightarrow Si φ est un produit scalaire, alors pour tout $\omega \in \Omega$, $X = \mathbf{1}_{\{\omega\}}$ est un élément non nul de \mathcal{F} , donc $\varphi(X, X) = E(\mathbf{1}_{\{\omega\}}^2) = E(\mathbf{1}_{\{\omega\}}) = P(\{\omega\}) > 0$.

\rightarrow Réciproquement, soit $X \in \mathcal{F}$ non nulle. Alors il existe $\omega_0 \in \Omega$ tel que $X(\omega_0) \neq 0$ et on a :

$$\varphi(X, X) = E(X^2) \geq X^2(\{\omega_0\})P(\{\omega_0\}) > 0$$

donc φ est bien un produit scalaire.

3. a) On a :

$$\begin{aligned}
 Y - a_0X - b_0 \in G^\perp &\iff \begin{cases} \varphi(Y - a_0X - b_0, \mathbf{1}_\Omega) = 0 \\ \varphi(Y - a_0X - b_0, X) = 0 \end{cases} \\
 &\iff \begin{cases} a_0E(X) + b_0 = E(Y) \\ a_0E(X^2) + b_0E(X) = E(XY) \end{cases} \\
 &\iff \begin{cases} a_0 = \frac{E(XY) - E(X)E(Y)}{V(X)} = \frac{\text{Cov}(X, Y)}{V(X)} \\ b_0 = E(Y) - E(X) \frac{\text{Cov}(X, Y)}{V(X)} \end{cases}
 \end{aligned}$$

Notons que $V(X) \neq 0$, puisque X n'est pas (quasi)-constante.

b) Comme $p_G(Y)$ est l'unique élément Z de G tel que $Y - Z \in G^\perp$, on a :

$$p_G(Y) = \frac{\text{Cov}(X, Y)}{V(X)} X + E(Y) - E(X) \frac{\text{Cov}(X, Y)}{V(X)}$$

c) On a : $E(p_G(Y)) = a_0E(X) + b_0 = E(Y)$ (calcul de a_0 , ou conséquence de $Y - p_G(Y) \perp \mathbf{1}_\Omega$)

d) Comme $A \neq \Omega$ et $A \neq \emptyset$, la variable aléatoire $X = \mathbf{1}_A$ n'est pas constante ; on peut donc utiliser la formule obtenue à la question b) :

• Or

$$\begin{aligned}
 \text{Cov}(\mathbf{1}_A, \mathbf{1}_B) &= E(\mathbf{1}_A \mathbf{1}_B) - E(\mathbf{1}_A)E(\mathbf{1}_B) = E(\mathbf{1}_{A \cap B}) - E(\mathbf{1}_A)E(\mathbf{1}_B) \\
 &= P(A \cap B) - P(A)P(B).
 \end{aligned}$$

• $V(\mathbf{1}_A) = P(A)P(\bar{A})$ (car $\mathbf{1}_A$ suit la loi de Bernoulli de paramètre $P(A)$).

Comme

$$\begin{aligned}
 b_0 &= E(\mathbf{1}_B) - E(\mathbf{1}_A) \frac{\text{Cov}(\mathbf{1}_A, \mathbf{1}_B)}{V(\mathbf{1}_A)} = P(B) - \frac{P(A \cap B) - P(A)P(B)}{P(\bar{A})} \\
 &= P_{\bar{A}}(B), \\
 a_0 + b_0 &= E(\mathbf{1}_B) + \frac{\text{Cov}(\mathbf{1}_A, \mathbf{1}_B)}{V(\mathbf{1}_A)} (1 - E(\mathbf{1}_A)) \\
 &= P(B) + \frac{P(A \cap B) - P(A)P(B)}{P(A)} \\
 &= \frac{P(A \cap B)}{P(A)} = P_A(B)
 \end{aligned}$$

on a bien :

$$p_G(\mathbf{1}_B) = a_0 \mathbf{1}_A + b_0 (\mathbf{1}_A + \mathbf{1}_{\bar{A}}) = (a_0 + b_0) \mathbf{1}_A + b_0 \mathbf{1}_{\bar{A}} = P_A(B) \mathbf{1}_A + P_{\bar{A}}(B) \mathbf{1}_{\bar{A}}.$$

Exercice 3.04.

On dit que le réel m est une médiane de la variable aléatoire à densité X si, et seulement si :

$$P(X \leq m) = P(X \geq m) = \frac{1}{2}$$

Partie I

Soit X une variable aléatoire réelle de densité f_X , à valeurs dans $X(\Omega) =]a, b[$, où $-\infty \leq a < b \leq +\infty$.

On note F_X sa fonction de répartition.

1. Montrer que l'équation $F_X(x) = \frac{1}{2}$ d'inconnue réelle x admet au moins une solution dans \mathbb{R} .

2. On suppose que F_X est strictement croissante sur $]a, b[$. Montrer que l'équation $F_X(x) = \frac{1}{2}$ d'inconnue réelle x admet une unique solution dans \mathbb{R} .

Cette valeur s'appelle **la médiane** de X et sera notée $m(X)$.

Partie II

Soit U et V deux variables indépendantes suivant la loi uniforme $\mathcal{U}(]0, 1[)$.

1. a) Déterminer $m(U)$ et $E(U)$.

b) Déterminer $m(U^2)$ et $E(U^2)$.

2. On définit la variable Z par $Z = \frac{U}{V}$.

a) Déterminer la loi de $-V$.

b) En déduire une densité de $U - V$.

c) En déduire la médiane de Z .

Solution :

Partie I.

1. La variable aléatoire X est à densité, donc F_X est continue sur \mathbb{R} . De plus $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$, et $0 < 1/2 < 1$, donc d'après le théorème des valeurs intermédiaires, l'équation $F_X(x) = 1/2$ admet au moins une solution dans \mathbb{R} .

2. Comme $\lim_{x \rightarrow a} F(x) = 0$ et $\lim_{x \rightarrow b} F(x) = 1$, la stricte croissance de F sur $]a, b[$ assure l'unicité de m dans $]a, b[$ (et il ne peut pas y en avoir d'autres en dehors de cet intervalle!).

Partie II.

1. a) Pour $x \in [0, 1]$, $F_U(x) = x$, donc $m(U) = \frac{1}{2}$. D'autre part, $E(U) = \frac{1}{2}$ (c'est du cours)

b) $U^2(\Omega) = [0, 1]$ et pour $x \in [0, 1]$,

$$F_{U^2}(x) = P(U^2 \leq x) = P(U \leq \sqrt{x}) = F_U(\sqrt{x}) = \sqrt{x},$$

$$\text{donc } m(U^2) = \frac{1}{4}.$$

$$\text{Par la formule de Koenig-Huygens, } E(U^2) = V(U) + E(U)^2 = \frac{1}{12} + \left(\frac{1}{2}\right)^2 = \frac{1}{3}.$$

La médiane n'est donc pas toujours la moyenne !

2. a) D'après le cours, comme V suit une loi uniforme, alors $-V$ suit aussi une loi uniforme. On a donc $-V \hookrightarrow \mathcal{U}([-1, 0])$.

b) Comme U et V sont indépendantes, alors d'après le lemme des coalitions, U et $-V$ le sont.

La fonction f_U étant bornée, on peut prendre pour densité de $U - V$:

$$g : x \mapsto \int_{-\infty}^{+\infty} f_U(t) f_{-V}(x-t) dt = \int_0^1 f_{-V}(x-t) dt.$$

Comme $-V \hookrightarrow \mathcal{U}([-1, 0])$, on a :

$$f_{-V}(x-t) = 1 \iff -1 \leq x-t \leq 0 \iff t \in [x, x+1].$$

$$\text{On a donc : } g(x) = \int_{\max(0, x)}^{\min(1, x+1)} 1 dt = \min(1, x+1) - \max(0, x).$$

Ainsi :

$$-1 \leq x \leq 0 \implies g(x) = \int_0^{x+1} 1 dt = x+1, \quad 0 < x \leq 1 \implies g(x) = \int_x^1 1 dt = 1-x,$$

sinon, $g(x) = 0$.

c) On a : $[Z \leq 1] \iff [U \leq V] \iff U - V \leq 0$.

$$\text{Ainsi, } P(Z \leq 1) = P(U - V \leq 0) = \int_{-\infty}^0 g(t) dt = \int_{-1}^0 (t+1) dt = \left[\frac{t^2}{2} + t \right]_{-1}^0 = \frac{1}{2}, \text{ donc :}$$

$$m(Z) = 1$$

Exercice 3.05.

On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires indépendantes définies sur le même espace probabilisé (Ω, \mathcal{A}, P) . On suppose que X_n suit une loi de Bernoulli de paramètre $p_n \in]0, 1[$.

Pour tout $n \in \mathbb{N}^*$, on pose $Y_n = X_n + X_{n+1}X_{n+2}$.

On admettra que si une suite réelle $(u_n)_{n \in \mathbb{N}^*}$ converge vers ℓ , alors la suite de ses moyennes $(\frac{1}{n} \sum_{k=1}^n u_k)_{n \in \mathbb{N}^*}$ converge aussi vers ℓ .

1. Déterminer l'espérance et la variance de Y_n .

2. Soient i et j deux entiers strictement positifs et distincts. Étudier l'indépendance de Y_i et Y_j selon les valeurs de i et de j . Calculer la covariance des variables aléatoires Y_i et Y_j .

3. Pour tout entier n strictement positif, on pose $Z_n = \frac{Y_1 + \dots + Y_n}{n}$.

a) Calculer l'espérance de Z_n .

b) Montrer que pour tout $n \in \mathbb{N}^*$, on a : $V(Z_n) \leq \frac{8}{n}$.

4. Dans cette question, on considère deux variables aléatoires U et V définies sur (Ω, \mathcal{A}, P) , un nombre réel c et un réel strictement positif ε .

a) Montrer que : $[|U - c| \geq \varepsilon] \subseteq [|U - V| \geq \frac{\varepsilon}{2}] \cup [|V - c| \geq \frac{\varepsilon}{2}]$.

b) En déduire que $P(|U - c| \geq \varepsilon) \leq P(|U - V| \geq \frac{\varepsilon}{2}) + P(|V - c| \geq \frac{\varepsilon}{2})$.

5. On suppose dans cette question que la suite $(p_n)_{n \in \mathbb{N}^*}$ converge vers p . Montrer que la suite $(Z_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers une variable aléatoire que l'on déterminera.

Solution :

1. En utilisant l'indépendance des variables de Bernoulli X_n , on trouve :

$$E(Y_n) = E(X_n) + E(X_{n+1})E(X_{n+2}) = p_n + p_{n+1}p_{n+2}$$

De plus, on a :

$$E(Y_n^2) = E(X_n + X_{n+1}X_{n+2} + 2X_nX_{n+1}X_{n+2}) = p_n + p_{n+1}p_{n+2} + 2p_n p_{n+1}p_{n+2}.$$

D'où :

$$\begin{aligned} V(Y_n) &= E(Y_n^2) - E(Y_n)^2 = p_n + p_{n+1}p_{n+2} + 2p_n p_{n+1}p_{n+2} - (p_n + p_{n+1}p_{n+2})^2 \\ &= p_n + p_{n+1}p_{n+2} - p_n^2 - p_{n+1}^2 p_{n+2}^2 = p_n(1 - p_n) + p_{n+1}p_{n+2}(1 - p_{n+1}p_{n+2}). \end{aligned}$$

2. Par symétrie, on peut supposer que $i < j$. On voit (lemme des coalitions) que les variables Y_i et Y_j sont indépendantes dès que $j > i + 2$ et dans ce cas $\text{Cov}(Y_i, Y_j) = 0$. Elle semblent clairement dépendantes lorsque $j = i + 1$ ou $j = i + 2$ (ce qui sera confirmé par le calcul de la covariance).

Posons $q_k = 1 - p_k$, en utilisant la bilinéarité de la covariance et l'indépendance, il vient :

$$\begin{aligned} \text{Cov}(Y_i, Y_{i+1}) &= \text{Cov}(X_{i+1}X_{i+2}, X_{i+1}) + \text{Cov}(X_{i+1}X_{i+2}, X_{i+2}X_{i+3}) \\ &= E(X_{i+1}X_{i+2}) - E(X_{i+1}X_{i+2})E(X_{i+1}) + E(X_{i+1}X_{i+2}X_{i+3}) \\ &\quad - E(X_{i+1}X_{i+2})E(X_{i+2}X_{i+3}) \\ &= p_{i+1}p_{i+2}q_{i+1} + p_{i+1}p_{i+2}p_{i+3}q_{i+2} \end{aligned}$$

De façon analogue, on trouve $\text{Cov}(Y_i, Y_{i+2}) = p_{i+1}p_{i+2}q_{i+2}1$.

3. a) Par linéarité de l'espérance, on obtient : $E(Z_n) = \frac{1}{n} \sum_{k=1}^n (p_k + p_{k+1}p_{k+2})$.

b) Compte tenu des expressions trouvées précédemment, les propriétés de la variance nous donnent pour $n \geq 3$:

$$\begin{aligned} V(Z_n) &= \frac{1}{n^2} \left[\sum_{k=1}^n V(Y_k) + 2 \sum_{i=1}^{n-1} \text{Cov}(Y_i, Y_{i+1}) + 2 \sum_{i=1}^{n-2} \text{Cov}(Y_i, Y_{i+2}) \right] \\ &\leq \frac{2n + 4(n-1) + 2(n-2)}{n^2} \leq \frac{8}{n}. \end{aligned}$$

On vérifie que le résultat final reste valide pour les premières valeurs de n .

4. a) Si ω appartient au complémentaire de l'ensemble de droite, l'inégalité triangulaire nous dit que ω est dans le complémentaire de l'ensemble de gauche.

b) Il suffit d'utiliser la question précédente et le fait que pour deux événements A et B on a toujours $P(A \cup B) \leq P(A) + P(B)$.

5. Comme la suite $(p_n)_{n \geq 1}$ converge vers p , on déduit de la propriété admise que $E(Z_n)$ tend vers $p + p^2$, donc que la suite de variables certaines $(E(Z_n))_{n \geq 1}$ converge en probabilité vers la variable certaine $p + p^2$. L'inégalité de Bienaymé-Tchebychev et la question 4. b) permettent d'écrire :

$$P(|Z_n - (p^2 + p)| \geq \varepsilon) \leq \frac{32}{n\varepsilon^2} + P(|E(Z_n) - p^2 - p| \geq \frac{\varepsilon}{2}) \xrightarrow{n \rightarrow \infty} 0$$

La suite de variables aléatoires $(Z_n)_n$ converge donc en probabilité vers la variable constante égale $p^2 + p$.

Exercice 3.06.

Dans cet exercice, les variables aléatoires sont définies sur un espace probabilisé (Ω, \mathcal{A}, P) .

Pour $\alpha > 0$ donné, on dit qu'une variable aléatoire X est α -sous-gaussienne si pour tout t réel,

$$E(e^{tX}) \leq e^{\alpha^2 t^2 / 2}$$

1. Soit X de loi normale centrée réduite. Montrer que X est 1-sous-gaussienne.

2. Montrer que pour tout t réel $\frac{e^t + e^{-t}}{2} \leq e^{t^2/2}$. (on pourra utiliser l'écriture de l'exponentielle sous forme de série).

3. Soit t réel. Montrer que pour $x \in [-1, 1]$, on a

$$e^{tx} \leq \frac{1+x}{2} e^t + \frac{1-x}{2} e^{-t}$$

4. Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes et α -sous-gaussiennes.

Soit μ_1, \dots, μ_n des réels tels que $\sum_{i=1}^n \mu_i^2 = 1$. Montrer que $\sum_{i=1}^n \mu_i X_i$ est α -sous gaussienne.

5. Soit X α -sous-gaussienne et $\lambda > 0$.

a) Montrer que pour tout $t > 0$, on a

$$P(X \geq \lambda) \leq \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right)$$

b) En déduire que

$$P(|X| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2\alpha^2}\right)$$

Solution :

1. Par le théorème de transfert et parce que $u \rightarrow e^{tu-t^2/2}$ est continue sur \mathbb{R} et négligeable devant $1/u^2$ au voisinage de l'infini :

$$E(e^{tX}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tu-u^2/2} du = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-1/2(t-u)^2} du = e^{t^2/2}$$

X est bien 1-sous gaussienne.

2. On sait que $e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}$. Donc :

$$\frac{e^t + e^{-t}}{2} = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} \leq \sum_{n=0}^{\infty} \frac{t^{2n}}{2^n n!} = e^{t^2/2}$$

3. On remarque que $0 \leq \frac{1-x}{2}$, que $\frac{1+x}{2} \leq 1$ et que $\frac{1-x}{2} + \frac{1+x}{2} = 1$. Il reste à utiliser la convexité de la fonction exponentielle sur l'intervalle $[-t, t]$.

4. Par le lemme des coalitions et par indépendance, en notant $S_n = \sum_{i=1}^n \mu_i X_i$:

$$E(e^{tS_n}) = E\left(\prod_{i=1}^n e^{t\mu_i X_i}\right) = \prod_{i=1}^n E(e^{t\mu_i X_i}) \leq \prod_{i=1}^n e^{(\alpha^2 \mu_i^2 t^2)/2} = e^{(\alpha^2 t^2)/2}$$

5. En utilisant la croissance de la fonction exponentielle, $t > 0$, et l'inégalité de Markov, il vient

$$P(X \geq \lambda) = P(e^{tX} \geq e^{t\lambda}) \leq \frac{E(e^{tX})}{e^{t\lambda}} \leq \frac{e^{\alpha^2 t^2/2}}{e^{t\lambda}} = \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right)$$

De la même façon, car $-X$ est α -sous gaussienne par la question précédente :

$$P(-X \geq \lambda) = P(e^{-tX} \geq e^{t\lambda}) \leq \frac{E(e^{-tX})}{e^{t\lambda}} \leq \frac{e^{\alpha^2 t^2/2}}{e^{t\lambda}} = \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right)$$

Donc

$$P(|X| \geq \lambda) \leq 2 \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right)$$

ceci pour tout $t > 0$.

En prenant $t = \frac{\lambda}{\alpha^2}$, (une étude rapide de la fonction associée montre que c'est la valeur qui minimise l'exposant de l'exponentielle), on obtient le résultat demandé.

Exercice 3.07.

Dans cet exercice, X_1, \dots, X_n désignent des variables aléatoires **non nécessairement indépendantes** définies sur un espace probabilisé (Ω, \mathcal{A}, P) .

Pour $n \in \mathbb{N}^*$, on note $M_n = \sup_{1 \leq i \leq n} X_i$ l'application définie sur Ω par, pour tout $\omega \in \Omega$:

$$M_n(\omega) = \max_{1 \leq i \leq n} X_i(\omega)$$

1. Montrer que M_n est une variable aléatoire réelle.
2. Montrer que pour tout réel t , $P(M_n > t) \leq \sum_{i=1}^n P(X_i > t)$.
3. On suppose dans cette question que les X_i suivent la même loi admettant un moment d'ordre $p + 1$, avec $p \in \mathbb{N}^*$.

Montrer que $\frac{1}{n^{1/p}} \sup_{1 \leq i \leq n} |X_i|$ converge en probabilité vers 0.

4. On suppose dans cette question que les X_i suivent la même loi normale centrée réduite.

- a) Montrer que pour tout $t > 0$, $P(|X_1| > t) \leq \frac{e^{-t^2/2}}{t}$.
- b) En déduire un majorant de $P(\sup_{1 \leq i \leq n} |X_i| \geq t)$, pour $t > 0$.
- c) Montrer que pour tout $s > \sqrt{2}$, $\lim_{n \rightarrow +\infty} [P(\sup_{1 \leq i \leq n} |X_i| \geq s(\ln n)^{1/2})] = 0$.

Solution :

1. Pour montrer que M_n est une variable aléatoire, il suffit d'écrire que pour tout t , $[M_n < t] = \bigcap_{i=1}^n [X_i < t]$, intersection d'éléments de la tribu \mathcal{A} donc élément de ladite tribu.

2. On a

$$\{\omega/M_n(\omega) > t\} = \{\omega/\exists i/X_i(\omega) > t\} = \bigcup_{i=1}^n \{\omega/X_i(\omega) > t\}$$

et

$$P(M_n > t) = P\left(\bigcup_{i=1}^n [X_i > t]\right) \leq \sum_{i=1}^n P(X_i > t)$$

3. Dans ce cas, la majoration précédente devient : $P(M_n > t) \leq nP(X > t)$.

On utilise alors l'inégalité de Markov et $|X_i|$ au lieu de X_i . On note $Y_n = \sup_{1 \leq i \leq n} |X_i|$ et il

vient, pour $\varepsilon > 0$:

$$\begin{aligned} P(Y_n > n^{1/p}\varepsilon) &\leq nP(|X| > n^{1/p}\varepsilon) = nP(|X|^{p+1} > n^{1+1/p}\varepsilon^{p+1}) \\ &\leq n \frac{E(|X|^{p+1})}{n^{1+1/p}\varepsilon^{p+1}} = \frac{C(\varepsilon, p)}{n^{1/p}} \end{aligned}$$

dernière quantité qui tend vers 0 lorsque n tend vers $+\infty$.

4. a) On utilise une intégration par parties :

$$\begin{aligned} P(|X| \geq t) &= \frac{2}{\sqrt{2\pi}} \int_t^{+\infty} e^{-u^2/2} du = \frac{2}{\sqrt{2\pi}} \int_t^{+\infty} \frac{ue^{-u^2/2}}{u} du \\ &= \sqrt{\frac{2}{\pi}} \left(\left[-\frac{e^{-u^2/2}}{u} \right]_t^{+\infty} - \int_t^{+\infty} \frac{e^{-u^2/2}}{u^2} du \right) \leq \frac{e^{-t^2/2}}{t} \end{aligned}$$

b) D'après les questions précédentes, pour $t > 0$

$$P\left(\sup_{1 \leq i \leq n} |X_i| \geq t\right) \leq n \frac{e^{-t^2/2}}{t}$$

c) Enfin, pour $s^2 > 2$:

$$\begin{aligned} P\left(\sup_{1 \leq i \leq n} |X_i| \geq s(\ln n)^{1/2}\right) &\leq nP(|X| < s\sqrt{\ln n}) \\ &\leq n \frac{e^{-s^2 \ln n/2}}{s\sqrt{\ln n}} \leq n \frac{e^{-2 \ln n/2}}{s\sqrt{\ln n}} = \frac{1}{s\sqrt{\ln n}} \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

Exercice 3.08.

Dans cet exercice, $(X_n)_{n \geq 1}$ est une suite de variables aléatoires positives à densité, indépendantes de même loi. De même $(Y_n)_{n \geq 1}$ est une autre suite de variables aléatoires positives, à densité, indépendantes de même loi. On suppose que les densités de X_1 et Y_1 sont continues strictement positives sur \mathbb{R} et que X_1 et Y_1 admettent une espérance.

Enfin, on suppose que pour tout $n \geq 1$, on a $E\left(\min_{1 \leq i \leq n} X_i\right) = E\left(\min_{1 \leq i \leq n} Y_i\right)$.

1. Soit Z une variable aléatoire positive à densité admettant une espérance.

Montrer que $E(Z) = \int_0^{+\infty} P(Z > t) dt$.

2. On note F la fonction de répartition de X_1 et $G = 1 - F$. Exprimer la fonction de répartition F_n de $\min_{1 \leq i \leq n} X_i$, puis $G_n = 1 - F_n$ en fonction de G .

3. Montrer que pour tout $n \geq 1$, $\int_0^{+\infty} [G(t)]^n dt = \int_0^1 nu^{n-1}G^{-1}(u) du$, où G^{-1} désigne la bijection réciproque de G (on effectuera le changement de variable $u = G(t)$).

On admet le résultat suivant : si f est une fonction continue sur un segment $[a, b]$ et si pour tout polynôme $Q \in \mathbb{R}[X]$, on a $\int_a^b Q(t)f(t) dt = 0$, alors f est la fonction nulle sur $[a, b]$.

4. Montrer que X_1 et Y_1 suivent la même loi.

Solution :

1. Soit F la fonction de répartition de Z et f une densité. Pour tout $A > 0$:

$$\int_0^A tf(t) dt = [t(F(t) - 1)]_0^A + \int_0^A (1 - F(t)) dt$$

Or $A(1 - F(A)) = A \int_A^{+\infty} f(u) du \leq \int_A^{+\infty} uf(u) du \xrightarrow{A \rightarrow +\infty} 0$ (reste d'une intégrale convergente).

Donc, par passage à la limite : $E(Z) = \int_0^{+\infty} (1 - F(t)) dt = \int_0^{+\infty} P(Z > t) dt$.

2. Pour tout $t > 0$: $\inf_{1 \leq i \leq n} (X_i > t) = \bigcap_{i=1}^n [X_i > t]$.

et, par indépendance :

$$G_n(t) = P(\inf_{1 \leq i \leq n} X_i > t) = P(\bigcap_{i=1}^n (X_i > t)) = \prod_{i=1}^n P(X_i > t) = [G(t)]^n.$$

3. La variable aléatoire X_1 étant à densité strictement positive, la fonction G est strictement croissante et de classe C^1 . Donc G^{-1} existe. En utilisant le changement de variable $u = G(t)$ de classe C^1 , il vient

$$\int_0^A G(t)^n dt = \int_1^{G(A)} u^n d(G^{-1})$$

Une intégration par parties donne

$$\int_1^{G(A)} u^n d(G^{-1}) = [u^n G^{-1}(u)]_1^{G(A)} - \int_1^{G(A)} nu^{n-1}G^{-1}(u) du$$

On a $G^{-1}(1) = 0$ et $\lim_{A \rightarrow +\infty} G(A) = 0$. Ainsi si $Z = \inf_{1 \leq i \leq n} X_i$, $E(Z)$ existe et comme

$$AG(A)^n = AP(Z > t) = A \int_A^{+\infty} f_Z(t) dt \leq \int_A^{+\infty} t f_Z(t) dt \xrightarrow{A \rightarrow +\infty} 0$$

On a finalement

$$E(Z) = \int_0^{+\infty} G(t)^n dt = \int_0^1 nu^{n-1} G^{-1}(u) du$$

4. Si F_1 et F_2 représentent les fonctions de répartition de X_1 et Y_1 , la question précédente donne : pour tout $n \geq 1$

$$\int_0^1 u^{n-1} F_1^{-1}(t) dt = \int_0^1 u^{n-1} F_2^{-1}(t) dt$$

Par linéarité de l'intégration et le théorème admis, $F_1^{-1}(t) = F_2^{-1}(t)$ pour tout $t > 0$ et par bijectivité $F_1(t) = F_2(t)$, pour tout $t > 0$.

Donc X_1 et Y_1 suivent la même loi.

Exercice 3.09.

Les variables aléatoires de cet exercice sont, soit à densité définie et continue sur \mathbb{R} , soit discrètes à valeurs dans \mathbb{Z} . Elles sont définies sur le même espace probabilisé (Ω, \mathcal{A}, P) .

On dit qu'une variable aléatoire X est *symétrique* si pour tout $x \in X(\Omega)$, on a : $P(X \leq x) = P(X \geq -x)$.

1. a) Donner un exemple de variable aléatoire symétrique.

b) Soit X une variable aléatoire suivant la loi de Bernoulli de paramètre $p \in]0, 1[$. Déterminer la loi de $Y = 2X - 1$. Pour quelles valeurs de p , Y est-elle une variable aléatoire symétrique ?

c) Soit X une variable aléatoire symétrique admettant une espérance $E(X)$. Calculer $E(X)$.

2. a) Dans cette question, X et Y sont deux variables aléatoires symétriques et indépendantes. Montrer que $X + Y$ est une variable aléatoire symétrique. On ne fera la démonstration que dans le cas où X et Y sont à densité et on l'admet dans le cas discret.

b) Montrer que si X_1, X_2, \dots, X_n sont des variables aléatoires symétriques indépendantes, alors leur somme $S_n = \sum_{k=1}^n X_k$ est symétrique.

3. Soit X_1, X_2, \dots, X_n des variables aléatoires symétriques et indépendantes de même loi.

Pour tout $k \in \llbracket 1, n \rrbracket$, on pose $S_k = \sum_{j=1}^k X_j$. Soit $x \geq 0$ fixé. On définit une suite d'événements

$(\Omega_k)_{k \geq 1}$ par :

$$\Omega_1 = [X_1 > x], \text{ et pour } k \geq 2, \Omega_k = \left[\sup_{1 \leq j \leq k-1} S_j \leq x \right] \cap [S_k > x]$$

a) Soit $k \in \llbracket 1, n \rrbracket$. Montrer que $[S_n - S_k \geq 0] \cap \Omega_k \subseteq [S_n > x] \cap \Omega_k$.

b) Montrer que $P([S_n - S_k \geq 0] \cap \Omega_k) \geq \frac{1}{2} P(\Omega_k)$.

4. a) Montrer que $\bigcup_{k=1}^n \Omega_k = \left[\sup_{1 \leq j \leq n} S_j > x \right]$.

b) En déduire l'inégalité suivante :

$$P\left(\sup_{1 \leq j \leq n} S_j > x \right) \leq 2P(S_n > x)$$

Solution :

1. a) Un exemple est X suivant une loi normale centrée ou la variable constante nulle !!.

b) On a $Y(\Omega) = \{-1, 1\}$ et $P(Y = 1) = P(X = 1) = p$ et

$$P(Y = -1) = P(X = 0) = 1 - p.$$

Cette loi est symétrique si et seulement si $p = 1/2$, puisque Y ne prend que deux valeurs.

c) Si l'espérance existe, par convergence et symétrie : $E(X) = 0$

2. a) Soit U une variable à densité symétrique.

$$\text{On a pour tout } x \text{ réel } F(x) = P(U \leq x) = P(U \geq -x) = 1 - F(-x).$$

Par dérivation, il vient $F'_U(x) = F'_U(-x)$, ce qui montre que l'on peut choisir une densité f_U paire.

Réciproquement si f_U est une fonction paire, le changement de variable $t \rightarrow -t$ donne :

$$\int_{-\infty}^x f_U(t)dt = \int_{-x}^{+\infty} f_U(t)dt$$

et U est symétrique.

Par indépendance, une densité de $X + Y$ est

$$\begin{aligned} f_{X+Y}(x) &= \int_{\mathbb{R}} f_X(x-t)f_Y(t)dt = \int_{\mathbb{R}} f_X(-x+t)f_Y(-t)dt = \int_{\mathbb{R}} f_X(-x-t)f_Y(t)dt \\ &= f_{X+Y}(-x) \end{aligned}$$

Donc $X + Y$ est symétrique.

b) Par récurrence sur n et la question précédente.

3. a) Pour $k = 1$, $[S_n - S_1 \geq 0] \cap \Omega_1 = [S_n \geq S_1] \cap [X_1 > x] = [S_n > x] \cap \Omega_1$.

Pour $k \geq 2$, $[S_n - S_k \geq 0] \cap \Omega_k = [S_n \geq S_k] \cap [S_k > x] \cap \Omega_k \subseteq [S_n > x] \cap \Omega_k$.

b) En utilisant le système complet d'événements $[S_n - S_k \geq 0] \cap \Omega_k$ et $[S_n - S_k < 0] \cap \Omega_k$, il vient

$$P(\Omega_k) = P([S_n - S_k \geq 0] \cap \Omega_k) + P([S_n - S_k < 0] \cap \Omega_k)$$

puis comme $S_n - S_k = \sum_{i=k+1}^n X_i$ est symétrique.

$$\begin{aligned} P(\Omega_k) &= P([S_n - S_k \geq 0] \cap \Omega_k) + P([S_n - S_k > 0] \cap \Omega_k) \\ &\leq 2P([S_n - S_k \geq 0] \cap \Omega_k) \end{aligned}$$

4. a) Soit $\omega \in \bigcup_{k=1}^n \Omega_k$; il existe k tel que $\omega \in \Omega_k$, et donc $S_k > x$ ce qui entraîne que $\max_{1 \leq j \leq n} S_j > x$.

Réciproquement si $\omega \in \Omega$ vérifie $\max_{1 \leq j \leq n} S_j(\omega) > x$, alors il existe $k \in \llbracket 1, n \rrbracket$, tel que $S_k(\omega) > x$.

On note $k_1, < k_2 < \dots$ la suite croissante d'indices pour lesquels $S_{k_i}(\omega) > x$, alors pour tout $j < k_1$, on a $S_j(\omega) \leq x$ et $S_{k_1}(\omega) > x$; donc $\omega \in \Omega_{k_1}$.

Si cette suite commence en $k_1 = 1$, on a $\omega \in \Omega_1$.

b) On remarque que les (Ω_k) sont deux à deux disjoints.

Ainsi il existe un unique k tel que $(\max_{1 \leq j \leq n} S_j > x) = \Omega_k$.

$$\begin{aligned} P\left(\max_{1 \leq j \leq n} S_j > x\right) &= P(\Omega_k) \leq 2P([S_n - S_k \geq 0] \cap \Omega_k) \leq 2P([S_n > x] \cap \Omega_k) \\ &\leq 2P(S_n > x) \end{aligned}$$

Exercice 3.10.

Partie A

Soit X une variable aléatoire définie sur l'espace probabilisé (Ω, \mathcal{A}, P) de loi uniforme sur $[-\frac{1}{2}, \frac{1}{2}]$.

- Déterminer la fonction de répartition de X .
- Prouver que X et $-X$ ont même loi.
- Rappeler les valeurs de l'espérance et de la variance de X .
- Démontrer que pour tout réel $\lambda > 0$, $e^{\lambda X}$ admet une espérance et calculer sa valeur.

Partie B

Dans la suite de l'exercice, $(X_i)_{i \geq 1}$ est une suite de variables aléatoires définies sur (Ω, \mathcal{A}, P) mutuellement indépendantes et toutes de même loi que X . Pour $n \in \mathbb{N}^*$, on pose $S_n = \sum_{k=1}^n X_k$.

- Démontrer que S_n et S_n^2 admettent une espérance et calculer $E(S_n)$ et $E(S_n^2)$.
- Soit $f : \mathbb{R} \mapsto \mathbb{R}$ l'application définie par $f(x) = \frac{e^x - e^{-x}}{2}$. Prouver que pour tout réel $\lambda > 0$, on a :

$$E(e^{\lambda S_n}) = \left(\frac{f(\frac{\lambda}{2})}{\frac{\lambda}{2}}\right)^n$$

- Prouver que pour tout réel $u > 0$, on a : $\frac{f(u)}{u} \leq \exp\left(\frac{u^2}{6}\right)$ (on pourra commencer par démontrer que pour tout $k \in \mathbb{N}$, on a $(2k+1)!6^k k!$).
- Démontrer que pour tous réels t et $\lambda > 0$, on a : $P(S_n \geq t) \leq e^{-\lambda t + n \frac{\lambda^2}{24}}$.
- En déduire que pour tout réel $t > 0$, on a : $P(S_n \geq t) \leq e^{-\frac{6t^2}{n}}$.

6. Soit Z une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0, 1)$, exploiter la question précédente et le théorème limite central pour établir que, pour tout réel $t > 0$:

$$P(Z \geq t) \leq e^{-\frac{t^2}{2}}$$

Solution :

Partie A.

1. Si F est la fonction de répartition de X :

$$F(x) = \begin{cases} 0 & \text{si } x \leq -1/2 \\ \frac{1}{2} + x & \text{si } -1/2 \leq x \leq 1/2. \\ 1 & \text{si } x \geq 1/2 \end{cases}$$

2. Clair !

3. On a $E(X) = 0$ et $\sigma^2(X) = \frac{1}{12}$

4. Le théorème du transfert assure que

$$E(e^{\lambda X}) = \int_{-\infty}^{\infty} e^{\lambda t} f_X(t) dt = \int_{-1/2}^{1/2} e^{\lambda t} dt = \frac{1}{\lambda} (e^{\lambda/2} - e^{-\lambda/2}).$$

Partie B.

1. Immédiatement $E(S_n) = E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i) = 0$ par linéarité.

De plus $S_n^2 = (\sum_{i=1}^n X_i)^2 = \sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j$. Par indépendance des $(X_i)_{i \geq 1}$, on a :

$E(X_i X_j) = E(X_i)E(X_j) = 0$, d'où :

$$E(S_n^2) = E(\sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j) = \sum_{i=1}^n E(X_i^2) = \frac{n}{12}$$

2. Comme $S_n = \sum_{i=1}^n X_i$, on a $e^{\lambda S_n} = e^{\lambda \sum_{i=1}^n X_i} = \prod_{i=1}^n e^{\lambda X_i}$. Pour tout entier n non nul, l'indépendance mutuelle des $(X_i)_{1 \leq i \leq n}$, assure que les variables $(e^{\lambda X_i})_{1 \leq i \leq n}$ sont elles-mêmes mutuellement indépendantes et donc :

$$\begin{aligned} E(e^{\lambda S_n}) &= E\left(\prod_{i=1}^n e^{\lambda X_i}\right) = \prod_{i=1}^n E(e^{\lambda X_i}) = \prod_{i=1}^n \frac{1}{\lambda} (e^{\lambda/2} - e^{-\lambda/2}) = \prod_{i=1}^n \frac{1}{\lambda} 2f(\lambda/2) \\ &= \prod_{i=1}^n \left(\frac{f(\lambda/2)}{\lambda/2}\right) \end{aligned}$$

D'où : $E[e^{\lambda S_n}] = \left(\frac{f(\lambda/2)}{\lambda/2}\right)^n$.

3. Pour tout réel u :

$$f(u) = \frac{1}{2} \left(\sum_{i=0}^{\infty} \frac{u^i}{i!} - \sum_{i=0}^{\infty} \frac{(-1)^i u^i}{i!} \right) = \frac{1}{2} \left(\sum_{k=0}^{\infty} 2 \frac{u^{2k+1}}{(2k+1)!} \right) = \sum_{k=0}^{\infty} \frac{u^{2k+1}}{(2k+1)!}.$$

D'où pour tout réel $u > 0$: $\frac{f(u)}{u} = \sum_{k=0}^{\infty} \frac{u^{2k}}{(2k+1)!}$.

D'autre part : $e^{\frac{u^2}{6}} = \sum_{k=0}^{\infty} \frac{u^{2k}}{6^k k!}$.

Pour tout entier k , on considère la propriété : $\mathcal{P}(k) : (2k+1)! \geq 6^k k!$.

→ la relation $\mathcal{P}(0)$ est vérifiée.

→ Supposons $\mathcal{P}(k)$ vraie pour un certain rang k , on a

$$\begin{aligned} (2(k+1)+1)! &= (2k+3)(2k+2)(2k+1)! = 2(2k+3)(k+1)(2k+1)! \\ &\geq 6(k+1)(2k+1)! \geq 6(k+1)6^k k! \end{aligned}$$

Donc $(2(k+1)+1)! \geq 6^{k+1}(k+1)!$ et $\mathcal{P}(k+1)$ est vraie.

Ainsi $\mathcal{P}(k)$ est vraie pour tout entier k , en conséquence, pour tout k : $\frac{u^{2k}}{(2k+1)!} \leq \frac{u^{2k}}{6^k k!}$, ce qui établit que : $\forall u > 0, \frac{f(u)}{u} \leq e^{\frac{u^2}{6}}$ (*).

4. Pour tous réels t et $\lambda > 0$: $S_n \geq t \iff e^{\lambda S_n} \geq e^{\lambda t}$, d'où $P(S_n \geq t) = P(e^{\lambda S_n} \geq e^{\lambda t})$.

Comme $e^{\lambda S_n}$ est positive et admet une espérance, par l'inégalité de Markov, on a :

$$P(e^{\lambda S_n} \geq e^{\lambda t}) \leq \frac{E(e^{\lambda S_n})}{e^{\lambda t}} \text{ puis (1) entraîne } P(e^{\lambda S_n} \geq e^{\lambda t}) \leq \frac{\left(\frac{f(\lambda/2)}{\lambda/2}\right)^n}{e^{\lambda t}}$$

De (*) on déduit : $P(S_n \geq t) \leq e^{-\lambda t + n \frac{\lambda^2}{24}}$.

5. Pour t et n fixés, le polynôme $q : \lambda \mapsto -\lambda t + n \frac{\lambda^2}{24}$ est minimum pour $\lambda = \frac{12t}{n}$ et son minimum est $q\left(\frac{12t}{n}\right) = -\frac{6t^2}{n}$. On déduit pour $\lambda = \frac{12t}{n}$ que $P(S_n \geq t) \leq e^{-\frac{6t^2}{n}}$.

6. Pour tout entier i , $V(X_i) = \sigma^2 = \frac{1}{12}$, le théorème central limite s'applique.

Pour tout $t \geq 0$, $\lim_{n \rightarrow \infty} P\left(\frac{S_n}{\sqrt{n/12}} \geq t\right) = P(Z \geq t)$, où Z suit la loi $\mathcal{N}(0, 1)$.

Or $P(S_n \geq \sqrt{n} \sqrt{\frac{1}{12}} t) \leq e^{-\frac{6(\sqrt{n} \sqrt{\frac{1}{12}} t)^2}{n}}$, d'où $P\left(\frac{S_n}{\sqrt{n/12}} \geq t\right) \leq e^{-\frac{t^2}{2}}$ et il n'y a plus qu'à faire $n \rightarrow \infty$.

Exercice 3.11.

Toutes les variables aléatoires considérées dans cet exercice sont définies sur un espace probabilisé (Ω, \mathcal{A}, P) .

Soient λ et p deux réels tels que $\lambda > 0$ et $0 < p < 1$.

On considère le couple de variables aléatoires (X, Y) à valeurs dans \mathbb{N}^2 , de loi définie par :

$$\forall (k, n) \in \mathbb{N}^2, P(X = n, Y = k) = \begin{cases} \frac{e^{-\lambda} \lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} & \text{si } 0 \leq k \leq n \\ 0 & \text{sinon.} \end{cases}$$

1. Vérifier que la relation ci-dessus définit bien une loi de probabilité sur \mathbb{N}^2 .

2. Déterminer la loi marginale de la variable aléatoire X , puis celle de la variable aléatoire Y . Les variables aléatoires X et Y sont-elles indépendantes ?

3. Déterminer la loi conditionnelle de la variable aléatoire Y , sachant que $(X = n)$ est réalisé.
4. Soit Z la variable aléatoire définie par $Z = X - Y$. Déterminer la loi de la variable aléatoire Z .
5. Les variables aléatoires Y et Z sont-elles indépendantes ?

Solution :

Remarquons que la loi du couple (X, Y) peut s'écrire :

$$P[\{X = n, Y = k\}] = \frac{e^{-\lambda} \lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} \mathbf{1}_{\{0 \leq k \leq n\}}$$

1. Pour tout $(k, n) \in \mathbb{N}^2$, $P[\{X = n, Y = k\}] \geq 0$. Il suffit de vérifier que $\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} P[X = n, Y = k] = 1$.

On écrit :

$$\begin{aligned} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} P[X = n, Y = k] &= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} \mathbf{1}_{\{0 \leq k \leq n\}} \\ &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{k=0}^n \frac{n! p^k (1-p)^{n-k}}{k!(n-k)!} \\ &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \\ &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} (1 + (1-p))^n \\ \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} P[X = n, Y = k] &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1 \end{aligned}$$

2. Par définition, la loi marginale de la variable aléatoire X est donnée par, pour tout $n \in \mathbb{N}$:

$$P[X = n] = \sum_{k=0}^{\infty} P[X = n, Y = k] = \sum_{k=0}^n \frac{e^{-\lambda} \lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} = e^{-\lambda} \frac{\lambda^n}{n!}$$

Ainsi X suit la loi de Poisson de paramètre $\lambda > 0$.

Par définition, la loi marginale de la variable aléatoire Y est donnée par, pour tout $k \in \mathbb{N}$:

$$\begin{aligned} P[Y = k] &= \sum_{n=0}^{\infty} P[X = n, Y = k] = \frac{e^{-\lambda} (\lambda p)^k}{k!} \sum_{n=k}^{\infty} \frac{[\lambda(1-p)]^{n-k}}{(n-k)!} \\ &= \frac{e^{-\lambda} (\lambda p)^k}{k!} \sum_{n=0}^{\infty} \frac{[\lambda(1-p)]^n}{n!} \end{aligned}$$

$$P[Y = k] = \frac{e^{-\lambda} (\lambda p)^k}{k!} e^{\lambda(1-p)} = e^{-\lambda p} \frac{(\lambda p)^k}{k!} :$$

Y suit la loi de Poisson de paramètre $\lambda p > 0$.

On a : $P[X = 0] = e^{-\lambda}$, $P[Y = 0] = e^{-\lambda p}$, $P[X = 0, Y = 0] = e^{-\lambda}$, et $e^{-\lambda} e^{-\lambda p} \neq e^{-\lambda}$.

D'où on déduit que les variables aléatoires X et Y ne sont pas indépendantes.

3. La loi conditionnelle de la variable aléatoire Y sachant $\{X = n\}$ est bien définie et est donnée par, pour tout $k \in \mathbb{N}$:

$$P_{(X=n)}[Y = k] = \frac{P[X = n, Y = k]}{P[X = n]} = \frac{e^{-\lambda} \lambda^n p^k (1-p)^{n-k}}{k!(n-k)!} \mathbf{1}_{\{0 \leq k \leq n\}} \frac{n!}{e^{-\lambda} \lambda^n}$$

$$P_{(X=n)}[Y = k] = \binom{n}{k} p^k (1-p)^{n-k} \mathbf{1}_{\{0 \leq k \leq n\}}$$

Ainsi, la loi conditionnelle de la variable aléatoire Y sachant $\{X = n\}$, est la loi binomiale de paramètres n et p .

4. D'après la formule des probabilités totales on a, pour tout $n \in \mathbb{N}$:

$$P(Z = n) = \sum_{k=0}^{\infty} P((X = n+k) \cap (Y = k)) = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^{n+k} p^k (1-p)^n}{k!n!}$$

$$= e^{-\lambda} \frac{[\lambda(1-p)]^n}{n!} \sum_{k=0}^{\infty} \frac{(\lambda p)^k}{k!}$$

$$P(Z = n) = e^{-\lambda(1-p)} \frac{[\lambda(1-p)]^n}{n!}$$

On voit que la variable aléatoire Z suit la loi de Poisson de paramètre $\lambda(1-p) > 0$.

5. Nous avons, pour tout $(k, n) \in \mathbb{N}^2$:

$$\star P[Y = k]P[Z = n] = e^{-\lambda p} \frac{(\lambda p)^k}{k!} e^{-\lambda(1-p)} \frac{[\lambda(1-p)]^n}{n!} = \frac{e^{-\lambda} \lambda^{k+n} p^k (1-p)^n}{k!n!}$$

$$\star P[Y = k, Z = n] = P[Y = k, X - Y = n] = P[X = n+k, Y = k]$$

$$= \frac{e^{-\lambda} \lambda^{n+k} p^k (1-p)^n}{k!n!},$$

On en déduit que les variables aléatoires Y et Z sont indépendantes.

Exercice 3.12.

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes définies sur le même espace probabilisé (Ω, \mathcal{A}, P) et suivant toutes la loi de Poisson de paramètre $\theta > 0$ inconnu.

Pour $n \geq 2$, on pose $M_n = \frac{1}{n} \sum_{k=1}^n X_k$.

1. Montrer que pour tout $c > 0$, la variable aléatoire e^{-cM_n} admet une espérance, et la calculer.

2. Pour $n \geq 2$, soit $c_n = n \ln \left(\frac{n}{n-1} \right)$. On choisit T_n comme estimateur de $e^{-\theta}$, où T_n est défini par :

$$T_n(X_1, \dots, X_n) = e^{-c_n M_n}.$$

a) Montrer que T_n est un estimateur sans biais de $e^{-\theta}$.

b) Calculer la variance de T_n .

c) Montrer que, pour tout $\theta > 0$, $(T_n)_n$ converge en probabilité vers $e^{-\theta}$.

Solution :

1. La variable aléatoire M_n étant à valeurs dans \mathbb{N} et la constante c étant positive, on déduit que : $0 \leq e^{-cM_n} \leq 1$. Ainsi, la variable aléatoire e^{-cM_n} est bornée et admet donc une espérance.

$$E[e^{-cM_n}] = E\left[e^{-\frac{c}{n} \sum_{k=1}^n X_k}\right] = E\left[\prod_{k=1}^n e^{-\frac{c}{n} X_k}\right] = \prod_{k=1}^n E[e^{-\frac{c}{n} X_k}] = \left(E[e^{-\frac{c}{n} X_1}]\right)^n$$

D'après le théorème de transfert :

$$E[e^{-\frac{c}{n} X_1}] = e^{-\theta} \sum_{k=0}^{\infty} e^{-\frac{c}{n} k} \frac{\theta^k}{k!} = e^{-\theta} \sum_{k=0}^{\infty} \frac{(e^{-\frac{c}{n}} \theta)^k}{k!} = e^{-\theta} e^{\theta e^{-\frac{c}{n}}} = e^{-\theta(1-e^{-\frac{c}{n}})}.$$

On conclut que : $E[e^{-cM_n}] = e^{-n\theta(1-e^{-\frac{c}{n}})}$.

2. a) D'après la question précédente, nous savons que cette espérance existe et vaut :

$$E[T_n(X_1, \dots, X_n)] = E[e^{-c_n M_n}] = e^{-n\theta(1-e^{-\frac{c_n}{n}})}$$

De plus, $e^{-\frac{c_n}{n}} = e^{-\ln\left(\frac{n-1}{n}\right)} = e^{\ln\left(\frac{n-1}{n}\right)} = \frac{n-1}{n}$. Ainsi, $1 - e^{-\frac{c_n}{n}} = \frac{1}{n}$ et on conclut :

$$E[T_n(X_1, \dots, X_n)] = e^{-n\theta \frac{1}{n}} = e^{-\theta}.$$

Par définition, cela signifie que T_n est un estimateur sans biais de $e^{-\theta}$.

b) D'après la question 1, T_n admet un moment d'ordre 2 et par définition

$$V[T_n(X_1, \dots, X_n)] = E[e^{-2c_n M_n}] - E[e^{-c_n M_n}]^2 = E[e^{-2c_n M_n}] - e^{-2\theta},$$

d'après la question 2 a)

D'après la question 1, $E[e^{-2c_n M_n}] = e^{-n\theta(1-e^{-\frac{2c_n}{n}})}$.

De plus, $e^{-\frac{2c_n}{n}} = \left(\frac{n-1}{n}\right)^2$ et $1 - e^{-\frac{2c_n}{n}} = \frac{2n-1}{n^2}$. Ainsi :

$$E[e^{-2c_n M_n}] = e^{-\theta \frac{2n-1}{n}} = e^{-2\theta} e^{\frac{\theta}{n}}$$

et on conclut que :

$$V[T_n(X_1, \dots, X_n)] = e^{-2\theta} e^{\frac{\theta}{n}} - e^{-2\theta} = e^{-2\theta} \left(e^{\frac{\theta}{n}} - 1\right).$$

c) D'après la question 1, la variable aléatoire $e^{-c_n M_n}$ a un moment d'ordre 2, on peut donc appliquer l'inégalité de Bienaymé-Tchebychev.

Ainsi, pour tout $\varepsilon > 0$, on obtient en utilisant les questions 2 a) et b) :

$$P[|e^{-c_n M_n} - e^{-\theta}| > \varepsilon] \leq \frac{e^{-2\theta} (e^{\frac{\theta}{n}} - 1)}{\varepsilon^2}.$$

Étant donné que $e^{\frac{\theta}{n}}$ admet 1 comme limite quand n tend vers l'infini. On déduit de la question 2. c que, pour tout $\theta > 0$, pour tout $\varepsilon > 0$:

$$\lim_{n \rightarrow \infty} P[|T_n(X_1, \dots, X_n) - e^{-\theta}| > \varepsilon] = 0,$$

ce qui signifie que pour tout $\theta > 0$, $(T_n(X_1, \dots, X_n))_n$ converge en probabilité vers $e^{-\theta}$.

Exercice 3.13.

On considère une suite de variables aléatoires $(\Delta_n)_{n \in \mathbb{N}^*}$ indépendantes et de même loi exponentielle de paramètre 1, définies sur un espace probabilisé (Ω, \mathcal{F}, P) .

On pose $T_0 = 0$ et pour tout entier naturel n non nul, on note T_n la variable aléatoire définie par :

$$T_n = \sum_{i=1}^n \Delta_i$$

1. Pour tout entier naturel n , déterminer l'espérance et la variance de la variable aléatoire T_n .

2. Soit t un réel positif ou nul.

a) Pour tout entier naturel n strictement supérieur à t , justifier la relation :

$$(T_n < t) \subset (|T_n - n| \geq n - t)$$

b) En déduire la valeur de $\lim_{n \rightarrow \infty} P(T_n < t)$.

c) Montrer que l'événement $\bigcap_{k=1}^{\infty} (T_k < t)$ est de probabilité nulle.

3. Soit t un réel positif ou nul. Étant donné un élément ω de Ω , on note $N(t)(\omega)$ le plus grand élément de l'ensemble $\{n \in \mathbb{N} / T_n(\omega) \leq t\}$ (qui contient 0) si cet ensemble est fini, et 0 sinon.

Montrer que l'application $N(t)$ est une variable aléatoire réelle vérifiant :

$$P(N(t) = 0) = P(T_1 > t).$$

4. a) Pour tout entier naturel n non nul, reconnaître la loi de la variable aléatoire T_n .

b) Montrer que pour tout réel strictement positif t et pour tout entier naturel n on a :

$$\sum_{k=0}^n \frac{t^k e^{-t}}{k!} + e^{-t} \int_0^t \frac{(t-u)^n}{n!} e^u du = 1$$

c) En remarquant que pour tout entier naturel n , $P(N(t) \leq n) = P(T_{n+1} > t)$, en déduire l'égalité :

$$P(N(t) \leq n) = \sum_{k=0}^n \frac{t^k e^{-t}}{k!}$$

d) Pour tout réel t positif ou nul, reconnaître la loi de la variable aléatoire $N(t)$.

Solution :

1. La variable aléatoire T_n admet une espérance et une variance comme somme finie de variables aléatoires admettant une espérance et une variance.

$$E[T_n] = E\left[\sum_{k=1}^n \Delta_k\right] = \sum_{k=1}^n E[\Delta_k] = nE[\Delta_1] = n, \quad V[T_n] = V\left[\sum_{k=1}^n \Delta_k\right] = \sum_{k=1}^n V[\Delta_k]$$

$$\text{Soit } : E(T_n) = V(T_n) = \sum_{k=1}^n 1 = n$$

car les variables aléatoires $\Delta_1, \dots, \Delta_n$ sont indépendantes et car $E(\Delta_1) = V[\Delta_1] = 1$.

2. a) $(T_n < t) \subset (|T_n - n| < n - t)$ (le deuxième événement est $(T_n < t) \cup (T_n > 2n - t)$)

b) La variable aléatoire a une variance car somme finie de variables aléatoires ayant une variance. Ainsi, d'après l'inégalité de Bienaymé-Tchebychev,

$$\forall \varepsilon > 0, P[|T_n - E[T_n]| \geq \varepsilon] \leq \frac{V[T_n]}{\varepsilon^2} = \frac{n}{\varepsilon^2}$$

Pour $\varepsilon = n - t > 0$, et en utilisant la question 2 a), on obtient que pour tout $n \in \mathbb{N}$ tel que $n > t$:

$$0 \leq P\{T_n < t\} \leq P\{|T_n - n| \geq n - t\} \leq \frac{n}{(n - t)^2}.$$

Comme $\frac{n}{(n - t)^2}$ tend vers 0 quand n tend vers l'infini, on en déduit par encadrement que :

$$\lim_{n \rightarrow \infty} P\{T_n < t\} = 0.$$

c) Pour tout $k \in \mathbb{N}^*$, posons $A_k = \{T_k < t\}$. Alors, étant donné que les variables aléatoires Δ_k sont positives, la suite $(A_k)_{k \in \mathbb{N}}$ est une suite décroissante d'événements, ainsi d'après un théorème du cours :

$$P\left[\bigcap_{k=1}^{\infty} A_k\right] = \lim_{n \rightarrow \infty} P[A_n] = 0.$$

3. D'après l'énoncé :

$$\begin{aligned} P\{N(t) = 0\} &= P[(T_1 > t) \cup \{\bigcap_{k=1}^{\infty} (T_k < t)\}] = P[T_1 > t] + P\left[\bigcap_{k=1}^{\infty} (T_k < t)\right] \\ &= P\{T_1 > t\}, \end{aligned}$$

4. a) La variable aléatoire T_n , qui est somme de n variables aléatoires indépendantes de loi exponentielle de paramètre 1, suit une loi gamma de paramètres n et 1.

b) Pour tout réel t , $e^{-t}e^t = 1$. Le résultat est obtenu en remarquant que l'application \exp est de classe C^∞ sur \mathbb{R} , et en appliquant la formule de Taylor avec reste intégral à l'ordre $n + 1$ en 0.

c) D'après l'énoncé, $P\{N(t) \leq n\} = P\{T_{n+1} > t\}$. D'après la question 4 a), la variable aléatoire T_{n+1} suit une loi gamma de paramètres $n + 1$ et 1.

$$P[N(t) \leq n] = P[T_{n+1} > t] = 1 - P[T_{n+1} \leq t] = 1 - \int_0^t \frac{x^n}{n!} e^{-x} dx$$

En effectuant le changement de variables $x = t - u$ qui est C^1 , nous obtenons :

$$P[N(t) \leq n] = 1 - e^{-t} \int_0^t \frac{(t - u)^n}{n!} e^u du = \sum_{k=0}^n \frac{t^k e^{-t}}{k!}$$

d) Pour tout réel t strictement positif et pour tout entier naturel n , nous avons :

$$P[N(t) = n] = P[N(t) \leq n] - P[N(t) \leq n - 1] = \frac{t^n e^{-t}}{n!}$$

Ainsi, on conclut que pour tout réel t strictement positif, la variable aléatoire $N(t)$ suit la loi de Poisson de paramètre t .

Si $t = 0$, alors $P[N(0) = 0] = P[T_1 > 0] = 1$. Ainsi, $N(t)$ est la variable aléatoire constante égale à 0 (qui est aussi, si on ose, une variable aléatoire suivant la loi de Poisson de paramètre 0).

Exercice 3.14.

Soit (Ω, \mathcal{F}, P) un espace probabilisé et $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur cet espace, mutuellement indépendantes de même loi ayant pour densité la fonction :

$$f : t \mapsto \frac{2t}{\theta^2} \mathbf{1}_{[0, \theta]}(t)$$

où θ est un paramètre réel strictement positif.

Pour tout $n \in \mathbb{N}^*$, on pose $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$

1. a) Calculer $E(X_1)$ et en déduire que pour tout $n > 0$, la variable aléatoire $T_n = \frac{3}{2} \bar{X}_n$ est un estimateur sans biais du paramètre θ .

b) Calculer son risque quadratique $r_{T_n}(\theta)$ et étudier la convergence en probabilités de la suite d'estimateurs $(T_n)_{n > 0}$.

2. Appliquer le théorème limite central à (\bar{X}_n) . En déduire un intervalle de confiance au niveau 95% pour θ , basé sur l'estimateur \bar{X}_n . (Si on note Φ la fonction de répartition d'une variable aléatoire qui suit une loi $\mathcal{N}(0, 1)$, on donne, pour $t = 1,96$: $2\Phi(t) - 1 = 0,95$)

3. Soit la variable aléatoire : $M_n = \sup\{X_1, \dots, X_n\}$.

a) Calculer la fonction de répartition G de M_n .

b) Soit $\delta \geq 0$. Étudier la convergence de la série de terme général $P(|M_n - \theta| > \delta)$.

c) Calculer $E(M_n)$ et étudier les propriétés de la variable aléatoire $M'_n = \frac{2n+1}{2n} M_n$ en tant qu'estimateur du paramètre θ .

Solution :

1. a) On a $E(X) = \int_0^\theta t \times \frac{2t}{\theta^2} dt = \frac{2}{3} \theta$.

Pour $n > 0$, $E(T_n) = \frac{3}{2} E(\bar{X}) = \theta$ donc T_n est un estimateur sans biais du paramètre θ .

b) On a $r_{T_n}(\theta) = V(T_n) + (E(T_n) - \theta)^2 = V(T_n) + 0 = V(\frac{3}{2} \bar{X}) = \frac{9}{4} \times \frac{1}{n^2} \times n V(X_1)$ (car les X_i sont indépendants).

Or $V(X_1) = \int_0^\theta t^2 \times \frac{2t}{\theta^2} dt - (\frac{2}{3} \theta)^2 = \frac{\theta^2}{18n}$ d'où : $r_{T_n}(\theta) = \frac{\theta^2}{8n}$.

L'inégalité de Bienaymé Tchebychev permet de conclure à la convergence de la suite d'estimateurs $(T_n)_n$

2. On applique le théorème limite central à la suite de variables aléatoires $(X_n)_n$: la suite de variables aléatoires réelles : $Z_n = \sqrt{n} \frac{\bar{X} - \frac{2}{3} \theta}{\frac{\theta}{\sqrt{18}}} = \frac{\sqrt{18n}}{\theta} \bar{X}_n - 2\sqrt{2n}$ converge en loi vers une variable aléatoire Z de loi $\mathcal{N}(0, 1)$.

Ainsi $P(-t \leq Z_n \leq t) = 2\Phi(t) - 1 \geq 0.95 \iff \Phi(t) \geq 0.975 \iff t \geq 1.96$

soit

$$-t \leq Z_n \leq t \iff \sqrt{2n} - t \leq \frac{\sqrt{18n}}{\theta} \bar{X}_n \leq \sqrt{2n} + t$$

$$\iff \frac{3\sqrt{2n}\bar{X}_n}{2\sqrt{2n} + 1.96} \leq \theta \leq \frac{3\sqrt{2n}\bar{X}_n}{2\sqrt{2n} - 1.96}.$$

D'où l'intervalle de confiance cherché : $\left[\frac{3\sqrt{2n}\bar{X}_n}{2\sqrt{2n} + 1.96}, \frac{3\sqrt{2n}\bar{X}_n}{2\sqrt{2n} - 1.96} \right]$.

3. a) Soit un réel x , on a $P(M_n \leq x) = \prod_{i=1}^n P(X \leq x)$ (intersection d'événements indépendants

et de même probabilité) et $P(X \leq x) = \int_{-\infty}^x f(t)dt = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x^2}{\theta^2} & \text{si } 0 \leq x \leq \theta \\ 1 & \text{si } x \geq \theta \end{cases}$

D'où la fonction de répartition : $G(x) = P(M_n \leq x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \left(\frac{x}{\theta}\right)^{2n} & \text{si } 0 \leq x \leq \theta \\ 1 & \text{si } x \geq \theta \end{cases}.$

b) * Si $0 < \delta \leq \theta$. Alors :

$$P(|M_n - \theta| > \delta) = P(M_n < \theta - \delta) + 1 - P_\theta(M_n < \theta + \delta)$$

$$= G(\theta - \delta) + 1 - G(\theta + \delta) = \left(\frac{\theta - \delta}{\theta}\right)^{2n} + 1 - 1$$

$$= \left(\frac{\theta - \delta}{\theta}\right)^{2n}$$

On reconnaît une série géométrique de raison < 1 donc convergente.

* Si $\theta < \delta$, tous les termes de la série sont nuls.

c) Une densité de M_n est $g(x) = \frac{2n}{\theta^{2n}} x^{2n-1} \mathbf{1}_{[0,\theta]}(x)$.

On a : $E(M_n) = \int_0^\theta \frac{2n}{\theta^{2n}} x^{2n} dx = \frac{2n}{2n+1} \theta,$

M_n est un estimateur biaisé de θ ce défaut est corrigé avec $M'_n = \frac{2n+1}{2n} M_n,$

On a : $r_{M'_n}(\theta) = V_\theta(M'_n) + (E_\theta(M'_n) - \theta)^2 = V_\theta(M'_n) = \left(\frac{2n+1}{2n}\right)^2 V_\theta(M_n)$

Or $V(M_n) = E_\theta(M_n^2) - \left(\frac{2n}{2n+1}\right)^2 \theta^2 = \int_0^\theta \frac{2n}{\theta^{2n}} x^{2n+1} dx - \left(\frac{2n}{2n+1}\right)^2 \theta^2$

$$= \left(\frac{n}{n+1}\right) \theta^2 - \left(\frac{2n}{2n+1}\right)^2 \theta^2 = \frac{\theta^2}{(n+1)(2n+1)^2}$$

d'où $r_{M'_n}(\theta) = \frac{\theta^2}{4n^2(n+1)} < r_{T_n}(\theta).$

Cet estimateur de θ est donc de meilleure qualité que T_n .

Exercice 3.15.

On considère un jeu de roulette avec n issues possibles (les numéros allant de 1 à n). Une partie est constituée d'exactly n lancers successifs de la boule. Pour une partie donnée, on note :

- A_n l'événement : chaque numéro sort exactement une fois,
- X_i la variable aléatoire égale au nombre d'apparitions du numéro i ,
- $S_n = \sum_{i=1}^n \mathbf{1}_{(X_i=0)}$

1. Calculer la probabilité de l'événement A_n .

2. Donner un équivalent de $P(A_n)$ lorsque n tend vers l'infini. (On pourra utiliser la formule de Stirling $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$).

Dans toutes les questions suivantes, le nombre n est fixé.

3. a) Quelle est la loi de X_i ?

b) Pour $j, k \in \mathbb{N}$, calculer la probabilité conditionnelle $P_{(X_1+\dots+X_{n-1}=j)}(X_n = k)$. Les variables aléatoires X_i , $1 \leq i \leq n$ sont-elles indépendantes ?

4. Que représente $\frac{S_n}{n}$? Calculer $E\left(\frac{S_n}{n}\right)$ puis en donner un équivalent pour $n \rightarrow +\infty$.

5. On note S'_n le nombre de numéros sortis exactement une fois lors d'une partie et S''_n le nombre de numéros sortis au moins deux fois lors d'une partie. Calculer $E\left(\frac{S'_n}{n}\right)$ et $E\left(\frac{S''_n}{n}\right)$.

6. a) On effectue une suite de parties toutes de même nombre de lancers n fixé. Pour tout $j \in \mathbb{N}^*$, on note T_j, T'_j, T''_j respectivement les proportions de numéros sortis 0, 1 ou au moins 2 fois au cours des n lancers de la j ième partie. Montrer que les trois variables aléatoires $\frac{1}{N}(T_1 + \dots + T_N)$, $\frac{1}{N}(T'_1 + \dots + T'_N)$ et $\frac{1}{N}(T''_1 + \dots + T''_N)$ convergent en probabilité.

b) Pour une roulette standard : $n = 37$ numéros, on donne $\left(\frac{36}{37}\right)^{37} = 0.362$ et $\left(\frac{36}{37}\right)^{36} = 0.372$. Conclure.

Solution :

1. Les résultats des lancers sont indépendants ; chaque nouveau lancer doit faire apparaître un numéro différent : $P(A_n) = \frac{n}{n} \times \frac{n-1}{n} \times \dots \times \frac{1}{n} = \frac{n!}{n^n}$.

2. On a $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ donc $P(A_n) \sim \sqrt{2\pi n} e^{-n}$.

3. a) On a $X_i = \sum_{j=1}^n b_{i,j}$ où $b_{i,j}$, $1 \leq j \leq n$ est une variable aléatoire de Bernoulli qui vaut 1 si le numéro i sort lors du j ème lancer, son paramètre est $\frac{1}{n}$. Il y a répétition de n épreuves indépendantes donc X_i suit la loi binômiale $\mathcal{B}(n, 1/n)$.

b) On remarque : $X_1 + \dots + X_n = n$, d'où pour $j, k \geq 0$,

$$P_{(X_1+\dots+X_{n-1}=j)}(X_n = k) = \begin{cases} 1 & \text{si } k = n - j \\ 0 & \text{sinon} \end{cases}$$

Si les X_i étaient indépendantes alors les variables $X_1 + \dots + X_{n-1}$ et X_n devraient aussi être indépendantes (lemme des coalitions) ce qui contredit la calcul de la probabilité conditionnelle ci-dessus.

Donc les variables aléatoires $X_i, 1 \leq i \leq n$ ne sont pas indépendantes.

4. La fraction $\frac{S_n}{n}$ est la proportion de numéros jamais sortis au cours des n tirages. Par linéarité de l'espérance, et le fait que tous les X_i ont même loi, il vient :

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n} \sum_{i=1}^n P(X_i = 0) = P(X_1 = 0) = \binom{n}{0} \left(\frac{1}{n}\right)^0 \left(1 - \frac{1}{n}\right)^n = \left(1 - \frac{1}{n}\right)^n \sim e^{-1}$$

5. On a

$$\begin{aligned} E\left(\frac{S'_n}{n}\right) &= \frac{1}{n} \sum_{i=1}^n P(X_i = 1) = P(X_1 = 1) = \binom{n}{1} \left(\frac{1}{n}\right) \left(1 - \frac{1}{n}\right)^{n-1} = n \frac{1}{n} \left(1 - \frac{1}{n}\right)^{n-1} \\ &= \left(1 - \frac{1}{n}\right)^{n-1} \end{aligned}$$

et

$$S_n + S'_n + S''_n = n \implies E\left(\frac{S''_n}{n}\right) = 1 - E\left(\frac{S_n}{n}\right) - E\left(\frac{S'_n}{n}\right) = 1 - \left(2 - \frac{1}{n}\right) \left(1 - \frac{1}{n}\right)^{n-1}$$

6. a) Pour n fixé, les variables $T_j, 1 \leq j \leq N$ sont indépendantes de même loi et admettent une espérance $E(T_j) = \left(1 - \frac{1}{n}\right)^n$ et une variance ; la loi faible des grands nombres s'applique :

$$\forall \varepsilon > 0, \lim_{N \rightarrow \infty} P\left[\left|\frac{1}{N}(T_1 + \dots + T_N) - \left(1 - \frac{1}{n}\right)^n\right| > \varepsilon\right] = 0$$

Il en est de même pour T'_j et T''_j , d'où les convergences en probabilité :

$$\begin{aligned} \frac{1}{N}(T_1 + \dots + T_N) &\rightarrow \left(1 - \frac{1}{n}\right)^n = \left(\frac{n-1}{n}\right)^n \\ \frac{1}{N}(T'_1 + \dots + T'_N) &\rightarrow \left(1 - \frac{1}{n}\right)^{n-1} = \left(\frac{n-1}{n}\right)^{n-1} \\ \frac{1}{N}(T''_1 + \dots + T''_N) &\rightarrow 1 - \left(2 - \frac{1}{n}\right) \left(1 - \frac{1}{n}\right)^{n-1} \end{aligned}$$

b) Pour une roulette standard : $n = 37$ numéros, $\left(\frac{36}{37}\right)^{37} = 36,2\%$ des numéros ne sortent pas, $\left(\frac{36}{37}\right)^{36} = 37,2\%$ des numéros sortent 1 seule fois et les autres $26,4\%$ sortent au moins 2 fois.

Exercice 3.16.

1. Soit (a_n) une suite réelle croissante qui converge vers $\ell \in \mathbb{R}$.

Montrer que $\forall n \in \mathbb{N}, a_n \leq \ell$.

2. On pose pour tout entier $n > 0 : a_n = \frac{\sqrt{n}}{4^n} \binom{2n}{n}$.

a) Montrer grâce à la formule de Stirling : $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$ que la suite $(a_n)_{n \in \mathbb{N}^*}$ converge vers un réel à déterminer.

b) Montrer que pour tout entier $n > 0 :$

$$0 \leq a_{n+1} - a_n = a_n \frac{(\sqrt{n+1} - \sqrt{n})^2}{2\sqrt{n}\sqrt{n+1}} \leq \frac{1}{\sqrt{\pi}} \frac{1}{8n(n+1)} = \frac{1}{8\sqrt{\pi}} \left(\frac{1}{n} - \frac{1}{n+1} \right)$$

c) Montrer que pour tous les entiers k , tels que $1 \leq k < p$:

$$0 \leq a_p - a_k \leq \frac{1}{8k\sqrt{\pi}}$$

et en déduire, pour tout entier $k > 0$:

$$0 \leq \frac{1}{\sqrt{\pi}} - a_k \leq \frac{1}{8k\sqrt{\pi}}, \text{ puis } \frac{1}{\pi k} - \frac{1}{4\pi k^2} \leq \frac{a_k^2}{k} \leq \frac{1}{\pi k}$$

3. Dans un plan rapporté à un repère (O, \vec{i}, \vec{j}) , une puce située à un instant t au point de coordonnées $(i, j) \in \mathbb{Z}^2$ se déplace aléatoirement pour aller en l'un des 4 points $(i+1, j-1)$, $(i+1, j+1)$, $(i-1, j-1)$, $(i-1, j+1)$ avec la même probabilité à l'instant $t+1$. La puce est à l'origine O à l'instant $t=0$ et chaque saut est indépendant du précédent.

On note X_k et Y_k les coordonnées de la puce à l'instant k .

a) Calculer la probabilité des événements $(X_k = 0)$ et $(Y_k = 0)$.

b) Pour $n \geq 1$, on note U_n le nombre de passages à l'origine O entre les instants 1 et $2n$.

Montrer que : $E(U_n) = \sum_{k=1}^n \frac{a_k^2}{k}$.

c) En déduire un équivalent de $E(U_n)$ lorsque n tend vers $+\infty$. (On pourra utiliser : $\sum_{k=1}^n \frac{1}{k} \sim \ln(n)$ lorsque $n \rightarrow +\infty$.)

Solution :

1. Par l'absurde : s'il existe n_0 tel que : $a_{n_0} > \ell$ alors $\forall n \geq n_0, a_n \geq \ell + \epsilon$ (en posant $\epsilon = \frac{a_{n_0} - \ell}{2}$)

d'où en passant à la limite : $\ell \geq \ell + \epsilon \implies \epsilon \leq 0$. Ce qui est absurde.

2. a) Soit $n > 0$. On a $a_n = \frac{\sqrt{n}}{4^n} \binom{2n}{n} \sim \frac{\sqrt{n}}{4^n} \frac{\left(\frac{2n}{e}\right)^{2n} \sqrt{4\pi n}}{\left(\frac{n}{e}\right)^{2n} 2\pi n} = \frac{1}{\sqrt{\pi}}$. Donc $(a_n)_{n \in \mathbb{N}^*}$ converge vers $\frac{1}{\sqrt{\pi}}$.

b) Pour tout entier $n > 0$:

• $a_{n+1} - a_n = a_n \left(\frac{a_{n+1}}{a_n} - 1 \right) = a_n \frac{2n+1 - 2\sqrt{n}\sqrt{n+1}}{2\sqrt{n}\sqrt{n+1}} = a_n \frac{(\sqrt{n+1} - \sqrt{n})^2}{2\sqrt{n}\sqrt{n+1}}$, donc $a_{n+1} - a_n \geq 0$

• $(\sqrt{n+1} - \sqrt{n})^2 = \frac{(n+1) - n}{(\sqrt{n+1} + \sqrt{n})^2} \leq \frac{1}{4\sqrt{n(n+1)}}$,

car $(\sqrt{n+1} + \sqrt{n})^2 - 4\sqrt{n(n+1)} = (\sqrt{n+1} - \sqrt{n})^2 \geq 0$

• d'où : $0 \leq a_{n+1} - a_n = a_n \frac{(\sqrt{n+1} - \sqrt{n})^2}{2\sqrt{n}\sqrt{n+1}} \leq a_n \frac{1}{4\sqrt{n(n+1)}} \frac{1}{2\sqrt{n(n+1)}} \leq \frac{a_n}{8n(n+1)}$

- enfin, $a_n \leq \frac{1}{\sqrt{\pi}}$ et $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$ entraînent : $a_{n+1} - a_n \leq \frac{1}{8\sqrt{\pi}} \left[\frac{1}{n} - \frac{1}{n+1} \right]$.

c) Soit $1 \leq k < p$:

- $a_p - a_k \geq 0$ car la suite (a_n) est croissante.
- et par télescopage : $0 \leq a_p - a_k \leq \frac{1}{8\sqrt{\pi}} \left(\frac{1}{k} - \frac{1}{p} \right) \leq \frac{1}{8k\sqrt{\pi}}$
- pour tout entier $k > 0$, on fait tendre p vers $+\infty$, on obtient : $0 \leq \frac{1}{\sqrt{\pi}} - a_k \leq \frac{1}{8k\sqrt{\pi}}$
- puis : $\frac{1}{k\pi} \left(1 - \frac{1}{8k} \right)^2 \leq \frac{a_k^2}{k} \leq \frac{1}{\pi k}$, ce qui implique que $\frac{1}{\pi k} - \frac{1}{4\pi k^2} \leq \frac{a_k^2}{k} \leq \frac{1}{\pi k}$

3. a) Par symétrie : $P(X_k = 0) = P(Y_k = 0)$.

Pour k impair : il est impossible de revenir en 0 après un nombre impair de déplacements.
 Pour $k = 2p$ pair : il faut autant de déplacements vers la droite (+1) que de déplacements vers la gauche (-1)

$$P(X_{2p} = 0) = \binom{2p}{p} \frac{1}{2^{2p}} = \frac{a_p}{\sqrt{p}}$$

b) Si $n \geq 1$, $U_n = \sum_{k=1}^{2n} 1_{X_k=Y_k=0}$. Par indépendance des variables aléatoires X_k et Y_k :

$$\begin{aligned} E(U_n) &= \sum_{k=1}^{2n} P(X_k = 0 \cap Y_k = 0) = \sum_{k=1}^{2n} P(X_k = 0)P(Y_k = 0) \\ &= \sum_{k=1}^n P(X_{2k} = 0)P(Y_{2k} = 0) = \sum_{k=1}^n (P(X_{2k} = 0))^2 = \sum_{k=1}^n \frac{a_k^2}{k} \end{aligned}$$

On utilise la question 2.c : $\frac{1}{\pi k} - \frac{1}{4\pi k^2} \leq \frac{a_k^2}{k} \leq \frac{1}{\pi k}$, d'où en sommant :

$$\frac{1}{\pi} \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{1}{4\pi k^2} \leq E(U_n) \leq \frac{1}{\pi} \sum_{k=1}^n \frac{1}{k}$$

Or la série harmonique diverge vers $+\infty$ et la série $\sum_{k=1}^n \frac{1}{4\pi k^2}$ est une série de Riemann

convergente d'où : $\lim_{n \rightarrow +\infty} \frac{E(U_n)}{n} = \frac{1}{\pi}$ et $E(U_n) \sim \frac{1}{\pi} \sum_{k=1}^n \frac{1}{k} \sim \frac{\ln(n)}{\pi}$ lorsque n tend vers $+\infty$.

Exercice 3.17.

1. Dans un schéma de Bernoulli de probabilité de succès $p \in]0, 1[$, on définit, pour $n \in \mathbb{N}^*$, la variable aléatoire X_n égale au nombre d'échecs avant le $n^{\text{ème}}$ succès. Déterminer la loi de X_n .

On dit que X_n suit la loi binomiale négative de paramètres n et p .

Soit (Ω, \mathcal{A}, P) un espace probabilisé, et X une variable aléatoire définie sur (Ω, \mathcal{A}, P) à valeurs dans \mathbb{N} ; on note : $\forall k \in \mathbb{N}, P(X = k) = p_k$.

On dit que X appartient à \mathcal{E} , si X vérifie la propriété suivante notée (\mathcal{P}) :

$$\exists a \in]-\infty, 1[, \exists b \in \mathbb{R}_+^*, \text{ tels que } \begin{cases} a = 0 \text{ ou } b/a \in \mathbb{Z} \\ \forall k \in \mathbb{N}^*, p_k = \left(a + \frac{b}{k}\right) p_{k-1} \end{cases}$$

On cherche à déterminer les lois des variables aléatoires appartenant à \mathcal{E} .

2. a) On suppose $a = 0$.

Quelles variables aléatoires vérifiant (\mathcal{P}) obtient-on ?

b) Vérifier que toute variable aléatoire suivant une loi binomiale $\mathcal{B}(n, p)$ appartient à \mathcal{E} .
Que valent alors a et b ?

c) Même question pour une loi binomiale négative.

d) On suppose que $X \in \mathcal{E}$ admet une espérance. Exprimer $E(X)$ en fonction de a et b .
Vérifier ce résultat pour les variables aléatoires obtenues aux questions a) et b).

3. Dans cette question, on suppose que X vérifie (\mathcal{P}) avec $a \neq 0$.

a) Montrer que : $\exists c \in \mathbb{R}, \forall k \in \mathbb{N}^*, p_k = p_0 \frac{a^k}{k!} \prod_{i=0}^{k-1} (c + i)$.

b) Déterminer la loi de X si $a < 0$.

c) Déterminer la loi de X si $a \in]0, 1[$.

d) Conclure.

Solution :

1. C'est une question classique : $X_n(\Omega) = \mathbb{N}$ et $P(X_n = k) = \binom{n+k-1}{k} p^n (1-p)^k$.

2. a) Si $a = 0$ alors $p_k = \frac{b^k}{k!} p_0$ (par récurrence), et $p_0 = e^{-b}$ par poids total de 1, donc $X \hookrightarrow \mathcal{P}(b)$. On vérifie la réciproque.

b) Si $X \hookrightarrow \mathcal{B}(n, p)$, alors, pour $1 \leq k \leq n$,

$$p_k = \binom{n}{k} p^k q^{n-k} \Rightarrow \frac{p_k}{p_{k-1}} = \frac{n-k+1}{k} \frac{p}{q} = \left(-1 + \frac{n+1}{k}\right) \frac{p}{q}$$

d'où $a = -p/q$ et $b = -(n+1)a$, avec $b/a \in \mathbb{Z}$, qui conviennent ; on a bien $p_k = 0$ pour $k > n$.

c) Si $X \hookrightarrow \mathcal{BN}(n, p)$, pour $k \in \mathbb{N}^*$,

$$\frac{p_k}{p_{k-1}} = \frac{n+k-1}{k} q = \left(1 + \frac{n-1}{k}\right) q$$

donc $a = q$ et $b = (n-1)a$ conviennent, avec $b/a \in \mathbb{Z}$.

d) En sommant pour $k \in \mathbb{N}^*$,

$$k p_k = a(k-1) p_{k-1} + a p_{k-1} + b p_{k-1} \Rightarrow E(X) = a E(X) + a + b \Rightarrow E(X) = \frac{a+b}{1-a}$$

Puis :

- $X \hookrightarrow \mathcal{P}(b) \Rightarrow E(X) = b$;
- $X \hookrightarrow \mathcal{B}(n, p) \Rightarrow E(X) = np$.

3. a) $k = 1$ donne nécessairement $c = 1 + b/a$; puis vérification par récurrence.

b) Si $a < 0$ et $b/a \in \mathbb{Z}$, on pose $n = -\frac{b}{a} - 1 \in \mathbb{N}$, d'où

$$p_k = p_0 \frac{a^k}{k!} \prod_{i=0}^{k-1} (-n + i) \Rightarrow p_k = \begin{cases} 0 & \text{si } k > n \\ p_0 \binom{n}{k} (-a)^k & \text{si } k \in \llbracket 0, n \rrbracket \end{cases}$$

Puis $p_0 = \frac{1}{(1-a)^n}$ par poids total 1, donc $X \hookrightarrow \mathcal{B}\left(n, \frac{-a}{1-a}\right)$.

c) De même, si $a \in]0, 1[$, on pose $n = 1 + \frac{b}{a} \in \mathbb{N}$, puis

$$\forall k, p_k = p_0 \binom{n+k-1}{k} a^k$$

avec $p_0 = (1-a)^n$, donc $X \hookrightarrow \mathcal{BN}(n, 1-a)$.

d) Les lois des $X \in \mathcal{E}$ sont les lois binomiales, binomiales négatives ou de Poisson.

Exercice 3.18.

Soit h la fonction définie sur $[0, 1]$ par : $h(x) = \begin{cases} -x \frac{\ln x}{\ln 2} & \text{si } x \in]0, 1] \\ 0 & \text{si } x = 0 \end{cases}$.

1. Soit n un entier naturel tel que $n2$.

a) Soit $\mathcal{O} = \{(p_1, \dots, p_{n-1}) \in (\mathbb{R}_+^*)^{n-1} / 1 - p_1 - \dots - p_{n-1} > 0\}$.

Pour $(p_1, \dots, p_{n-1}) \in \mathcal{O}$, on pose

$$h_n((p_1, \dots, p_{n-1})) = h(1 - p_1 - \dots - p_{n-1}) + \sum_{k=1}^{n-1} h(p_k).$$

Montrer que \mathcal{O} est un ouvert de \mathbb{R}^{n-1} . Montrer que h_n admet au plus un extremum sur \mathcal{O} .

b) Soit $(p_1, \dots, p_{n-1}, p_n) \in]0, 1]^n$ tel que $\sum_{k=1}^n p_k = 1$. Montrer que $\sum_{k=1}^n h(p_k) \leq \frac{\ln n}{\ln 2}$.

c) Soit X une variable aléatoire telle que $X(\Omega) = \{x_1, \dots, x_n\}$.

On pose $H(X) = \sum_{x \in X(\Omega)} h(P(X = x))$.

Montrer que $H(X) \leq \frac{\ln n}{\ln 2}$. Quand a-t-on égalité ?

2. Soit Y une variable aléatoire suivant la loi géométrique de paramètre $p \in]0, 1]$. On note $p_k = P(Y = k)$ et $m = E(Y)$.

a) Montrer que $H(Y)$ existe et déterminer sa valeur.

b) Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}^*$, $E(X) = m$ et $H(X)$ existe.

On suppose que pour tout $k \in \mathbb{N}^*$, $q_k = P(X = k) > 0$.

Montrer que pour tout $k \in \mathbb{N}^*$, $\ln(p_k) - \ln(q_k) \leq \frac{p_k}{q_k} - 1$. En déduire que $H(X) \leq H(Y)$.

Solution :

1. a) L'ensemble \mathcal{O} est ouvert car défini à partir d'une inéquation stricte polynomiale. On cherche alors les points critique de h_n sur \mathcal{O} . Il vient,

$$h_n(x) = \frac{1}{\ln 2} \left[\left(\sum_{k=1}^{n-1} p_k - 1 \right) \ln \left(1 - \sum_{k=1}^{n-1} p_k \right) - \sum_{k=1}^{n-1} p_k \ln(p_k) \right]$$

et pour $i \in \llbracket 1, n-1 \rrbracket$ $\partial_i h_n = \frac{1}{\ln 2} \left[\ln \left(1 - \sum_{k=1}^{n-1} p_k \right) - \ln(p_i) \right]$

On trouve ainsi un unique point critique qui est $p_1 = p_2 = \dots = p_{n-1} = \frac{1}{n}$

b) On vérifie que la fonction $-h$ est convexe, donc h concave. Comme $p_n = 1 - \sum_{k=1}^{n-1} p_k$, on a

$$h\left(\frac{1}{n} \sum_{k=1}^n p_k\right) \geq \frac{1}{n} \sum_{k=1}^n h(p_k)$$

et comme $\sum_{k=1}^n p_k = 1$, il vient $\sum_{k=1}^n h(p_k) \leq nh(1/n) = \frac{\ln n}{\ln 2}$

c) On pose $p_k = P(X = k)$.

• si $p_k > 0$ pour tout k , par la question précédente $H(X) \leq \frac{\ln n}{\ln 2}$.

Si X suit la loi uniforme sur (x_1, \dots, x_n) , alors $p_k = 1/n$ et on vérifie que $H(X) = \frac{\ln n}{\ln 2}$.

Réciproquement si on a égalité $H(X) = \frac{\ln n}{\ln 2}$, alors

$$H(X) = h(p_n) + \sum_{k=1}^{n-1} h(p_k) = h\left(1 - \sum_{k=1}^{n-1} p_k\right) + \sum_{k=1}^{n-1} h(p_k) = h_n(p_1, \dots, p_{n-1})$$

Ainsi h_n atteint son maximum en un point critique et par la question précédente, on a pour tout $k \in \llbracket 1, n-1 \rrbracket$, $p_k = 1/n$ et donc $p_n = 1/n$.

• sinon on peut supposer que $0 < p_1 < p_2 < \dots < p_m$ et $p_{m+1} = \dots = p_n = 0$.

On a alors $H(X) \sum_{k=1}^m h(p_k) \leq \frac{\ln m}{\ln 2} \leq \frac{\ln n}{\ln 2}$.

Si X suit la loi uniforme, la démonstration est identique à la précédente et réciproquement, en cas d'égalité $\frac{\ln n}{\ln 2} \leq H(X) \leq \frac{\ln m}{\ln 2} \leq \frac{\ln n}{\ln 2}$. Donc $m = n$.

2. a) On a $m = 1/p$, $h(p_k) = h(pq^{k-1}) = \frac{1}{\ln 2} (-\ln(p/q)pq^{k-1} - (\ln q)kpq^{k-1})$ et

$$H(Y) = \frac{\ln(q/p)}{\ln 2} - \frac{1}{p} \frac{\ln q}{\ln 2} = \frac{-q \ln q - p \ln p}{p \ln 2}$$

b) Comme $\ln x \leq x - 1$, on a $\ln p_k - \ln q_k \leq p_k/q_k - 1$. En s'affranchissant du dénominateur $\ln 2$, il vient

$$\begin{aligned}
 H(Y) - H(X) &= \sum_{k=1}^n q_k \ln(q_k) - p_k \ln(p_k) = \sum_{k=1}^n q_k (\ln(q_k) - \ln(p_k)) + (q_k - p_k) \ln(p_k) \\
 &= \sum_{k=1}^n q_k (\ln(q_k) - \ln(p_k)) + \sum_{k=1}^n (q_k - p_k) \ln(pq^{k-1}) \\
 H(Y) - H(X) &= \sum_{k=1}^n q_k (\ln(q_k) - \ln(p_k)) + \sum_{k=1}^n (q_k - p_k) (\ln p + (k-1) \ln q) \\
 &= \sum_{k=1}^n q_k (\ln(q_k) - \ln(p_k)) \quad (\text{même espérance}) \\
 &= \sum_{k=1}^n q_k (\ln(q_k) - \ln(p_k)) - \sum_{k=1}^n q_k \left(1 - \left(\frac{p_k}{q_k}\right)\right) \\
 &= \sum_{k=1}^n q_k \left(\ln(q_k) - \ln(p_k) - 1 + \frac{p_k}{q_k}\right) \\
 & \qquad \qquad \qquad H(Y) - H(X) \geq 0
 \end{aligned}$$

Exercice 3.19.

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

1. Soit T une variable aléatoire suivant la loi exponentielle de paramètre λ et $m \in \mathbb{R}_+^*$. Déterminer la loi de la variable aléatoire $[mT] + 1$ (où $[\cdot]$ désigne la fonction partie entière).

2. Soit T une variable aléatoire positive telle que, pour tout $n \in \mathbb{N}$, $X_n = [2^n T] + 1$ est une variable aléatoire suivant une loi géométrique. On note $p_n \in]0, 1[$ son paramètre et l'on pose $q_n = 1 - p_n$.

a) Montrer que, pour tout réel x , $[x] = 0$ si et seulement si $[2x] = 0$ ou $[2x] = 1$.

b) En considérant $P(X_n = 1)$, établir que pour tout $n \in \mathbb{N}$, $q_{n+1} = \sqrt{q_n}$.

c) Montrer que, pour tout $(x, n) \in \mathbb{R}_+ \times \mathbb{N}$, on a :

$$(X_n \geq [2^n x] + 2) \subset (T \geq x) \subset (X_n \geq [2^n x])$$

d) En déduire que T suit une loi exponentielle et déterminer son paramètre en fonction de q_0 .

Solution :

1. On note $Y = [mT] + 1$. On remarque que $Y(\Omega) = \mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned}
 P(Y = n) &= P(n-1 \leq mT < n) = \int_{\frac{n-1}{m}}^{\frac{n}{m}} \lambda e^{-\lambda t} dt = -e^{\lambda n/m} + e^{-\lambda(n-1)/m} \\
 &= (1-p)^{n-1} p.
 \end{aligned}$$

où $p = 1 - e^{-\lambda/m}$. On en déduit que T suit la loi géométrique de paramètre $1 - e^{-\lambda/m}$.

2. a) Il suffit de revenir à la définition de la partie entière. En effet,

▷ si $[x] = 0$, alors $0 \leq 2x < 2$, et donc $[2x] = 0$ ou $[2x] = 1$.

▷ si $\lfloor 2x \rfloor = 0$, alors $0 \leq 2x < 1$, donc $\lfloor x \rfloor = 0$ et si $\lfloor 2x \rfloor = 1$, alors $\frac{1}{2} \leq x < 1$, donc là encore $\lfloor x \rfloor = 0$.

b) Soit $n \in \mathbb{N}$. On a $(X_n = 1) = (\lfloor 2^n T \rfloor = 0)$. D'après 2.a., cet événement est donc la réunion disjointe des événements $(\lfloor 2^{n+1} T \rfloor = 0)$ et $(\lfloor 2^{n+1} T \rfloor = 1)$. Ainsi,

$$P(X_n = 1) = P(X_{n+1} = 1) + P(X_{n+1} = 2).$$

soit $p_n = p_{n+1} + (1 - p_{n+1})p_{n+1}$, que l'on peut réécrire, sous la forme :

$$1 - q_n = 1 - q_{n+1} + (1 - q_{n+1})q_{n+1}.$$

Ainsi, $q_{n+1} = \sqrt{q_n}$.

c) Soit $(x, n) \in \mathbb{R}_+ \times \mathbb{N}$. Si $X_n \geq \lfloor 2^n x \rfloor + 2$, alors $\lfloor 2^n T \rfloor \geq \lfloor 2^n x \rfloor + 1 \geq 2^n x$, et donc $2^n T \geq 2^n x$, i.e. $T \geq x$. De plus, si $T \geq x$, alors $2^n T \geq 2^n x \geq \lfloor 2^n x \rfloor$ et donc $X_n \geq \lfloor 2^n x \rfloor$. On en déduit que

$$(X_n \geq \lfloor 2^n x \rfloor + 2) \subset (T \geq x) \subset (X_n \geq \lfloor 2^n x \rfloor).$$

d) Soit $(x, n) \in \mathbb{R}_+ \times \mathbb{N}$. De l'inclusion de la question précédente, on déduit que

$$P(X_n \geq \lfloor 2^n x \rfloor + 2) \leq P(T \geq x) \leq P(X_n \geq \lfloor 2^n x \rfloor).$$

Ainsi, puisque X_n suit la loi géométrique de paramètre $p_n = 1 - q_n$,

$$q_n^{\lfloor 2^n x \rfloor + 1} \leq P(T \geq x) \leq q_n^{\lfloor 2^n x \rfloor - 1}.$$

Comme $q_n = q_0^{2^{-n}}$, on obtient

$$q_0^{\lfloor 2^n x \rfloor / 2^n + 1 / 2^n} \leq P(T \geq x) \leq q_0^{\lfloor 2^n x \rfloor / 2^n - 1 / 2^n}.$$

Or on sait que $\lim_{n \rightarrow +\infty} \frac{\lfloor 2^n x \rfloor}{2^n} = x$. On déduit de ce qui précède qu'en faisant tendre n vers $+\infty$, que $P(T \geq x) = q_0^x = e^{-\lambda x}$ où $\lambda = -\ln(q_0)$.

On reconnaît ainsi la fonction de répartition d'une variable suivant la loi exponentielle de paramètre $-\ln(q_0)$.

Exercice 3.20.

Soit $p \in]0, 1[$, λ et μ deux réels strictement positifs. On pose

$$f(x) = \begin{cases} p\mu e^{\mu x} & \text{si } x < 0 \\ (1-p)\lambda e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

1. Montrer que f est une densité de probabilité.

On note X une variable aléatoire de densité f .

2. Soit $(Y_n)_n$ une suite de variables aléatoires indépendantes définies sur (Ω, \mathcal{A}, P) et de même loi que X . Pour tout $n \geq 1$, on pose :

$$I_n(\omega) = \begin{cases} 1 & \text{si } Y_n(\omega) > 0 \\ 0 & \text{si } Y_n(\omega) \leq 0 \end{cases}$$

Montrer que $(I_n)_n$ est une suite de variables aléatoires indépendantes de même loi que l'on précisera.

3. Soit $Z_n = I_n Y_n$.

a) Montrer que $(Z_n)_n$ est une suite de variables aléatoires indépendantes de même loi. Déterminer la fonction de répartition de Z_1 .

b) On pose $S_n = \sum_{k=1}^n I_k$ et $V_n = \sum_{k=1}^n Z_k$. Les deux variables aléatoires S_n et V_n sont-elles indépendantes? Donner la loi de S_n .

4. a) Déterminer la loi suivie par la somme de n variables aléatoires indépendantes, toutes de loi exponentielle $\mathcal{E}(\alpha)$ (on pourra commencer par le cas $n = 2$).

b) Déterminer, à l'aide d'une intégrale, pour tout $n \geq 1, k \in \llbracket 1, n \rrbracket$ et $x \geq 0$: $P_{(S_n=k)}(V_n \leq x)$.

c) En déduire une expression de la loi de V_n .

Solution :

1. La fonction f est continue sur \mathbb{R} sauf en 0 et positive. De plus

$$\int_{-\infty}^{+\infty} f(t) dt = p \int_{-\infty}^0 \mu e^{\mu t} dt + (1-p) \int_0^{+\infty} \lambda e^{-\lambda t} dt = p + 1 - p = 1$$

2. Les variables aléatoires I_n sont indépendantes, car I_n ne dépend que de Y_n et les (Y_n) sont indépendantes.

I_n suit une loi de Bernoulli de paramètre $P(Y_n > 0) = P(X > 0) = 1 - p$.

3. a) Les variables aléatoires Z_n sont indépendantes, car Z_n ne dépend que de I_n et Y_n . On rappelle que les variables Y_n ne chargent pas les points. La loi de Z_n est ainsi donnée par $P(Z_n < 0) = 0, P(Z_n = 0) = P(Y_n \leq 0) = p$ et pour $x > 0$

$$P(Z_n \leq x) = P([I_n = 1] \cap [0 < Y_n \leq x]) = P(0 < Y_n \leq x) = (1-p)(1 - e^{-\lambda x})$$

car $I_n = 1$ si et seulement si $Y_n > 0$.

Les Z_n suivent la même loi de fonction de répartition

$$F_1(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } x = 0 \\ (1-p)(1 - e^{-\lambda x}) & \text{si } x > 0 \end{cases}$$

b) Les variables aléatoires S_n et V_n ne sont pas indépendantes, car

$$[S_n = n] = \bigcap_{j=1}^n (I_j = 1), \quad [V_n = 0] = \bigcap_{j=1}^n (I_j = 0) \text{ et } P([S_n = n] \cap [V_n = 0]) = 0$$

La loi suivie par S_n est la loi binomiale $\mathcal{B}(n, 1 - p)$.

4. a) La loi suivie par la somme de n variables indépendantes, toutes de loi exponentielle $\mathcal{E}(\alpha)$ est la loi Gamma de paramètres (α, n) , de densité :

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{e^{-\alpha x} \alpha^n x^{n-1}}{(n-1)!} & \text{si } x \geq 0 \end{cases}$$

b) Réaliser l'événement $[S_n = k]$ c'est choisir parmi I_1, \dots, I_n les k variables qui prennent alors la valeur 1, les autres prenant la valeur 0 ; et alors $[V_n \leq x] = [V_k \leq x]$. On a également la réciproque de manière immédiate. Ainsi

$$P_{(S_n=k)}(V_n \leq x) = P(V_k \leq x) = \int_0^x \frac{e^{-\lambda t} \lambda^k t^{k-1}}{(k-1)!} dt$$

c) La famille $([S_n = k])_{0 \leq k \leq n}$ forme un système complet d'événements. Ainsi, pour tout x :

$$P(V_n \leq x) = \sum_{k=0}^n P_{(S_n=k)}(V_n \leq x) P(S_n = k)$$

et

- si $x < 0$, $F_n(x) = 0$
- $F_n(0) = P_{(S_n=0)}(V_n \leq 0) P(S_n = 0) = P(X \leq 0)^n (p)^n = p^{2n}$
- si $x > 0$, $F_n(x) = p^{2n} + \sum_{k=1}^n \binom{n}{k} (1-p)^k p^{n-k} \int_0^x \frac{e^{-\lambda t} \lambda^k t^{k-1}}{(k-1)!} dt.$