

PROBABILITÉS

Exercice 3.1.

On considère deux variables aléatoires indépendantes, X et Y , définies sur le même espace probabilisé (Ω, \mathcal{A}, P) , telles que X suit la loi géométrique $\mathcal{G}(p)$ où $p \in]0, 1[$ et Y suit la loi uniforme $\mathcal{U}([0, 2])$, on pose $q = 1 - p$.

On pose : $T = X + Y$, $Z = \lfloor T \rfloor$, où $\lfloor T \rfloor$ désigne la partie entière de T . On admet que T et Z sont des variables aléatoires définies sur l'espace (Ω, \mathcal{A}, P) .

On dit qu'une variable aléatoire V est à densité généralisée si sa fonction de répartition F_V est continue sur \mathbb{R} et de classe C^1 sur \mathbb{R} sauf sur un ensemble dénombrable de points. Une densité f_V de V s'obtient alors, en tout point où F_V est dérivable, par la relation $F_V' = f_V$.

1. On note F_T la fonction de répartition de T .

a) Donner, pour tout réel t , l'expression de $F_T(t)$ en distinguant a priori les cas :

$t < 1$, $1 \leq t < 2$, $2 \leq t < 3$ et $k \leq t < k + 1$, où k est un entier naturel supérieur ou égal à 3.

(on remarquera que le cas particulier $2 \leq t < 3$ rejoint le cas général, $k \geq 3$).

b) Vérifier que T est une variable aléatoire à densité généralisée.

c) Donner l'expression d'une densité de T .

2. a) Donner la loi de Z .

b) Calculer $E(Z)$.

c) Donner la loi de $\lfloor Y \rfloor$ ainsi que son espérance.

d) En remarquant que si x est un entier naturel et y un nombre réel, alors $\lfloor x + y \rfloor = x + \lfloor y \rfloor$, retrouver la loi de Z et donner la valeur de son espérance.

Solution :

1. a) On a $X(\Omega) = \mathbb{N}^*$ et, pour tout k de \mathbb{N}^* , $P(X = k) = pq^{k-1}$. D'autre part $Y(\Omega) = [0, 2]$ et pour tout y de $[0, 2]$, on $F_Y(y) = \frac{y}{2}$. On a donc $T(\Omega) = [1, +\infty[$.

i) Si $t < 1$, on a évidemment $F_T(t) = 0$.

ii) Si $t \in [1, 2[$, on a :

$$(X + Y \leq t) = ([X = 1] \cap [X + Y \leq t]) = ([X = 1] \cap [Y \leq t - 1])$$

On a donc : $P(X + Y \leq t) = P([X = 1] \cap [Y \leq t - 1])$ et comme les variables X et Y sont indépendantes, on obtient :

$$P(X + Y \leq t) = P([X = 1])P([Y \leq t - 1]) = \frac{p(t-1)}{2}$$

iii) Si $t \in [2, 3[$, on a :

$$\begin{aligned} (X + Y \leq t) &= (([X = 1] \cap [X + Y \leq t]) \cup ([X = 2] \cap [X + Y \leq t])) \\ &= (([X = 1] \cap [Y \leq t - 1]) \cup ([X = 2] \cap [Y \leq t - 2])). \end{aligned}$$

Par incompatibilité et par indépendance, on obtient :

$$P(X + Y \leq t) = P([X = 1])P([Y \leq t - 1]) + P([X = 2])P([Y \leq t - 2])$$

Soit :

$$P(X + Y \leq t) = \frac{p(t-1)}{2} + \frac{pq(t-2)}{2}$$

iv) Soit $k \geq 3$ et $t \in [k, k + 1[$. On a alors :

$$\begin{aligned} [T \leq t] &= [X \leq k - 2] \cup ([X = k - 1] \cap [X + Y \leq t]) \cup ([X = k] \cap [X + Y \leq t]) \\ &= [X \leq k - 2] \cup ([X = k - 1] \cap [Y \leq t - k + 1]) \cup ([X = k] \cap [Y \leq t - k]) \end{aligned}$$

Là encore, par indépendance et incompatibilité, on obtient :

$$\begin{aligned} P([T \leq t]) &= P([X \leq k - 2]) + P([X = k - 1])P([Y \leq t - k + 1]) \\ &\quad + P([X = k])P([Y \leq t - k]) \end{aligned}$$

$$\text{Or } P([X \leq k - 2]) = \sum_{j=1}^{k-2} P([X = j]) = \sum_{j=1}^{k-2} pq^{j-1} = p \frac{1 - q^{k-2}}{1 - q} = 1 - q^{k-2}$$

(on peut aussi, si on préfère, calculer $P(X > k - 2)$. D'où :

$$P([T \leq t]) = 1 - q^{k-2} + pq^{k-2} \frac{t-k+1}{2} + pq^{k-1} \frac{t-k}{2}$$

On constate que le cas $k = 2$ rejoint le cas général, on a donc finalement :

$$\begin{cases} \text{si } t < 1, F_T(t) = 0 \\ \text{si } t \in [1, 2[, F_T(t) = \frac{p(t-1)}{2} \\ \text{si } t \in [k, k+1[, k \geq 2, F_T(t) = 1 - q^{k-2} + pq^{k-2} \frac{t-k+1}{2} + pq^{k-1} \frac{t-k}{2} \end{cases}$$

b) F_T est clairement continue et de classe C^1 sur \mathbb{R} sauf en tout point de \mathbb{N}^* , ensemble qui est bien dénombrable. Il reste à étudier la continuité aux points de \mathbb{N}^* .

i) Continuité en 1 : on a : $\lim_{t \rightarrow 1^-} F_T(t) = 0$ et $\lim_{t \rightarrow 1^+} F_T(t) = \lim_{t \rightarrow 1^+} \frac{p(t-1)}{2} = 0$, ce qui prouve la continuité de F_T au point 1.

ii) Continuité en 2 :

$$\begin{aligned} \lim_{t \rightarrow 2^-} F_T(t) &= \lim_{t \rightarrow 2^-} \frac{p(t-1)}{2} = \frac{p}{2} \text{ et} \\ \lim_{t \rightarrow 2^+} F_T(t) &= \lim_{t \rightarrow 2^+} \frac{p(t-1)}{2} + \frac{pq(t-2)}{2} = \frac{p}{2}. \end{aligned}$$

En conclusion, F_T est continue en 2.

iii) Continuité en k où k est un entier naturel supérieur ou égal à 3 :

$$\begin{aligned} \lim_{t \rightarrow k^-} F_T(t) &= \lim_{t \rightarrow k^-} \left(1 - q^{k-3} + pq^{k-3} \frac{t-k+2}{2} + pq^{k-2} \frac{t-k+1}{2} \right) \\ &= 1 - q^{k-3} + pq^{k-3} + \frac{pq^{k-2}}{2} \end{aligned}$$

$$\text{Soit } \lim_{t \rightarrow k^-} F_T(t) = 1 - q^{k-2} + \frac{pq^{k-2}}{2}.$$

D'autre part :

$$\begin{aligned} \lim_{t \rightarrow k^+} F_T(t) &= \lim_{t \rightarrow k^+} \left(1 - q^{k-2} + pq^{k-2} \frac{t-k+1}{2} + pq^{k-1} \frac{t-k}{2} \right) \\ &= 1 - q^{k-2} + \frac{pq^{k-2}}{2}. \end{aligned}$$

Ainsi, F_T est bien continue en tout point k entier tel que $k \geq 3$.

En conclusion : F_T est continue sur \mathbb{R} et est de classe C^1 sur \mathbb{R} sauf éventuellement aux points appartenant à \mathbb{N}^* qui est un ensemble dénombrable. T est donc bien une variable à densité.

c) On obtient une densité de T en dérivant F_T sur tous les intervalles ouverts où elle est dérivable et en affectant une valeur arbitraire à sa dérivée f_T aux points de \mathbb{N}^* .

On obtient :

$$\begin{cases} \text{si } t < 1, & f_T(t) = 0 \\ \text{si } t \in [1, 2[, & f_T(t) = \frac{p}{2} \\ \text{si } t \in [k, k+1[, k \geq 2, & f_T(t) = \frac{pq^{k-2}}{2} + \frac{pq^{k-1}}{2} \end{cases}$$

2. a) On a $Z(\Omega) = \mathbb{N}^*$ et $P(Z = 1) = P(T < 2) = F_T(2) = \frac{p}{2}$.

Pour $k \geq 2$: $P(Z = k) = P(k \leq T < k+1) = F_T(k+1) - F_T(k)$, d'où :

$$P(Z = k) = \left[1 - q^{k-2} + pq^{k-2} + \frac{pq^{k-1}}{2}\right] - \left[1 - q^{k-2} + \frac{pq^{k-2}}{2}\right]$$

Finalement :

$$\begin{cases} P(Z = 1) = \frac{p}{2} \\ \forall k \geq 2, P(Z = k) = \frac{pq^{k-2}(1+q)}{2} \end{cases}$$

b) on a :

$$\begin{aligned} E(Z) &= \frac{p}{2} + \frac{p(1+q)}{2} \sum_{k=2}^{+\infty} kq^{k-2} = \frac{p}{2} + \frac{p(1+q)}{2q} \sum_{k=2}^{+\infty} kq^{k-1} \\ &= \frac{p}{2} + \frac{p(1+q)}{2q} \left[\frac{1}{(1-q)^2} - 1 \right] \end{aligned}$$

Après simplification :

$$E(Z) = \frac{p+2}{2p}$$

c) On a $\lfloor Y \rfloor(\Omega) = \{0, 1\}$, et $P(\lfloor Y \rfloor = 0) = P(Y < 1) = \frac{1}{2}$.

De même $P(\lfloor Y \rfloor = 1) = P(1 \leq Y < 2) = \frac{1}{2}$. Donc :

$$\lfloor Y \rfloor \hookrightarrow \mathcal{B}(1/2), \text{ et } E(\lfloor Y \rfloor) = \frac{1}{2}$$

d) On a défini la variable Z par $Z = \lfloor X + Y \rfloor$. Or X prend ses valeurs dans \mathbb{N}^* , d'où : $\lfloor X + Y \rfloor = X + \lfloor Y \rfloor$.

On a donc, par indépendance : $P(Z = 1) = P(X = 1 \cap \lfloor Y \rfloor = 1) = p \times \frac{1}{2}$.

Pour $k \geq 2$, $P(Z = k) = P(((X = k) \cap (\lfloor Y \rfloor = 0)) \cup ((X = k-1) \cap (\lfloor Y \rfloor = 1)))$.

Par incompatibilité et indépendance, on obtient : $P(Z = k) = pq^{k-1} \times \frac{1}{2} + pq^{k-2} \times \frac{1}{2}$.

Enfin, X et $\lfloor Y \rfloor$ sont des variables aléatoires discrètes admettant des espérances, d'où : $E(Z) = E(X) + E(\lfloor Y \rfloor)$.

On obtient donc :

$$E(Z) = \frac{1}{p} + \frac{1}{2}$$

Exercice 3.2.

1. Soit $(u_n)_{n \geq 1}$ une suite réelle convergeant vers un réel λ .

Pour tout $n \geq 1$, on pose $v_n = \frac{1}{n} \sum_{k=1}^n u_k$.

En revenant à la définition de la limite d'une suite, montrer que la suite (v_n) converge vers λ .

2. Soit Y une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) telle que $Y(\Omega) = \mathbb{N}^*$. On suppose que Y admet une espérance.

Montrer que $E(Y) = \sum_{k=1}^{\infty} P(Y \geq k)$.

Soit X une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) telle que $X(\Omega) = \mathbb{N}$ et admettant une espérance.

On se propose de déterminer un développement de $S_n = \sum_{k=1}^n P(X < k)$, lorsque n tend vers $+\infty$.

3. a) Déterminer un équivalent de S_n , lorsque n tend vers $+\infty$.

b) Déterminer un équivalent de $S_n - n$, lorsque n tend vers $+\infty$.

4. On suppose que X admet un moment d'ordre 2, $E(X^2)$.

a) Montrer que : $S_n - n + E(X) \leq \frac{E(X^2)}{n}$.

b) En déduire que, lorsque n tend vers $+\infty$, on a :

$$S_n = n - E(X) + \varepsilon(n), \text{ avec } \lim_{n \rightarrow +\infty} \varepsilon(n) = 0.$$

Solution :

1. Traduisons le fait que $\lim_{n \rightarrow +\infty} u_n = \lambda$.

Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que si $n \geq N$, on a $|u_n - \lambda| < \varepsilon/2$.

On peut écrire, pour $n \geq N + 1$:

$$|v_n - \lambda| = \frac{1}{n} \left| \sum_{k=1}^n u_k - n\lambda \right| \leq \frac{1}{n} \sum_{k=1}^N |u_k - \lambda| + \frac{1}{n} \sum_{k=N+1}^n |u_k - \lambda|$$

Or : $\frac{1}{n} \sum_{k=N+1}^n |u_k - \lambda| \leq \frac{(n - N)}{n} \times \frac{\varepsilon}{2} < \frac{\varepsilon}{2}$

et : $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^N |u_k - \lambda| = \lim_{n \rightarrow +\infty} \frac{C_N}{n} = 0$, il existe donc $N' \in \mathbb{N}$ tel que si

$n \geq N'$, alors $\frac{1}{n} \sum_{k=1}^N |u_k - \lambda| < \frac{\varepsilon}{2}$.

Pour $n \geq \max(N + 1, N')$, on a donc $|v_n - \lambda| \leq \varepsilon$, ce qui prouve le résultat demandé.

2. La permutation des signes \sum étant autorisée pour des termes positifs ou nuls, on écrit :

$$\begin{aligned} \sum_{k=1}^{\infty} P(Y \geq k) &= \sum_{k=1}^{\infty} \sum_{i=k}^{\infty} P(Y = i) = \sum_{i=1}^{\infty} \sum_{k=1}^i P(Y = i) \\ &= \sum_{i=1}^{\infty} iP(Y = i) = E(Y) \end{aligned}$$

3. a) On sait que $\lim_{k \rightarrow +\infty} P(X < k) = 1$. Par la première question :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n P(Y < k) = \lim_{n \rightarrow +\infty} \frac{S_n}{n} = 1$$

Donc $S_n \sim n$, lorsque n tend vers ∞ .

$$\text{b) On a : } S_n - n = \sum_{k=1}^n (P(X < k) - 1) = - \sum_{k=1}^n P(X \geq k)$$

Cette dernière quantité tend vers $-E(X)$ par la question 2. Donc, comme $E(X) \neq 0$, $S_n - n \sim -E(X)$.

4. a) On a :

$$S_n - n + E(X) = - \sum_{k=1}^n P(X \geq k) + \sum_{k=1}^{\infty} P(X \geq k) = \sum_{k=n+1}^{\infty} P(X \geq k)$$

Par l'inégalité de Markov, $P(X \geq k) \leq \frac{E(X^2)}{k^2}$. Donc :

$$S_n - n + E(X) \leq E(X^2) \sum_{k=n+1}^{\infty} \frac{1}{k^2}$$

En utilisant une comparaison série/intégrale, avec la fonction $x \mapsto \frac{1}{x^2}$, il

vient facilement $\sum_{k=n+1}^{\infty} \frac{1}{k^2} \leq \frac{1}{n}$ et donc :

$$S_n - n + E(X) \leq \frac{E(X^2)}{n}$$

b) On obtient donc le développement asymptotique de S_n :

$$S_n = n - E(X) + O\left(\frac{1}{n}\right)$$

Ce qui contient en particulier le résultat demandé.

Exercice 3.3.

Soient X_1 et X_2 , deux variables aléatoires de densités respectives f_1 et f_2 strictement positives et dérivables sur \mathbb{R} .

On suppose qu'il existe une fonction g définie et dérivable sur \mathbb{R}^+ telle que pour tout $(x, y) \in \mathbb{R}^2$:

$$f_1(x)f_2(y) = g(x^2 + y^2)$$

1. On suppose dans cette question uniquement que X_1 et X_2 suivent la même loi normale centrée réduite. Montrer que la condition précédente est réalisée et déterminer g .

2. a) Exprimer, pour tout $x \in \mathbb{R}^*$ et pour tout $y \in \mathbb{R}$, $\frac{f_1'(x)}{xf_1(x)}$ en fonction de $g(x^2 + y^2)$ et $g'(x^2 + y^2)$.

b) Soit h la fonction définie sur \mathbb{R}^* par : $h(x) = \frac{f_1'(x)}{xf_1(x)}$.

Montrer que h est constante. On note a cette constante.

3. a) On définit la fonction k sur \mathbb{R} par : $k(x) = f_1(x)e^{-ax^2/2}$

Montrer que la fonction k est constante sur \mathbb{R} (on étudiera dans un premier temps les restrictions de k à \mathbb{R}_+^* , puis à \mathbb{R}_-^*).

En déduire une expression de f_1 .

b) Montrer que $a < 0$.

c) Montrer que X_1 suit une loi normale. Quels sont ses paramètres ?

4. a) Calculer $g(1)$.

b) Montrer que X_2 suit la même loi que X_1 .

Solution :

1. Si X_1 et X_2 suivent la loi normale centrée réduite, on a :

$$f_1(x)f_2(y) = \frac{1}{2\pi}e^{-(x^2+y^2)/2} = g(x^2 + y^2)$$

avec $g(t) = \frac{1}{2\pi}e^{-t/2}$, et g est bien définie et dérivable sur \mathbb{R}^+ .

2. a) On sait que pour tous x, y réels, $f_1(x)f_2(y) = g(x^2 + y^2)$, avec g dérivable sur \mathbb{R}^+ . Ainsi pour tous x et y :

$$f_1'(x)f_2(y) = 2xg'(x^2 + y^2)$$

Comme $f_1(x) \neq 0$, il vient, pour $x \neq 0$ et y réel :

$$\frac{f_1'(x)}{xf_1(x)} = 2 \times \frac{g'(x^2 + y^2)}{g(x^2 + y^2)}$$

b) Soit $x_1 \neq x_2$ deux réels non nuls, on a alors en prenant $x = x_1$ et $y = x_2$ puis $x = x_2$ et $y = x_1$:

$$h(x_1) = \frac{f'_1(x_1)}{x_1 f_1(x_1)} = 2 \times \frac{g'(x_1^2 + x_2^2)}{g(x_1^2 + x_2^2)} = 2 \times \frac{g'(x_2^2 + x_1^2)}{g(x_2^2 + x_1^2)} = \frac{f'_1(x_2)}{x_2 f_1(x_2)} = h(x_2)$$

ce qui montre que h est constante.

3. a) La fonction k est dérivable sur \mathbb{R} , avec pour $x \neq 0$:

$$k'(x) = f'_1(x)e^{-ax^2/2} - axf_1(x)e^{-ax^2/2} = xf_1(x)e^{-ax^2/2}(h(x) - a) = 0$$

Ainsi k est constante sur \mathbb{R}_-^* et \mathbb{R}_+^* et étant dérivable sur \mathbb{R} elle est continue sur \mathbb{R} et finalement constante sur \mathbb{R} .

$$\exists C \in \mathbb{R}, \forall x \in \mathbb{R}, f_1(x) = Ce^{ax^2/2}$$

b) La fonction f_1 est une densité de probabilité.

Donc $C > 0$ et $a < 0$, car autrement $\int_{-\infty}^{+\infty} f_1(x) dx$ divergerait ou ne pourrait pas avoir pour valeur 1 (qui est strictement positif!).

c) Posons $\sigma_1 = \sqrt{-\frac{1}{a}}$. On reconnaît une densité d'une loi normale centrée d'écart-type σ_1 , ce qui impose de prendre $C = \frac{1}{\sigma_1 \sqrt{2\pi}}$ et X_1 suit la loi normale $\mathcal{N}(0, \sigma_1^2)$.

Par un raisonnement analogue, on montre que X_2 suit une loi normale $\mathcal{N}(0, \sigma_2^2)$ pour une certaine valeur de σ_2 .

4. a) On a :

$$g(1) = f_1(1) \times f_2(0) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-1/2\sigma_1^2} \times \frac{1}{\sigma_2 \sqrt{2\pi}}$$

b) Comme on a aussi $g(1) = f_1(0) \times f_2(1)$, il vient :

$$\frac{1}{\sigma_1 \sqrt{2\pi}} e^{-1/2\sigma_1^2} \times \frac{1}{\sigma_2 \sqrt{2\pi}} = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-1/2\sigma_2^2} \times \frac{1}{\sigma_1 \sqrt{2\pi}}$$

Soit $e^{-1/2\sigma_1^2} = e^{-1/2\sigma_2^2}$ et par injectivité de la fonction exponentielle : $\sigma_1^2 = \sigma_2^2$, puis, par positivité $\sigma_1 = \sigma_2$.

Ainsi X_1 et X_2 suivent la même loi normale centrée.

Exercice 3.4.

Un individu gravit un escalier. A chaque fois, avant de faire un pas, il lance une pièce non équilibrée donnant pile avec la probabilité p (avec $0 < p < \frac{1}{2}$) et progresse d'une marche s'il obtient « pile » et enjambe deux marches d'un coup s'il obtient « face ».

1. Pour $n \in \mathbb{N}^*$, soit X_n le nombre de marches gravies à l'issue des n premiers pas et X'_n le nombre de fois ou l'individu a progressé par enjambées de 2 marches au cours des n premiers pas.

- a) Déterminer une relation simple liant X_n et X'_n . En déduire la loi de X_n .
- b) Donner les valeurs de l'espérance et de la variance de X_n .

2. Pour $n \in \mathbb{N}^*$, soit Y_n le nombre aléatoire de pas justes nécessaires pour atteindre ou dépasser la $n^{\text{ème}}$ marche et $E(Y_n)$ l'espérance de Y_n .

- a) Quelles sont les valeurs prises par la variable aléatoire Y_n ?
- b) Déterminer la loi de Y_1 , puis celle de Y_2 et préciser l'espérance de ces deux variables aléatoires.
- c) Montrer que pour tout entier naturel k , et tout entier $n \geq 3$, on a :

$$P(Y_n = k) = p \times P(Y_{n-1} = k - 1) + (1 - p) \times P(Y_{n-2} = k - 1)$$

- d) Montrer que pour $n \geq 3$, $E(Y_n) = p \cdot E(Y_{n-1}) + (1 - p)E(Y_{n-2}) + 1$.

3. On considère l'ensemble E des suites $(u_n)_{n \in \mathbb{N}^*}$ telles que pour tout $n \geq 3$, on ait :

$$u_n = pu_{n-1} + (1 - p)u_{n-2} + 1$$

a) Montrer qu'il existe un réel α , que l'on déterminera, tel que la suite $(\alpha n)_{n \in \mathbb{N}^*}$ appartient à E .

b) Montrer que u appartient à E si et seulement si la suite $v : n \mapsto u_n - \alpha n$ vérifie la relation : $\forall n \in \mathbb{N}^*, v_n = pv_{n-1} + (1 - p)v_{n-2}$.

- c) En déduire la valeur de $E(Y_n)$.

Solution :

1. a) On a : $X_n = n + X'_n$ (n pas font n marches, plus une marche à chaque fois que l'on gravit deux marches d'un coup). Ainsi, $X_n(\Omega) = \llbracket n, 2n \rrbracket$ et si l'on note $q = 1 - p$, comme X'_n suit la loi binomiale $\mathcal{B}(n, q)$:

$$P(X_n = n + k) = \binom{n}{k} q^k p^{n-k}$$

- b) On a évidemment $E(X_n) = n + nq$ et $V(X_n) = npq$.

2. a) On a $Y_n(\Omega) = \llbracket \lfloor \frac{n+1}{2} \rfloor, n \rrbracket$ (pour aller vite on ne fait que des enjambées de deux marches, on peut prendre son temps et monter les marches une par une et toutes les situations intermédiaires sont possibles).

- b) \star Comme $Y_1(\Omega) = \{1\}$, on a $Y_1 = 1$ et $E(Y_1) = 1$.

★ De même $Y_2(\Omega) = \llbracket 1, 2 \rrbracket$, et $P(Y_2 = 2) = p, P(Y_2 = 1) = q$. Ainsi $E(Y_2) = 2p + q = 1 + p$.

c) Soit A l'événement «le premier pas permet de gravir une marche», et B l'événement «le premier pas permet de gravir deux marches». La famille (A, B) forme évidemment un système complet d'événements utile et, pour tout k :

$$\begin{aligned} P(Y_n = k) &= P_A(Y_n = k)P(A) + P_B(Y_n = k)P(B) \\ &= p \times P(Y_{n-1} = k-1) + q \times P(Y_{n-2} = k-1) \end{aligned}$$

d) Les variables aléatoires en jeu étant finies, les espérances existent et :

$$\begin{aligned} \sum_k kP(Y_n = k) &= p \sum_k kP(Y_{n-1} = k-1) + q \sum_k kP(Y_{n-2} = k-1) \\ &= pE(Y_{n-1} + 1) + qE(Y_{n-2} + 1) \\ &= pE(Y_{n-1}) + qE(Y_{n-2}) + 1 \end{aligned}$$

3. a) Si la suite (α_n) vérifie la récurrence, il vient :

$$\alpha_n = p\alpha_{n-1} + q\alpha_{n-2} + 1$$

D'où $\alpha = \frac{1}{1+q}$

b) Il suffit de faire la différence pour montrer cette question.

c) L'équation caractéristique de la suite est $r^2 - pr - q = 0$. Ses racines sont $-q$ et 1 . Ainsi, il existe λ, μ réels tels que pour tout $n \geq 1$:

$$E(Y_n) = \lambda(1)^n + \mu(-q)^n + \frac{n}{1+q}$$

Les calculs de $E(Y_1)$ et $E(Y_2)$ donnent alors :

$$\lambda = \frac{2p^2 - 6p + 3}{(2-p)^2}, \quad \mu = \frac{p^2 - 3p + 1}{(1-p)(2-p)^2}$$

Exercice 3.5.

1. On ouvre un guichet au temps 0. Des clients se présentent à ce guichet aux instants aléatoires T_1, T_2, \dots . On suppose que les variables aléatoires E_1, E_2, \dots définies sur un espace probabilisé (Ω, \mathcal{A}, P) par :

$$\begin{cases} E_1 = T_1 \\ E_2 = T_2 - T_1 \\ E_3 = T_3 - T_2 \\ \dots \end{cases}$$

sont mutuellement indépendantes et suivent une même loi exponentielle de paramètre $\lambda > 0$.

On note X le nombre de clients arrivant dans l'intervalle $[0, 1]$.

- a) Déterminer, pour tout n de \mathbb{N}^* , la loi de T_n .
- b) Calculer $P[X = 0]$.
- c) Montrer que : $\forall n \geq 1, P[X = n] = P[T_{n+1} > 1] - P[T_n > 1]$.
- d) En déduire la loi de X .

2. On pose, pour tout $i \in \mathbb{N}^*$, $E_i = -\frac{1}{\lambda} \ln(U_i)$, où U_1, U_2, \dots sont des variables aléatoires réelles mutuellement indépendantes de même loi uniforme sur $[0, 1]$.

On considère :

$$Y = \begin{cases} 0 & \text{si } U_1 < e^{-\lambda} \\ \min\{n \in \mathbb{N}^* \mid U_1 U_2 \cdots U_n < e^{-\lambda}\} & \text{sinon} \end{cases}$$

- a) Montrer que Y est une variable aléatoire.
- b) Montrer que les variables aléatoires X et Y ont la même loi.

3. En déduire une fonction Pascal qui simule une loi de Poisson de paramètre $\lambda > 0$.

Solution :

1. a) On a : $T_n = E_1 + \dots + E_n$, avec $n \geq 1$. Ainsi T_n suit une loi $\Gamma(\lambda, n)$ comme somme de variables aléatoires indépendantes de même loi $\mathcal{E}(\lambda) = \Gamma(\lambda, 1)$. Une densité de T_n est :

$$f_n(x) = \begin{cases} \frac{x^{n-1}}{(n-1)!} \lambda^n e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

- b) On a : $[X = 0] = [T_1 > 1] = [E_1 > 1]$, donc $P[X = 0] = 1 - (1 - e^{-\lambda}) = e^{-\lambda}$.

c) De même : $[X = n] = [T_n \leq 1 < T_{n+1}]$.

Or :

$$\begin{aligned} [1 < T_{n+1}] &= ([T_n \leq 1] \cup [T_n > 1]) \cap [1 < T_{n+1}] \\ &= ([T_n \leq 1] \cap [1 < T_{n+1}]) \cup ([T_n > 1] \cap [1 < T_{n+1}]) \\ &\hspace{15em} \text{(événements incompatibles)} \end{aligned}$$

Donc :

$$\begin{aligned} P[1 < T_{n+1}] &= P([T_n \leq 1] \cap [1 < T_{n+1}]) + P([T_n > 1] \cap [1 < T_{n+1}]) \\ &= P([T_n \leq 1] \cap [1 < T_{n+1}]) + P([T_n > 1]) \\ &= P[X = n] + P([T_n > 1]) \end{aligned}$$

d'où :

$$\forall n \geq 1, P[X = n] = P[1 < T_{n+1}] - P[1 < T_n].$$

d) Ainsi

$$P[X = n] = \int_1^{+\infty} \frac{x^n}{n!} \lambda^{n+1} e^{-\lambda x} dx - \int_1^{+\infty} \frac{x^{n-1}}{(n-1)!} \lambda^n e^{-\lambda x} dx$$

et, en intégrant par parties :

$$\begin{aligned} P[X = n] &= \left[\frac{e^{-\lambda x}}{-\lambda} \frac{x^n}{n!} \lambda^{n+1} \right]_1^{+\infty} - \int_1^{+\infty} \frac{e^{-\lambda x}}{-\lambda} n \frac{x^{n-1}}{n!} \lambda^{n+1} dx \\ &\quad - \int_1^{+\infty} \frac{x^{n-1}}{(n-1)!} \lambda^n e^{-\lambda x} dx \\ &= e^{-\lambda} \frac{\lambda^n}{n!} \end{aligned}$$

On reconnaît la loi de Poisson de paramètre λ .

2. a) Soit $t \in \mathbb{R}$,

$$[Y \leq t] = \begin{cases} \emptyset & \text{si } t < 0 \\ [U_1 < e^{-\lambda}] & \text{si } 0 \leq t < 1 \\ [U_1 < e^{-\lambda}] \cup [U_1 U_2 < e^{-\lambda}] \cup \dots \cup [U_1 \dots U_n < e^{-\lambda}] & \text{sinon} \end{cases}$$

On a donc pour tout t , $[Y \leq t] \in \mathcal{P}(\Omega)$, ce qui prouve que Y est bien une variable aléatoire.

b) On a :

$$[X = 0] = [T_1 > 1] = [E_1 > 1] = \left[-\frac{1}{\lambda} \ln U_1 > 1 \right] = [U_1 < e^{-\lambda}] = [Y = 0]$$

Donc $P([X = 0]) = e^{-\lambda}$.

Et pour $n \geq 1$:

$$\begin{aligned} [X = n] &= [E_1 + E_2 + \dots + E_n \leq 1 < E_1 + E_2 + \dots + E_n + E_{n+1}] \\ &= \left[-\frac{1}{\lambda} \ln(U_1) - \dots - \frac{1}{\lambda} \ln(U_n) \leq 1 < -\frac{1}{\lambda} \ln(U_1) - \dots - \frac{1}{\lambda} \ln(U_{n+1}) \right] \\ &= [U_1 U_2 \dots U_n \geq e^{-\lambda} > U_1 U_2 \dots U_n U_{n+1}] = [Y = n]. \end{aligned}$$

Donc $\forall n \in \mathbb{N}$, $P[X = n] = P[Y = n]$, les variables aléatoires X et Y suivent donc la même loi.

3. Voici une proposition de fonction Pascal de simulation d'une loi de Poisson de paramètre $\lambda > 0$.

```
function Simul_Poisson_Escp(lambda : real) :integer ;
Var N : integer ;
    P : real ;
Begin
N := 0 ;
P := random ;
```

```

while (P ≥ exp(-lambda)) do
  begin
    P := P*random ;
    N := N+1 ;
  end ;
Simul_Poisson_Escp := N ;
End.

```

Exercice 3.6.

Soit α un réel strictement positif. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , indépendantes et de loi donnée par :

$$\forall n \in \mathbb{N}^*, \quad P(X_n = n^\alpha) = P(X_n = -n^\alpha) = \frac{1}{2}$$

Pour $n \geq 1$, on pose $S_n = \sum_{j=1}^n X_j$ et pour $t \in \mathbb{R}$, on pose $\cosh(t) = \frac{e^t + e^{-t}}{2}$.

On admet que pour tout réel t : $\cosh(t) \leq e^{\frac{t^2}{2}}$.

1. a) Soit β un réel strictement positif. À l'aide d'une comparaison avec une intégrale, montrer que :

$$1^\beta + 2^\beta + \dots + n^\beta \sim \frac{n^{\beta+1}}{\beta + 1}$$

quand n tend vers l'infini.

b) Calculer la variance de S_n et en donner un équivalent quand n tend vers l'infini.

c) Montrer que si $\alpha < \frac{1}{2}$, la suite $(\frac{S_n}{n})_n$ tend vers 0 en probabilité.

2. Soit Y une variable aléatoire discrète à valeurs positives d'espérance $E(Y)$, et a un réel strictement positif. Montrer que :

$$E(Y) \geq a.P([Y \geq a])$$

3. On suppose $\alpha < \frac{1}{2}$. Soient $\varepsilon > 0$ et $\tau > 0$.

a) Montrer que : $P(S_n \geq n\varepsilon) \leq \exp(-n^{1-2\alpha}\varepsilon\tau)E(e^{\tau \frac{S_n}{n^{2\alpha}}})$.

b) Montrer que : $P(S_n \geq n\varepsilon) \leq \exp(-n^{1-2\alpha}\varepsilon\tau) \exp(\frac{\tau^2}{2}n^{1-2\alpha})$.

c) En déduire que : $P(\frac{S_n}{n} \geq \varepsilon) \leq \exp(-\frac{n^{1-2\alpha}\varepsilon^2}{2})$.

Solution :

1.a) On a : $\int_{k-1}^k t^\beta dt \leq k^\beta \leq \int_k^{k+1} t^\beta dt$, puis en sommant :

$$\int_0^n t^\beta dt \leq \sum_1^n k^\beta \leq \int_1^{n+1} t^\beta dt$$

c'est-à-dire : $\frac{n^{\beta+1}}{\beta+1} \leq \sum_{k=1}^n k^\beta \leq \frac{(n+1)^{\beta+1}}{\beta+1} - 1$.

D'où l'équivalent souhaité : $\sum_{k=1}^n k^\beta \sim \frac{n^{\beta+1}}{\beta+1}$

b) On a : $V(S_n) = E(S_n^2)$ car S_n est centrée.

Or $E(S_n^2) = E\left(\sum_{j=1}^n X_j^2 + 2 \sum_{i<j} X_i X_j\right)$, puis par linéarité de l'espérance :

$$\begin{aligned} V(S_n) &= \sum_{j=1}^n E(X_j^2) + 2 \sum_{i<j} E(X_i X_j) = \sum_{j=1}^n \left(\frac{1}{2}(j^\alpha)^2 + \frac{1}{2}(-j^\alpha)^2\right) + 2 \sum_{i<j} 0 \\ &= \sum_{j=1}^n j^{2\alpha} \text{ (en effet par indépendance } E(X_i X_j) = E(X_i)E(X_j) = 0) \end{aligned}$$

D'où $V(S_n) \underset{(n \rightarrow \infty)}{\sim} \frac{n^{2\alpha+1}}{2\alpha+1}$.

c) On peut écrire :

$$P\left(\left|\frac{S_n}{n} - 0\right| > \varepsilon\right) = P(|S_n - 0| > n\varepsilon) \leq \frac{V(S_n)}{n^2 \varepsilon^2}$$

Or $\frac{V(S_n)}{n^2 \varepsilon^2} \sim \frac{n^{2\alpha-1}}{(2\alpha+1)\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$ (car $2\alpha-1 < 0$), ce qui prouve la convergence demandée.

2. Il s'agit de l'inégalité de Markov, qui fait partie du cours.

3. a) On a

$$E\left(\exp\left(\tau \frac{S_n}{n^{2\alpha}}\right)\right) = E\left(\exp\left(\tau \frac{S_n}{n^{2\alpha}}\right)(\mathbf{1}_{S_n > n\varepsilon} + \mathbf{1}_{S_n \leq n\varepsilon})\right) \geq E\left(\exp\left(\tau \frac{S_n}{n^{2\alpha}}\right)\mathbf{1}_{S_n > n\varepsilon}\right)$$

Donc :

$$E\left(\exp\left(\tau \frac{S_n}{n^{2\alpha}}\right)\right) \geq \exp\left(\tau \frac{n\varepsilon}{n^{2\alpha}}\right) E(\mathbf{1}_{S_n > n\varepsilon}) = \exp\left(\tau \frac{n\varepsilon}{n^{2\alpha}}\right) P(S_n \geq n\varepsilon)$$

ce qui est l'inégalité souhaitée à l'ordre des termes près.

b) Par indépendance :

$$E\left(\exp\left(\tau \frac{S_n}{n^{2\alpha}}\right)\right) = E\left(\prod_{j=1}^n \exp\left(\frac{\tau}{n^{2\alpha}} X_j\right)\right) = \prod_{j=1}^n E\left(\exp\left(\frac{\tau}{n^{2\alpha}} X_j\right)\right)$$

Or :

$$E\left(\exp\left(\frac{\tau}{n^{2\alpha}} X_j\right)\right) = \frac{1}{2} \exp\left(\frac{\tau}{n^{2\alpha}} n^\alpha\right) + \frac{1}{2} \exp\left(\frac{\tau}{n^{2\alpha}} (-n^\alpha)\right) = \cosh(\tau n^{-\alpha})$$

$$\leq \exp\left(\frac{(\tau n^{-\alpha})^2}{2}\right)$$

Donc :

$$E\left(\exp\left(\tau \frac{S_n}{n^{2\alpha}}\right)\right) \leq \exp\left(\frac{\tau}{2} n^{-2\alpha+1}\right)$$

c) On a alors : $P\left(\frac{S_n}{n} > \varepsilon\right) \leq \exp\left(-n^{1-2\alpha}\left(\varepsilon\tau - \frac{\tau^2}{2}\right)\right)$

On pose $f(\tau) = \varepsilon\tau - \frac{\tau^2}{2}$. Alors $f(\tau)$ est maximale pour $\tau = \varepsilon$. On a donc bien :

$$P\left(\frac{S_n}{n} > \varepsilon\right) \leq \exp\left(-n^{1-2\alpha}\frac{\varepsilon^2}{2}\right)$$

Exercice 3.7.

On considère une suite de variables aléatoires indépendantes $(X_k)_{k \in \mathbb{N}^*}$, définies sur le même espace probabilisé (Ω, \mathcal{A}, P) et suivant toutes la loi uniforme sur $[0, 1]$.

On pose, pour tout entier n supérieur ou égal à 2, $Z_n = \inf(X_1, \dots, X_n)$, c'est à dire que :

$$\text{pour tout } \omega \in \Omega, Z_n(\omega) = \min(X_1(\omega), \dots, X_n(\omega)).$$

On admet que Z_n est une variable aléatoire définie sur le même espace (Ω, \mathcal{A}, P) .

1. Donner la fonction de répartition et une densité de Z_n .
2. Soit U une variable aléatoire définie sur (Ω, \mathcal{A}, P) , indépendante des X_k et suivant la loi uniforme sur $[0, 1]$.
 - a) Déterminer une densité de $-U$.
 - b) Déterminer une densité g de $Z_n - U$.
 - c) En considérant la variable aléatoire M_n définie par $M_n = \inf(X_2, \dots, X_n)$, déduire de ce qui précède la valeur de $P(Z_n = X_1)$.
Ce résultat était-il prévisible ? La variable aléatoire $Z_n - X_1$ est-elle une variable à densité ? La variable aléatoire $Z_n - X_1$ est-elle discrète ?
3. On pose $W_n = \ln(Z_n)$.
 - a) Déterminer une densité de W_n .
 - b) Soit $(Y_k)_{k \in \mathbb{N}^*}$ une suite de variables indépendantes suivant toutes la loi exponentielle de paramètre 1.
 - i) Déterminer une densité de $T_{n+1} = -\frac{Y_{n+1}}{n+1}$.

ii) On pose, pour $n \geq 1$, $S_n = - \sum_{k=1}^n \frac{Y_k}{k}$.

Montrer que S_n et W_n possèdent la même densité.

c) Montrer, sans calculer explicitement l'intégrale, que

$$\int_{-\infty}^0 x \cdot e^x (1 - e^x)^{n-1} dx = -\frac{1}{n} \sum_{k=1}^n \frac{1}{k}.$$

Solution :

1. On commence classiquement par chercher la fonction d'antirépartition de Z_n :

$P(Z_n \geq x) = P((X_1 \geq x) \cap \dots \cap (X_n \geq x))$ et comme les variables X_1, X_2, \dots, X_n sont supposées indépendantes, on en déduit :

$$P(Z_n \geq x) = (1 - F_1(x))^n.$$

En remplaçant $F_1(x)$ par sa valeur, on obtient :

$$F_{Z_n}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - (1 - x)^n & \text{si } x \in [0, 1] \\ 1 & \text{si } x \geq 1 \end{cases}$$

On obtient une densité en dérivant sur chacun des intervalles ouverts et en rajoutant des valeurs quelconques aux points manquants ; on obtient par exemple :

$$f_{Z_n}(x) = \begin{cases} n(1 - x)^{n-1} & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

2. a) Soit par le calcul, soit directement, $-U$ suit la loi uniforme sur $[-1, 0]$. Une densité de $-U$ est donc la fonction f_{-U} définie par :

$$\forall x \in \mathbb{R}, f_{-U}(x) = \mathbf{1}_{[-1, 0]}(x)$$

b) Comme les variables $-U$ et Z_n sont indépendantes, on peut effectuer un produit de convolution pour déterminer la fonction g :

$$\forall x \in \mathbb{R}, g(x) = \int_{-\infty}^{+\infty} f_{Z_n}(t) f_{-U}(x - t) dt.$$

Dans un premier temps, il reste : $g(x) = \int_0^1 n(1 - t)^{n-1} f_{-U}(x - t) dt$.

On remarque alors que la variable $Z_n - U$ prend ses valeurs dans $[-1, 1]$ et deux cas se présentent :

- si $-1 \leq x \leq 0$, alors $g(x) = \int_0^{x+1} n(1 - t)^{n-1} dt = 1 - (-x)^n$,

• si $0 \leq x \leq 1$, alors $g(x) = \int_x^1 n(1-t)^{n-1} dt = (1-x)^n$.

c) On a $P(Z_n = X_1) = P(X_1 \leq M_n) = P(M_n - X_1 \geq 0)$. On applique alors ce qui précède, où M_n joue le rôle de Z_n et X_1 celui de U .

On obtient donc : $P(Z_n = X_1) = \int_0^1 (1-t)^{n-1} dt = \frac{1}{n}$.

Le résultat est logique pour des raisons d'indépendance et de symétrie.

★ La variable $Z_n - X_1$ n'est pas à densité puisque $P(Z_n - X_1 = 0) \neq 0$.

★ La variable $Z_n - X_1$ n'est pas une variable discrète puisque $(Z_n - X_1)(\Omega)$ n'est pas un ensemble dénombrable.

3. a) On a $W_n(\Omega) =]-\infty, 0]$ et $P(W_n \leq x) = P(\ln Z_n \leq x) = P(Z_n \leq e^x)$.

En conclusion :

$$F_{W_n}(x) = \begin{cases} 1 - (1 - e^x)^n & \text{si } x \leq 0 \\ 1 & \text{si } x \geq 0 \end{cases}$$

On en déduit une densité de W_n :

$$f_{W_n}(x) = \begin{cases} ne^x(1 - e^x)^{n-1} & \text{si } x \leq 0 \\ 0 & \text{si } x > 0 \end{cases}$$

b) i) $P(T_{n+1} \leq x) = P(Y_{n+1} \geq -(n+1)x) = 1 - P(Y_{n+1} \leq -(n+1)x)$.

Ainsi T_{n+1} admet comme fonction de répartition

$$F_{T_{n+1}}(x) = \begin{cases} e^{(n+1)x} & \text{si } x \leq 0 \\ 1 & \text{si } x \geq 0 \end{cases}$$

On en déduit une densité de T_{n+1} :

$$f_{T_{n+1}}(x) = \begin{cases} (n+1)e^{(n+1)x} & \text{si } x \leq 0 \\ 0 & \text{si } x > 0 \end{cases}$$

ii) On montre le résultat par récurrence sur $n \in \mathbb{N}^*$.

Notons h_n une densité de S_n .

★ Pour $n = 1$, on a $S_1 = -Y_1$ et une densité de S_1 est donc : $h_1 : x \mapsto e^x \mathbf{1}_{]-\infty, 0]}(x)$, qui est bien une densité de W_1 .

★ Supposons que pour un certain n , une densité de S_n soit f_{W_n} et déterminons une densité de S_{n+1} .

On a $S_{n+1} = S_n - \frac{Y_{n+1}}{n+1} = S_n + T_{n+1}$. Le lemme des coalitions permet d'affirmer que S_n et T_{n+1} sont indépendantes ; on peut donc effectuer un produit de convolution.

Comme $S_{n+1}(\Omega) = \mathbb{R}^-$, on effectue le calcul pour $x < 0$:

$$f_{S_{n+1}}(x) = \int_{-\infty}^{+\infty} f_{S_n}(t) f_{T_{n+1}}(x-t) dt = \int_x^0 ne^t(1-e^t)^{n-1}(n+1)e^{(n+1)(x-t)} dt$$

$$\begin{aligned}
&= n(n+1)e^{(n+1)x} \int_x^0 e^{-t} (e^{-t} - 1)^{n-1} dt \\
&= n(n+1)e^{(n+1)x} \left[-\frac{1}{n} (e^{-t} - 1)^n \right]_0^x
\end{aligned}$$

Enfin : $f_{S_{n+1}}(x) = (n+1)e^{(n+1)x}(1-e^x)^n$, ce qui est bien une densité de W_{n+1} et ce qui achève la récurrence.

c) Comme les variables Y_k admettent une espérance, la variable S_n admet une espérance et, par linéarité, $E(S_n) = -\sum_{k=1}^n \frac{E(Y_k)}{k} = -\sum_{k=1}^n \frac{1}{k}$.

Comme W_n suit la même loi que S_n , on en déduit que W_n admet une espérance, qui vaut donc $-\sum_{k=1}^n \frac{1}{k}$. Mais $E(W_n) = \int_{-\infty}^{+\infty} x f_{W_n}(x) dx = \int_{-\infty}^0 x n e^x (1-e^x)^{n-1} dx$.

$$\text{On a donc } \int_{-\infty}^0 x e^x (1-e^x)^{n-1} dx = -\frac{1}{n} \sum_{k=1}^n \frac{1}{k}$$

d'où le résultat demandé.

Exercice 3.8.

1. Soit X une variable aléatoire discrète définie sur un espace probabilisé (Ω, \mathcal{A}, P) . Soit h une fonction réelle positive et croissante sur \mathbb{R} .

Montrer que pour tout réel a :

$$P(X \geq a) \leq \frac{E(h(X))}{h(a)}$$

Soit (X_n) une suite de variables aléatoires indépendantes identiquement distribuées suivant la loi de Bernoulli de paramètre $p \in]0, 1[$ (on a donc $P(X = 1) = p$).

Pour tout n de \mathbb{N}^* , on pose :

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$$

2. Déterminer la limite en probabilité de la suite (\bar{X}_n) .

3. Soit a tel que $p < a < 1$.

a) Montrer que pour tout réel $\lambda > 0$:

$$P(\bar{X}_n \geq a) \leq (E(e^{\lambda X_1}))^n e^{-an\lambda}$$

b) En déduire que $P(\bar{X}_n \geq a) \leq e^{-nh_p(a)}$,

où pour tout $x \in]0, 1[$, on a posé : $h_p(x) = x \ln\left(\frac{x}{p}\right) + (1-x) \ln\left(\frac{1-x}{1-p}\right)$.

4. Soit a tel que $0 < a < p$. Deduire de ce qui précède une majoration de $P(\overline{X}_n \leq a)$.

Solution :

1. Par le théorème de transfert, et par croissance de la fonction h :

$$\begin{aligned} E(h(X)) &= \sum_{x_k} h(x_k)P(X = x_k) \geq \sum_{x_k \geq a} h(x_k)P(X = x_k) \\ &\geq h(a) \sum_{x_k \geq a} P(X = x_k) = h(a)P(X \geq a) \end{aligned}$$

D'où le résultat.

2. Nous sommes dans le cas d'utilisation de la loi faible des grands nombres. Ainsi, la suite (\overline{X}_n) converge en probabilité vers $p = E(X)$.

3. a) Soit $h(x) = e^{\lambda x}$. C'est une fonction positive croissante, si $\lambda > 0$. Ainsi par indépendance :

$$\begin{aligned} P(\overline{X}_n \geq a) &= P(n\overline{X}_n \geq na) \leq \frac{E(h(n\overline{X}_n))}{h(na)} = e^{-n\lambda a} E\left(\prod_{i=1}^n e^{\lambda X_i}\right) \\ &\leq e^{-n\lambda a} (E(e^{\lambda X_1}))^n \end{aligned}$$

b) Lorsque X suit la loi de Bernoulli de paramètre p , $E(e^{\lambda X}) = (1-p) + pe^\lambda$. Donc, pour tout $\lambda > 0$

$$P(\overline{X}_n \geq a) \leq e^{-n\lambda a} (1-p + pe^\lambda)^n$$

On étudie donc l'application $\varphi : \lambda \mapsto e^{-n\lambda a} (1-p + pe^\lambda)^n$ sur \mathbb{R}_+^* et on majore $P(\overline{X}_n \geq a)$ par le minimum de φ sur \mathbb{R}_+^* .

Des calculs simples montrent que la dérivée $\varphi'(\lambda)$ s'annule pour λ tel que

$$(1-a)pe^{\lambda(1-a)} = a(1-p)e^{-\lambda a},$$

soit pour :

$$\lambda = \lambda_0 = \ln\left(\frac{a(1-p)}{(1-a)p}\right)$$

et est décroissante sur $]0, \lambda_0[$, croissante sur $[\lambda_0, +\infty[$, ce qui donne le minimum de $\varphi(\lambda_0)$ et :

$$P(\overline{X}_n \geq a) \leq \exp\left[-na \ln\left(\frac{a(1-p)}{(1-a)p}\right)\right] \left(\frac{1-p}{1-a}\right)^n = e^{-nh_p(a)}$$

4. Lorsque $0 < a < p$, alors $1-p < 1-a < 1$ et $Y_n = 1 - X_n$ suit la loi de Bernoulli de paramètre $1-p$. La question précédente donne :

$$P(\overline{Y}_n \geq 1-a) \leq e^{-nh_{1-p}(1-a)}$$

ou :

$$P(\overline{X}_n \leq a) \leq e^{-nh_{1-p}(1-a)} = e^{-nh_p(a)}$$

Exercice 3.9.

Un vendeur de cycles vend des pédales de bicyclette qu'il se procure chez son grossiste par boîtes de deux ; toutes les boîtes sont supposées identiques et dans chaque boîte il y a une pédale droite et une pédale gauche.

★ Lorsqu'un client demande le remplacement de ses deux pédales de vélo, le commerçant lui vend une boîte complète et lui fait payer la somme de $2r$ euros.

★ Lorsqu'un client demande le remplacement d'une seule des deux pédales, le commerçant décide de ne pas obliger le client à acheter une boîte complète, mais majore le prix de la pédale dans une proportion α , c'est-à-dire lui fait payer la somme de $(1 + \alpha)r$ euros.

Pour la simplicité de l'étude, on suppose que l'on sait que le nombre de pédales à poser séparément pendant la durée de l'étude vaut $2n$, où n est un entier naturel non nul. On suppose que le vendeur ne dispose au départ que de boîtes complètes et en nombre suffisant. Soit p la probabilité qu'une demande d'un client qui ne demande qu'une pédale corresponde à une pédale droite (p n'est pas nécessairement égal à $1/2$) et X le nombre de boîtes nécessaires à la satisfaction de ces $2n$ demandes. (le commerçant n'ouvre une boîte que s'il ne dispose pas d'une boîte entamée lui permettant d'accéder à la demande du client)

1. Quelle est la loi de X ? On précisera l'ensemble des valeurs prises par X .
2. Montrer que X peut s'écrire : $X = a + |Y - b|$, où a et b sont des constantes qu'on précisera et Y une variable aléatoire qui suit une loi binomiale.

Donner l'expression l'espérance de $E(X)$ en fonction de n et p .

Dans la suite, on prendra la valeur $p = 1/2$.

3. Quelle majoration α le marchand de cycles doit-il appliquer au prix de chaque pédale vendue séparément pour qu'en moyenne le prix de vente des $2n$ pédales vendues séparément soit égal au prix de vente des X boîtes nécessaires vendues $2r$ euros chacune.

La valeur α trouvée dépend de n et on la note dorénavant α_n . Prouver que la suite (α_n) est décroissante. Donner un équivalent simple de α_n et la limite de α_n lorsque n tend vers $+\infty$.

[On admettra la formule de Stirling : $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$]

Solution :

1. Soit X_1 le nombre de pédales droites demandées (et X_2 le nombre de pédales gauches). On a $X_1 + X_2 = 2n$ et X_1 et X_2 suivent la loi binomiale de paramètres $2n$ et p .

Or $X = \max(X_1, X_2)$. D'où : $X(\Omega) = \llbracket n, 2n \rrbracket$, et la loi de X est donnée par :

$$\begin{cases} P(X = n) = P(X_1 = n) \\ \forall k \in \llbracket n + 1, 2n \rrbracket, P(X = k) = P(X_1 = k) + P(X_1 = 2n - k) \end{cases}$$

Soit, avec $q = 1 - p$:

$$\begin{cases} P(X = n) = \binom{2n}{n} p^n q^n \\ \forall k \in \llbracket n + 1, 2n \rrbracket, P(X = k) = \binom{2n}{k} [p^k q^{2n-k} + p^{2n-k} q^k] \end{cases}$$

2. $\max(a, b) = \frac{1}{2}(a + b) + \frac{1}{2}|a - b|$, donc

$$X = \frac{1}{2}(X_1 + X_2) + \frac{1}{2}|X_1 - X_2| = n + |X_1 - n|$$

On a :

$$E(X) = n \binom{2n}{n} p^n q^n + \sum_{k=n+1}^{2n} k \binom{2n}{k} (p^k q^{2n-k} + q^k p^{2n-k})$$

Ainsi pour $p = q = \frac{1}{2}$:

$$E(X) = n \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} + 2n \sum_{k=n+1}^{2n} \binom{2n-1}{k-1} \left(\frac{1}{2}\right)^{2n-1}$$

Et comme $\binom{2n-1}{k} = \binom{2n-1}{2n-1-k}$

$$\begin{aligned} E(X) &= n \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} + n \left(\frac{1}{2}\right)^{2n-1} \sum_{k=0}^{2n} \binom{2n-1}{k-1} \\ &= n \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} + n \left(\frac{1}{2}\right)^{2n-1} 2^{2n-1} \end{aligned}$$

$$E(X) = n \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} + n$$

3. Ainsi on veut $2rE(X) = 2nr(1 + \alpha_n)$, d'où :

$$\alpha_n = \frac{1}{n} E(X) - 1 = \left(\frac{1}{2}\right)^{2n} \binom{2n}{n}$$

On a alors, en revenant aux factorielles :

$$\frac{\alpha_n}{\alpha_{n-1}} = \frac{2n(2n-1)}{4n^2} \leq 1$$

Donc la suite (α_n) est décroissante.

De plus, la formule de Stirling donne après quelques simplifications :

$$\alpha_n = \frac{(2n)!}{(n!)^2} \times \frac{1}{2^{2n}} \sim \frac{1}{\sqrt{\pi n}}$$

et en particulier $\lim_{n \rightarrow \infty} \alpha_n = 0$.

Exercice 3.10.

On note n et r deux entiers vérifiant $n \geq 2$ et $r \geq 3$.

On considère une épreuve aléatoire pouvant aboutir à r résultats différents R_1, R_2, \dots, R_r de probabilités respectives x_1, x_2, \dots, x_r telles que, pour tout i de $\llbracket 1, r \rrbracket$, $0 < x_i < 1$.

On effectue n épreuves indépendantes du type de celle décrite ci-dessus. Pour tout i de $\llbracket 1, r \rrbracket$, on note X_i la variable aléatoire réelle qui vaut 1 si R_i n'est pas obtenu à l'issue de ces n épreuves et qui vaut 0 sinon.

On désigne par X la variable aléatoire réelle égale au nombre total de résultats qui n'ont pas été obtenus à l'issue des n épreuves.

1. a) Que vaut $\sum_{i=1}^r x_i$?

b) Exprimer la variable aléatoire X en fonction de X_1, X_2, \dots, X_r .

c) Donner la loi de X_i pour tout i de $\llbracket 1, r \rrbracket$.

d) En déduire que l'espérance de X est donnée par :

$$E(X) = \sum_{i=1}^r (1 - x_i)^n$$

2. On considère l'application f de $\Omega =]0, 1[)^r$ dans \mathbb{R} définie par :

$$\forall (x_1, x_2, \dots, x_r) \in \Omega, f(x_1, x_2, \dots, x_r) = \sum_{i=1}^r (1 - x_i)^n$$

a) Montrer que Ω est un ouvert convexe de \mathbb{R}^r .

b) Démontrer que f est de classe C^2 sur Ω et déterminer la hessienne $\nabla^2 f(x_1, x_2, \dots, x_r)$ de f , pour tout (x_1, x_2, \dots, x_r) de Ω .

c) Vérifier que la forme quadratique associée à la hessienne au point (x_1, x_2, \dots, x_r) de Ω prend des valeurs strictement positives sur les vecteurs non nuls.

3. On cherche les extremums de f sous la contrainte linéaire $(C) : x_1 + \dots + x_r = 1$.

a) Montrer que f admet pour cette optimisation sous contrainte un unique point critique A que l'on déterminera.

b) En utilisant l'égalité de Taylor-Lagrange à l'ordre 1, montrer que f présente en A un minimum global sous la contrainte (C) .

c) Donner la valeur de $E(X)$ correspondant à ce minimum.

Solution :

1. a) La famille R_1, R_2, \dots, R_r constituant un système complet d'événements, on a $\sum_{i=1}^r x_i = 1$.

b) On a : $X = \sum_{i=1}^r X_i$.

c) X_i est une variable de Bernoulli, de paramètre $P(X_i = 1) = (1 - x_i)^n$.

d) Donc : $E(X) = \sum_{i=1}^r E(X_i) = \sum_{i=1}^r (1 - x_i)^n$.

2. a) Ω est un ouvert convexe de \mathbb{R}^r en tant que produit d'ouverts convexes de \mathbb{R} (c'est un pavé ouvert).

b) La fonction f en tant que fonction polynôme des x_i , est de classe C^∞ sur Ω , avec :

$$\forall i \in \llbracket 1, r \rrbracket, \frac{\partial f}{\partial x_i}(x) = -n(1 - x_i)^{n-1}$$

puis :

$$\forall (i, j) \in \llbracket 1, r \rrbracket^2, \frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \begin{cases} 0 & \text{si } i \neq j \\ n(n-1)(1 - x_i)^{n-2} & \text{si } i = j \end{cases}$$

On en déduit :

$$\nabla^2 f(x) = n(n-1) \begin{pmatrix} (1-x_1)^{n-2} & 0 & \dots & 0 \\ 0 & (1-x_2)^{n-2} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & (1-x_r)^{n-2} \end{pmatrix}$$

c) Notons q_α la forme quadratique associée à la hessienne $\nabla^2 f(x)$ au point $x = (x_1, x_2, \dots, x_r)$. Alors :

$$\forall (h_1, h_2, \dots, h_r) \in \mathbb{R}^r, q_\alpha(h_1, h_2, \dots, h_r) = \sum_{i=1}^r n(n-1)(1-x_i)^{n-2} h_i^2$$

qui est strictement positif si les h_i ne sont pas tous nuls.

3. a) Si l'on pose :

$$\mathcal{H} = \left\{ H = (h_1, \dots, h_r) / \sum_{i=1}^r h_i = 0 \right\}$$

les points critiques A sous la contrainte donnée sont les solutions éventuelles du système :

$$\sum_{i=1}^r x_i = 1 \text{ et } \forall H \in \mathcal{H}, \langle \nabla f(A), H \rangle = 0$$

ou encore :

$$\sum_{i=1}^r x_i = 1 \text{ et } \nabla f(A) \in \mathcal{H}^\perp = \text{Vect}((1, \dots, 1))$$

ce qui indique qu'il existe $\lambda \in \mathbb{R}$ tel que :

$$\begin{cases} -n(1 - x_i)^{n-1} + \lambda = 0 \text{ si } i \in \llbracket 1, r \rrbracket \\ x_1 + \dots + x_r = 1 \end{cases}$$

dont l'unique solution est $A = \left(\frac{1}{r}, \dots, \frac{1}{r}\right)$.

b) Si $H = (h_1, \dots, h_r)$ est élément de \mathcal{H} , grâce à la convexité de Ω , le segment $[A, A + H]$ est inclus dans Ω , alors, d'après l'égalité de Taylor-Lagrange à l'ordre 1, il existe un réel $\theta \in]0, 1[$ tel que :

$$f(A + H) = f(A) + \langle \nabla f(A), H \rangle + \frac{1}{2} q_{A+\theta H}(H)$$

Comme $H \in \mathcal{H}$, on a $\langle \nabla f(A), H \rangle = 0$, donc :

$$f(A + H) - f(A) = \frac{1}{2} q_{A+\theta H}(H)$$

quantité positive d'après ce qui précède. Ainsi, f présente bien en A un minimum global sous la contrainte \mathcal{C} .

c) En ce point, la valeur minimum de l'espérance est :

$$E(X) = \sum_{i=1}^r \left(1 - \frac{1}{r}\right)^n = r \left(1 - \frac{1}{r}\right)^n = \frac{(r-1)^n}{r^{n-1}}$$

Exercice 3.11.

1. Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice de rang 1.

a) Montrer qu'il existe deux matrices colonnes U et V de $\mathcal{M}_{n,1}(\mathbb{R})$, telles que $M = U^t V$.

b) On pose $\lambda = {}^t V U$. Montrer que si $\lambda \neq 0$, alors $\frac{1}{\lambda} M$ est la matrice d'un projecteur dont on précisera l'image et le noyau.

2. Soit X et Y deux variables aléatoires définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , et à valeurs dans $\mathbb{N}_n = \{1, 2, \dots, n\}$. Soit $A = (a_{i,j})$ la matrice carrée d'ordre n , de terme général :

$$\forall (i, j) \in \mathbb{N}_n^2, a_{i,j} = P_{[Y=j]}(X = i)$$

a) Montrer que, pour tout $j \in \mathbb{N}_n$, on a : $\sum_{i=1}^n a_{i,j} = 1$.

b) Soit f l'endomorphisme de \mathbb{R}^n canoniquement associé à A . On pose :

$$B = \begin{pmatrix} P(X = 1) \\ P(X = 2) \\ \vdots \\ P(X = n) \end{pmatrix}$$

Montrer que $B \in \text{Im}(f)$

c) Montrer que X et Y sont indépendantes si, et seulement si, le rang de A est égal à 1.

Solution :

1. a) Le rang de M vaut 1, donc il existe une matrice colonne $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ (non nulle) telle que pour tout j de $\llbracket 1, n \rrbracket$, la $j^{\text{ème}}$ colonne C_j de M soit de la forme $\lambda_j C$, (l'un au moins des λ_j étant non nul). Ainsi on a :

$$M = (\lambda_1 C \quad \lambda_2 C \quad \dots \quad \lambda_n C) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} (\lambda_1 \quad \dots \quad \lambda_n)$$

On peut donc prendre $U = C$ et $V = {}^t(\lambda_1 \quad \dots \quad \lambda_n)$.

b) On pose $\lambda = {}^t V U$ (λ appartient à \mathbb{R}). On suppose $\lambda \neq 0$, alors :

$$\left(\frac{1}{\lambda} M\right)^2 = \frac{1}{\lambda^2} (U^t V)(U^t V) = \frac{1}{\lambda^2} U ({}^t V U) {}^t V = \frac{1}{\lambda^2} \lambda U {}^t V = \frac{1}{\lambda} M$$

donc $\frac{1}{\lambda} M$ est la matrice d'un projecteur. Son image est la droite engendrée par la colonne $C = U$ et son noyau est l'hyperplan d'équation $M X = 0$, soit $U^t V X = 0$, et comme $U \neq 0$ et ${}^t V X \in \mathbb{R}$, cela équivaut à ${}^t V X = 0$.

2. a) Soit $j \in \llbracket 1, n \rrbracket$, on a :

$$\sum_{i=1}^n a_{i,j} = \sum_{i=1}^n P_{[Y=j]}(X = i) = \sum_{i=1}^n \frac{P(X = i \cap Y = j)}{P(Y = j)} = \frac{\sum_{i=1}^n P(X = i \cap Y = j)}{P(Y = j)}$$

Or les événements $(X = i)_{1 \leq i \leq n}$ forment un système complet d'événements, donc :

$$\sum_{i=1}^n P(X = i \cap Y = j) = P(Y = j)$$

D'où

$$\sum_{i=1}^n a_{i,j} = \frac{P(Y = j)}{P(Y = j)} = 1$$

b) D'après la formule des probabilités totales, on a :

$$\forall i \in \llbracket 1, n \rrbracket, P(X = i) = \sum_{j=1}^n P_{[Y=j]}(X = i)P(Y = j) = \sum_{j=1}^n a_{i,j}P(Y = j)$$

Autrement dit :

$$\begin{pmatrix} P(X = 1) \\ P(X = 2) \\ \vdots \\ P(X = n) \end{pmatrix} = A \begin{pmatrix} P(Y = 1) \\ P(Y = 2) \\ \vdots \\ P(Y = n) \end{pmatrix}$$

ce qui prouve le résultat demandé.

c) \star Supposons X et Y indépendantes. Alors :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, P_{[Y=j]}(X = i) = P(X = i)$$

Donc toutes les colonnes de A sont égales à : $\begin{pmatrix} P(X = 1) \\ P(X = 2) \\ \vdots \\ P(X = n) \end{pmatrix}$ qui est non nulle

puisque $\sum_{i=1}^n P(X = i) = 1$. La matrice A est bien de rang 1.

\star Réciproquement, supposons le rang de A égal à 1. Alors $\text{Im}(A)$ est une droite vectorielle. D'après la question précédente, le vecteur $U =$

$\begin{pmatrix} P(X = 1) \\ P(X = 2) \\ \vdots \\ P(X = n) \end{pmatrix}$ engendre alors $\text{Im}(A)$. Donc :

$$\forall j \in \llbracket 1, n \rrbracket, \exists \lambda_j \in \mathbb{R}, \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix} = \lambda_j \begin{pmatrix} P(X = 1) \\ P(X = 2) \\ \vdots \\ P(X = n) \end{pmatrix}$$

D'après la question 2.a, on a pour tout $j \in \llbracket 1, n \rrbracket$:

$$1 = \sum_{i=1}^n a_{i,j} = \sum_{i=1}^n \lambda_j P(X = i) = \lambda_j \sum_{i=1}^n P(X = i) = \lambda_j$$

Donc pour tout $j \in \llbracket 1, n \rrbracket$, $\lambda_j = 1$. Par suite, on a :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, P_{[Y=j]}(X = i) = a_{i,j} = \lambda_j P(X = i) = P(X = i)$$

ce qui revient à écrire :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, P(X = i \cap Y = j) = P(X = i)P(X = j)$$

Les variables X et Y sont donc indépendantes.

Exercice 3.12.

Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{A}, P) , continue, de densité f et de fonction de répartition F . On note U une variable aléatoire réelle de loi uniforme sur $[0, 1]$.

1. Soit $\alpha \in \mathbb{R}_+^*$. Montrer que $-\frac{1}{\alpha} \ln(U)$ est une variable aléatoire réelle dont on déterminera la loi.

2. On suppose dans cette question seulement que f n'est jamais nulle.

a) Montrer que F est alors une bijection de \mathbb{R} sur $]0, 1[$. Est-il possible de lever l'hypothèse « f jamais nulle » en conservant le caractère bijectif de F ?

b) Montrer que $F^{-1}(U)$ est une variable aléatoire de densité f .

3. Pour tout réel $x \in]0, 1[$, on note :

$$I_x = \{a \in \mathbb{R}, F(a) \geq x\} \text{ et } F^{-1}(x) = \{t \in \mathbb{R} \mid F(t) = x\}$$

a) Montrer que I_x est une demi-droite fermée à gauche.

b) Montrer l'existence de $G(x) = \inf I_x$ et de $\inf F^{-1}(x)$, puis montrer leur égalité.

c) Montrer que : $\forall t \in \mathbb{R}, \forall x \in [0, 1], G(x) \leq t \implies F(t) \geq x$.

d) En déduire que les variables aléatoires X et $G(U)$ ont la même loi.

Solution :

1. Soit $Y = -\frac{1}{\alpha} \ln(U)$.

Pour tout $t \in \mathbb{R}$, $[Y \leq t] = [\ln U \geq -\alpha t] = [U \geq e^{-\alpha t}] \in \mathcal{P}(\Omega)$, car U est une variable aléatoire définie sur Ω . Donc Y est bien une variable aléatoire, et :

$$P[Y \leq t] = P[U \geq e^{-\alpha t}] = \begin{cases} 1 - e^{-\alpha t} & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases}$$

donc $Y \hookrightarrow \mathcal{E}(\alpha)$.

2. a) La fonction F est dérivable et $F' = f$ qui est à valeurs strictement positives. F est une application continue et strictement croissante de \mathbb{R} dans $F(\mathbb{R}) =]0, 1[$.

[0 ne peut être atteint sinon il existerait un réel x_0 tel que $F(x_0) = 0$ et la croissance de F entraînerait $\forall x \leq x_0, F(x) = 0$ et donc $\forall x < x_0, F'(x) = f(x) = 0$ ce qui est impossible. De la même manière 1 ne peut être atteint.] Ainsi F est une bijection croissante de \mathbb{R} dans $F(\mathbb{R}) =]0, 1[$.

Si on suppose seulement $f \geq 0$:

La fonction F est croissante au sens large car sa dérivée f est positive ou nulle. Si F était constante sur un intervalle $[a, b]$ avec $a < b$, alors sa dérivée f serait nulle sur cet intervalle. f ne peut donc s'annuler qu'en des points isolés si on veut conserver le caractère bijectif de F .

b) Pour tout $t \in \mathbb{R}$, $[F^{-1}(U) \leq t] = [U \leq F(t)]$ et $P([U \leq F(t)]) = F(t)$. Ainsi $F^{-1}(U)$ est donc une variable aléatoire de fonction de répartition F et de densité f .

3. a) Soit $a_1 \in I_x$, $F(a_1) \geq x \implies \forall a_2 \geq a_1, F(a_2) \geq F(a_1)$ (F croissante), donc I_x est une demi-droite.

F étant une fonction numérique continue, l'image réciproque du fermé $[x, +\infty[$ est un fermé, donc I_x est une demi-droite fermée à gauche.

b) • Existence : $\forall x \in]0, 1[$:

$$\lim_{a \rightarrow +\infty} F(a) = 1 \implies \exists a_1, F(a_1) \geq x, \text{ donc } I_x \text{ est non vide.}$$

$\lim_{a \rightarrow -\infty} F(a) = 0 \implies \exists a_2, \forall a \leq a_2, F(a) < x$, donc I_x est minorée par a_2 .

Ainsi I_x , partie non vide et minorée de \mathbb{R} , admet une borne inférieure.

Le théorème des valeurs intermédiaires sur $[a_2, a_1]$ prouve que $F^{-1}\{x\}$ est non vide, minorée par a_2 ; c'est donc aussi une partie non vide et minorée de \mathbb{R} , il admet une borne inférieure, d'où l'existence de $\inf F^{-1}\{x\}$.

• Égalité : on a : $F^{-1}\{x\} \subset I_x \implies G(x) \leq \inf F^{-1}\{x\}$.

$\alpha = \inf F^{-1}\{x\} \implies \forall a < \alpha, F(a) < x$ (F croissante) $\implies \alpha \leq G(x)$, d'où :

$$\inf F^{-1}\{x\} = G(x)$$

c) Pour tout $t \in \mathbb{R}$, pour tout $x \in [0, 1]$,

$$G(x) \leq t \iff \inf F^{-1}\{x\} \leq t \iff F(t) \geq x$$

d) $[G(U) \leq t] = [F(t) \geq U]$, donc :

$$P[G(U) \leq t] = P[F(t) \geq U] = F(t) = P[X \leq t].$$

Les variables aléatoires X et $G(U)$ suivent donc la même loi.

Exercice 3.13.

Dans un gratte-ciel, un ascenseur n'assure que la descente. Il part du sommet à l'étage n , et à chaque fermeture de la porte, il descend pour s'arrêter aléatoirement à un étage strictement inférieur jusqu'à ce qu'il parvienne au rez-de-chaussée (étage 0) : on suppose qu'à chaque fois le numéro de l'étage

d'arrêt suit une loi uniforme sur l'ensemble des numéros des étages encore accessibles.

On note $A(p, n)$ la probabilité que, lors de sa descente, l'ascenseur s'arrête à l'étage p , avec $0 \leq p < n$.

1. Calculer $A(0, n), A(n - 1, n), A(n - 2, n)$.
2. Démontrer, en utilisant la position du premier arrêt, la relation :

$$\mathcal{R}(n) : \forall p \in \llbracket 0, n - 2 \rrbracket, A(p, n) = \frac{1}{n} \left[1 + \sum_{j=p+1}^{n-1} A(p, j) \right]$$

3. En utilisant $\mathcal{R}(n)$ et $\mathcal{R}(n - 1)$, en déduire que :

$$\forall p \in \llbracket 0, n - 3 \rrbracket, A(p, n) = A(p, n - 1).$$

4. Déterminer $A(p, n)$ pour $0 \leq p < n$.
5. Soit p tel que $0 \leq p \leq n - 1$; on note E_p la variable aléatoire valant 1 si lors de sa descente, l'ascenseur s'arrête à l'étage p et 0 sinon.

a) Déterminer la loi de E_p . Les variables aléatoires E_0, E_1, \dots, E_{n-1} sont-elles 2 à 2 indépendantes ?

b) Soit X la variable aléatoire égale au nombre d'arrêts lors de la descente. Exprimer l'espérance $E(X)$ et la variance $V(X)$ de X .

Solution :

1. $\star A(0, n) = 1$: à force de descendre, on arrive au rez-de-chaussée !
- $\star A(n - 1, n) = \frac{1}{n}$: on descend directement d'un étage.
- $\star A(n - 2, n) = \frac{1}{n} + \frac{1}{n} \times \frac{1}{n - 1} = \frac{1}{n - 1}$: on descend directement de deux étages, ou bien 2 fois d'un étage.

2. On a :

$$A(p, n) = \sum_{j=0}^{n-1} P(X_1 = j) \times P(\text{un arrêt à l'étage } p / \text{premier arrêt à l'étage } j)$$

Or :

$$P(\text{un arrêt à l'étage } p / \text{premier arrêt à l'étage } j) = \begin{cases} 1 & \text{si } p = j \\ 0 & \text{si } p > j \\ A(p, j) & \text{si } p < j \end{cases}$$

d'où :

$$\forall p \in \llbracket 0, n - 2 \rrbracket, A(p, n) = \frac{1}{n} \left[1 + \sum_{j=p+1}^{n-1} A(p, j) \right]$$

3. Par le même procédé en partant de l'étage $n - 1$, on obtient :

$$\forall p \in \llbracket 0, n - 3 \rrbracket, A(p, n - 1) = \frac{1}{n - 1} \left[1 + \sum_{j=p+1}^{n-2} A(p, j) \right]$$

Ainsi :

$$1 + \sum_{j=p+1}^{n-2} A(p, j) = (n - 1)A(p, n - 1) = nA(p, n) - A(p, n - 1)$$

Donc :

$$A(p, n) = A(p, n - 1), \text{ pour tout } p \text{ de } \llbracket 0, n - 3 \rrbracket$$

4. Pour tout $p \in \llbracket 0, n - 3 \rrbracket$, $A(p, n) = A(p, n - 1)$, d'où par récurrence :

$$\forall p \in \llbracket 0, n - 3 \rrbracket, A(p, n) = A(p, p + 3).$$

Donc :

$$\forall p \in \llbracket 0, n - 3 \rrbracket, A(p, n) = A(p, p + 3) = \frac{1}{p + 1}.$$

Or pour $p \in \{n - 2, n - 1\}$, on a déjà $A(p, n) = \frac{1}{p + 1}$. D'où :

$$\forall p \in \llbracket 0, n - 1 \rrbracket, A(p, n) = \frac{1}{p + 1}$$

5. $\star P([E_p = 1]) = A(n, p) = \frac{1}{p + 1}$, donc $E_p \hookrightarrow \mathcal{B}\left(\frac{1}{p + 1}\right)$.

\star Soient i et j tels que $0 \leq i < j \leq n - 1$. Alors :

$$\begin{aligned} P([E_i = 1] \cap [E_j = 1]) &= P([E_i = 1] / [E_j = 1]) \cdot P([E_j = 1]) \\ &= A(j, i) A(n, j) = \frac{1}{i + 1} \times \frac{1}{j + 1} \\ &= P([E_i = 1]) \cdot P([E_j = 1]) \end{aligned}$$

d'où l'indépendance.

b) On a : $X = E_0 + E_1 + \dots + E_{n-1}$. Ainsi par indépendance pour le calcul de la variance

- $E(X) = E(E_0) + E(E_1) + \dots + E(E_{n-1}) = 1 + \frac{1}{2} + \dots + \frac{1}{n}$.
- $V(X) = V(E_0) + V(E_1) + \dots + V(E_{n-1}) = 1 \times 0 + \frac{1}{2} \times \frac{1}{2} + \frac{1}{3} \times \frac{2}{3} + \dots + \frac{1}{n} \times \frac{n-1}{n}$

$$E(X) = \sum_{k=1}^n \frac{1}{k} \quad ; \quad V(X) = \sum_{k=1}^n \frac{k-1}{k^2}$$

Exercice 3.14.

Soient U , V , W trois variables aléatoires indépendantes définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , telles que U et W suivent une loi de Poisson de paramètre $\lambda > 0$, et V suit une loi de Poisson de paramètre $\mu > 0$. On pose : $X = U + V$ et $Y = V + W$.

1. Déterminer les lois de X et de Y (on redémontrera ce résultat du cours).
2. a) Montrer que $\text{Cov}(X, Y)$ existe et la calculer.
 b) En déduire le coefficient de corrélation linéaire de X et de Y .
3. Soit n un entier naturel.
 a) Déterminer la loi conditionnelle de V sachant $[X = n]$.
 b) En déduire que l'espérance conditionnelle de Y sachant $[X = n]$ est égale à $\lambda + \frac{n\mu}{\lambda + \mu}$.
 c) Montrer que cette espérance est supérieure ou égale à n si et seulement si on a : $E(X) \geq n$.
4. On suppose que la taille d'un individu est la somme de deux variables aléatoires indépendantes, l'une représentant l'effet du patrimoine génétique, l'autre celui du mode de vie, et que ces variables aléatoires sont indépendantes et suivent des lois de Poisson de paramètres respectifs λ et μ . La variable aléatoire X désigne alors la taille d'un père et Y celle de son fils. Donner une interprétation du résultat de la question 3.

Solution :

1. $U + V$ prend ses valeurs dans \mathbb{N} et pour tout n de \mathbb{N} , on a par indépendance :

$$\begin{aligned} P(U + V = n) &= \sum_{i=0}^n P((U = i) \cap (V = n - i)) = \sum_{i=0}^n P(U = i)P(V = n - i) \\ &= \sum_{i=0}^n e^{-\lambda} \frac{\lambda^i}{i!} \times e^{-\mu} \frac{\mu^{n-i}}{(n-i)!} = \frac{e^{-(\lambda+\mu)}}{n!} \sum_{i=0}^n \binom{n}{i} \lambda^i \mu^{n-i} \\ &= \frac{e^{-(\lambda+\mu)}}{n!} (\lambda + \mu)^n \end{aligned}$$

Ainsi $X = U + V$ suit la loi de Poisson de paramètre $\lambda + \mu$, idem pour $V + W$.

2. a) Comme X et Y admettent chacune une espérance, $\text{Cov}(X, Y)$ existe si et seulement si XY a une espérance.

Or : $XY = UV + UW + VW + V^2$ et U, V, W sont indépendantes admettant une espérance, donc les produits deux à deux de ces variables admettent une espérance (valant le produit des espérances) et V admet une variance, donc un moment d'ordre deux. Bref $E(XY)$ existe et :

$$\begin{aligned} E(XY) &= E(U)E(V) + E(U)E(W) + E(V)E(W) + \text{Var}(V) + E(V)^2 \\ &= \lambda\mu + \lambda^2 + \lambda\mu + \mu + \mu^2 = (\lambda + \mu)^2 + \mu \end{aligned}$$

D'où :

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = (\lambda + \mu)^2 + \mu - (\lambda + \mu)^2 = \mu$$

b) Comme, par indépendance : $\text{Var}(X) = \text{Var}(Y) = \lambda + \mu$, il vient :

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mu}{\lambda + \mu}$$

3. a) Si $(X = n)$ est réalisé, alors V peut prendre les valeurs $0, 1, \dots, n$ et :

$$\begin{aligned} P_{(X=n)}(V = k) &= \frac{P((X = n) \cap (V = k))}{P(X = n)} = \frac{P((X = n) \cap (U = n - k))}{P(X = n)} \\ &= \frac{P(X = n)P(U = n - k)}{P(X = n)} = \frac{e^{-\mu} \frac{\mu^k}{k!} e^{-\lambda} \frac{\lambda^{n-k}}{(n-k)!}}{e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^n}{n!}} \\ P_{(X=n)}(V = k) &= \binom{n}{k} \left(\frac{\mu}{\lambda + \mu}\right)^k \left(\frac{\lambda}{\lambda + \mu}\right)^{n-k} \end{aligned}$$

Autrement dit, la loi de V conditionnellement à la réalisation de l'événement $(X = n)$ est la loi binomiale de paramètres n et $\frac{\mu}{\lambda + \mu}$.

b) Ainsi $E_{(X=n)}(V) = \frac{n\mu}{\lambda + \mu}$.

Comme W est indépendante de X , on a $E_{(X=n)}(W) = E(W) = \lambda$, et par linéarité :

$$E_{(X=n)}(Y) = \lambda + \frac{n\mu}{\lambda + \mu}$$

c) $\lambda + \frac{n\mu}{\lambda + \mu} \geq n \iff \lambda^2 + \lambda\mu + n\mu \geq n\lambda + n\mu \iff \lambda + \mu \geq n$.

4. Dans une telle population, il y a une tendance à la stabilisation des tailles autour de la moyenne dans le sens où les fils nés de pères plus petits que la moyenne sont en moyenne plus grands que leur père, alors que les fils nés de pères plus grands que la moyenne sont en moyenne plus petits que leur père.

Exercice 3.15.

Une urne contient R boules rouges et B boules blanches. On pose $N = R + B$. On effectue une suite de tirages au hasard d'une boule de cette urne selon le processus suivant :

- lorsqu'on tire une boule rouge, celle-ci est enlevée de l'urne et remplacée par une boule blanche, avant de passer au tirage suivant.
- lorsqu'on tire une boule blanche, celle-ci est enlevée de l'urne et remplacée par une boule rouge, avant de passer au tirage suivant.

L'expérience est modélisée par un espace probabilisé (Ω, \mathcal{A}, P) .

Pour tout entier n de \mathbb{N}^* , on note A_n l'événement «le $n^{\text{ème}}$ tirage amène une boule rouge». Soit R_n la variable aléatoire représentant le nombre de boules rouges contenues dans l'urne à l'issue du $n^{\text{ème}}$ tirage.

1. Soit $n \geq 0$. Déterminer la probabilité conditionnelle $P_{[R_n=k]}(A_{n+1})$ de A_{n+1} conditionnellement à l'événement $[R_n = k]$.

En déduire $P(A_{n+1})$ en fonction de l'espérance $E(R_n)$

2. a) Soit $\mathbf{1}_{n+1}$ la fonction indicatrice de A_{n+1} . Exprimer R_{n+1} en fonction de R_n et $\mathbf{1}_{n+1}$.

b) En déduire une relation de récurrence sur la suite $(E(R_n))_n$.

c) Exprimer $E(R_n)$ en fonction de n , $E(R_0)$ et de N . En déduire la valeur de $\lim_{n \rightarrow +\infty} E(R_n)$.

3. a) Soit $Z_n = R_n \cdot \mathbf{1}_{n+1}$. Déterminer la loi de Z_n en fonction de la loi de R_n .

b) En déduire que $E(Z_n) = \frac{E(R_n^2)}{N}$.

Solution :

1. Si $[R_n = k]$ est un événement non quasi-impossible et est réalisé, alors il y a k boules rouges dans l'urne avant le tirage et :

$$P_{[R_n=k]}(A_{n+1}) = \frac{k}{N}$$

En ne conservant en fait que les événements non quasi-impossibles, on a :

$$P(A_{n+1}) = \sum_{k=0}^N P(A_{n+1} \cap [R_n = k]) = \sum_{k=0}^N \frac{k}{N} P([R_n = k]) = \frac{1}{N} E(R_n)$$

2. a) Si on obtient une boule blanche, alors on ajoute une boule rouge et si on obtient une boule rouge on enlève une boule rouge, soit :

$$R_{n+1} = R_n + 1 - 2 \cdot \mathbf{1}_{n+1}$$

b) Comme $E(\mathbf{1}_{n+1}) = \frac{1}{N} E(R_n)$, il vient par linéarité de l'espérance :

$$E(R_{n+1}) = \left(1 - \frac{2}{N}\right) E(R_n) + 1$$

c) On est en présence d'une suite arithmético-géométrique, de point fixe $\frac{N}{2}$ et on obtient classiquement :

$$E(R_n) = \left(E(R_0) - \frac{N}{2}\right) \left(1 - \frac{2}{N}\right)^n + \frac{N}{2}; \quad \lim_{n \rightarrow \infty} E(R_n) = \frac{N}{2}$$

3. a) \star Si $k \in \llbracket 1, N \rrbracket$, on a $[Z_n = k] = [R_n = k] \cap [\mathbf{1}_{n+1} = 1] = [R_n = k] \cap A_{n+1}$ et donc par le résultat de la question 1. :

$$P(Z_n = k) = P(R_n = k) \times \frac{k}{N}$$

\star Donc :

$$P(Z_n = 0) = 1 - \sum_{k=1}^N P(Z_n = k) = 1 - \frac{1}{N} \sum_{k=0}^N kP(R_n = k) = 1 - \frac{1}{N} E(R_n)$$

$$\text{b) Ainsi } E(Z_n) = \sum_{k=0}^N kP(Z_n = k) = \frac{1}{N} \sum_{k=1}^N k^2 P(R_n = k) = \frac{1}{N} E(R_n^2).$$

Exercice 3.16.

Sur le cercle de centre O et de rayon 1, on considère un point A fixe et un point B choisi aléatoirement (la suite du texte précisera le sens à donner à cette expression). On note L la variable aléatoire égale à la distance de O à la corde AB : L est une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{B}, P) prenant ses valeurs dans $[0, 1]$.

On considère la mesure θ de l'angle $(\overrightarrow{OA}, \overrightarrow{OB})$ prise dans $[0, 2\pi[$.

1. Exprimer L en fonction de θ .

2. a) Montrer que la restriction de la fonction \cos à $[0, \pi]$ admet une fonction réciproque qui est dérivable sur $] -1, 1[$.

b) On note \arccos cette fonction réciproque. Calculer pour tout $x \in] -1, 1[$, la dérivée $\arccos'(x)$.

3. On suppose que θ est une variable aléatoire sur l'espace probabilisé (Ω, \mathcal{B}, P) , suivant une loi uniforme sur $[0, 2\pi[$. Donner la fonction de répartition F de L et montrer que L admet pour densité la fonction f définie par :

$$f(x) = \begin{cases} \frac{2}{\pi} \times \frac{1}{\sqrt{1-x^2}} & \text{si } x \in [0, 1[\\ 0 & \text{si } x < 0 \text{ ou } x \geq 1 \end{cases}$$

4. Démontrer que pour tout $k \in \mathbb{N}$, L^k admet une espérance que l'on notera m_k .

5. a) Calculer m_0 et m_1 .

b) Trouver une relation entre m_{k+2} et m_k . En déduire les valeurs de m_{2k} et m_{2k+1} pour tout $k \in \mathbb{N}$.

c) Déterminer $\lim_{k \rightarrow +\infty} \frac{m_k}{m_{k+1}}$.

d) Étudier la suite $(S_k)_{k \in \mathbb{N}}$ définie par $S_k = (k+1)m_k m_{k+1}$, et en déduire un équivalent de m_k lorsque $k \rightarrow +\infty$.

Solution :

1. Soit H le milieu de $[A, B]$, le triangle OHA est rectangle en H et $\frac{OH}{OA} = |\cos(\frac{\theta}{2})|$, soit :

$$L = OH = |\cos(\frac{\theta}{2})|$$

2. a) La restriction de la fonction \cos au segment $[0, \pi]$ réalise une bijection strictement décroissante de $[0, \pi]$ sur $[-1, 1]$. Comme la dérivée de la fonction \cos est non nulle sur $]0, \pi[$, la bijection réciproque Arc cos est dérivable sur $] -1, 1[$.

b) Avec :
$$\text{Arc cos}'(t) = \frac{1}{\cos'(\text{Arc cos } t)} = -\frac{1}{\sin(\text{Arc cos } t)}$$

Or sur le domaine utile, la fonction \sin est positive et :

$$\sin(\text{Arc cos } t) = \sqrt{1 - \cos^2(\text{Arc cos } t)} = \sqrt{1 - t^2}$$

soit :

$$\text{Arc cos}'(t) = -\frac{1}{\sqrt{1 - t^2}}$$

3. L prend ses valeurs dans $[0, 1]$, et pour $x \in [0, 1]$:

$$F_L(x) = P(L \leq x) = P(|\cos(\theta/2)| \leq x) = P(-x \leq \cos(\theta/2) \leq x)$$

et par décroissance de la fonction \cos sur le domaine considéré :

$$\begin{aligned} F_L(x) &= P(\text{Arc cos } x \leq \frac{\theta}{2} \leq \text{Arc cos}(-x)) \\ &= P(2 \text{Arc cos } x \leq \theta \leq \pi - 2 \text{Arc cos}(x)) = \frac{1}{2\pi}(\pi - 4 \text{Arc cos } x) \end{aligned}$$

F_L est bien de classe \mathcal{C}^1 sur $[0, 1[$ et L est une variable à densité. Par dérivation on peut alors prendre pour densité la fonction f définie par :

$$f(x) = \begin{cases} \frac{2}{\pi} \times \frac{1}{\sqrt{1 - x^2}} & \text{si } x \in [0, 1[\\ 0 & \text{si } x < 0 \text{ ou } x \geq 1 \end{cases}$$

4. L est une variable aléatoire bornée, elle admet donc des moments de tous ordres.

5. a) $\star m_0 = \frac{2}{\pi} \int_0^1 \frac{dx}{\sqrt{1 - x^2}} = \frac{2}{\pi} [-\text{Arc cos } x]_0^1 = 1,$

$\star m_1 = \frac{2}{\pi} \int_0^1 \frac{x}{\sqrt{1 - x^2}} dx = \frac{2}{\pi} [-\sqrt{1 - x^2}]_0^1 = \frac{2}{\pi}.$

b) On écrit : $m_{k+2} = \frac{2}{\pi} \int_0^1 x^{k+1} \times \frac{2x}{2\sqrt{1 - x^2}} dx$ et on procède à l'intégration par parties ainsi préparée :

$$\begin{aligned}
m_{k+2} &= \frac{2}{\pi} [x^{k+1}(-\sqrt{1-x^2})]_0^1 + \frac{2}{\pi} (k+1) \int_0^1 x^k \sqrt{1-x^2} dx \\
&= \frac{2}{\pi} (k+1) \int_0^1 x^k \frac{(1-x^2)}{\sqrt{1-x^2}} dx = (k+1)(m_k - m_{k+2})
\end{aligned}$$

D'où :

$$m_{k+2} = \frac{k+1}{k+2} m_k$$

On en déduit, par récurrence :

$$m_{2k} = \frac{2k-1}{2k} \times \frac{2k-3}{2k-2} \times \cdots \times \frac{1}{2} \times 1; \quad m_{2k+1} = \frac{2k}{2k+1} \times \frac{2k-2}{2k-1} \times \cdots \times \frac{2}{3} \times \frac{2}{\pi}$$

On peut, si on le souhaite, « compacter » ces écritures à l'aide de factorielles, ce qui donne :

$$m_{2k} = \frac{(2k)!}{2^{2k} (k!)^2}; \quad m_{2k+1} = \frac{(2^k k!)^2}{(2k+1)!} \times \frac{2}{\pi}$$

c) Par encadrement des fonctions à intégrer, on a pour tout k de \mathbb{N}^* :

$$0 < m_{k+1} \leq m_k \leq m_{k-1}$$

Comme $\lim_{k \rightarrow \infty} \frac{m_{k-1}}{m_{k+1}} = \lim_{k \rightarrow \infty} \frac{k+1}{k} = 1$, il vient, par encadrement :

$$\lim_{k \rightarrow +\infty} \frac{m_k}{m_{k+1}} = 1$$

d) La relation de récurrence donne : $(k+2)m_{k+2}m_{k+1} = (k+1)m_{k+1}m_k$. Ainsi la suite (S_k) est constante égale à $S_0 = m_0m_1 = \frac{2}{\pi}$.

Comme $m_k \sim m_{k+1}$, il vient : $km_k^2 \sim \frac{2}{\pi}$, et par positivité de m_k :

$$m_k \underset{(\infty)}{\sim} \sqrt{\frac{2}{k\pi}}$$

Exercice 3.17.

On considère une suite d'épreuves de Bernoulli indépendantes.

On note p , ($0 < p < 1$) la probabilité d'un succès S et $q = 1 - p$ la probabilité d'un échec E . On appelle séquence de succès de longueur r , toute suite ininterrompue de r succès consécutifs. On considère la variable aléatoire T égale au nombre d'épreuves précédant la première séquence de succès de longueur 3. La suite d'épreuves s'arrête alors au troisième succès de cette séquence. Par exemple l'événement élémentaire $\{SSESSS\}$ appartient à l'événement $[T = 3]$. On admet que T est une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) et on note pour tout entier k , $t_k = P(T = k)$.

1. Calculer les probabilités t_0 , t_1 , t_2 et t_3 .

2. Pour tout $n \in \mathbb{N}^*$, on note q_n la probabilité de l'événement A_n : « n'obtenir aucune séquence de succès de longueur 3 dans une suite de n épreuves ». Exprimer, pour $n \geq 1$, t_n en fonction de q_{n-1} .

3. En posant $q_0 = 1$, montrer que, pour tout $n \geq 3$:

$$q_n = q(q_{n-1} + pq_{n-2} + p^2q_{n-3})$$

Calculer les nombres q_0, q_1, q_2 et q_3 .

4. Étudier la monotonie de la suite (q_n) . Quelle est la valeur la plus probable de T ?

5. a) Montrer que pour tout s réel tel que $|s| < 1$, la série de terme général $q_n s^n$ est convergente.

On pose alors $F(s) = \sum_{n=0}^{+\infty} q_n s^n$.

b) Montrer que :
$$F(s) = \frac{1 + ps + p^2 s^2}{1 - qs(1 + ps + p^2 s^2)}$$

c) En admettant que $\sum_{n=0}^{+\infty} q_n = \lim_{s \rightarrow 1} F(s)$, retrouver la valeur de la somme

de la série $\sum_{k=0}^{+\infty} t_k$.

Solution :

1. $t_0 = P(SSS) = p^3$, $t_1 = P(ESSS) = qp^3$, $t_2 = P(.ESSS) = qp^3$, où le point signifie que le résultat du premier essai est sans intérêt.

De même $t_3 = P(..ESSS) = qp^3$ (même remarque).

2. Réaliser $[T = n]$, c'est ne pas avoir trois succès consécutifs au cours des $(n - 1)$ premières épreuves, avoir un échec au rang n et enfin trois succès consécutifs. Par indépendance, il vient :

$$t_n = q_{n-1}qp^3$$

3. Pour $n \geq 3$, on regarde ce qui se passe au début. Réaliser A_n c'est :

→ Soit commencer par un échec, puis ne pas avoir de triplet de succès au cours des $(n - 1)$ épreuves suivantes.

→ Commencer par un succès suivi d'un échec, puis ne pas avoir de triplet de succès au cours des $(n - 2)$ épreuves suivantes.

→ Commencer par deux succès suivi d'un échec, puis ne pas avoir de triplet de succès au cours des $(n - 3)$ épreuves suivantes.

Donc, par incompatibilité et indépendance :

$$\forall n \geq 3, q_n = qq_{n-1} + pq_{n-2} + p^2q_{n-3}$$

Comme $q_0 = q_1 = q_2 = 1$, il vient alors :

$$q_3 = q(1 + p + p^2) = 1 - p^3 \text{ (ce qui se voit directement)}$$

4. Comme $A_n \subset A_{n-1}$, la suite (q_n) est décroissante. La valeur la plus probable de T est donc clairement 0.

5. a) Pour $|s| < 1$, on a : $0 \leq |q_n s^n| \leq |s|^n$ et la convergence de la série géométrique de terme général $|s|^n$ donne la convergence (absolue) de la série définissant $F(s)$.

b) On écrit :

$$\begin{aligned} F(s) &= \sum_{n=0}^{\infty} q_n s^n = 1 + s + s^2 + \sum_{n=3}^{\infty} q(q_{n-1} + pq_{n-2} + p^2q_{n-3})s^n \\ &= 1 + s + s^2 + qs \sum_{n=3}^{\infty} q_{n-1} s^{n-1} + pqs^2 \sum_{n=3}^{\infty} q_{n-2} s^{n-2} + qp^2s^3 \sum_{n=3}^{\infty} q_{n-3} s^{n-3} \\ &= 1 + s + s^2 + qs(F(s) - 1 - s) + pqs^2(F(s) - 1) + qp^2s^3F(s) \end{aligned}$$

et en développant et en regroupant :

$$F(s) = \frac{1 + ps + p^2s^2}{1 - qs(1 + ps + p^2s^2)}$$

c) En admettant :

$$\sum_{n=0}^{\infty} q_n = \frac{1 + p + p^2}{1 - q(1 + p + p^2)} = \frac{1 - p^3}{1 - p - q(1 - p^3)} = \frac{1 - p^3}{qp^3}$$

On en déduit :

$$\sum_{k=0}^{\infty} t_k = p^3 + qp^3 \frac{1 - p^3}{qp^3} = 1$$

On retrouve ainsi le fait que T est bien une variable aléatoire, puisque définie « presque partout ».

Exercice 3.18.

Soit $I =]1/2, 1[$, p un paramètre réel de I et $q = 1 - p$. Soit X une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans $\{2, 4, 6, \dots, 2x, \dots\}$, telle que pour tout $x \in \mathbb{N}^*$, on a :

$$P([X = 2x]) = \begin{cases} 1 - 2pq & \text{si } x = 1 \\ (2pq)^2(1 - 2pq)^{x-2} & \text{si } x \geq 2 \end{cases}$$

Pour $N \in \mathbb{N}^*$, on désigne par (X_1, \dots, X_N) un N -échantillon i.i.d. de la loi de X .

Soit $(2x_1, \dots, 2x_N)$ un N -échantillon de réalisations du N -échantillon (X_1, \dots, X_N) tel que, pour tout k de $\llbracket 1, N \rrbracket$, $x_k \in \mathbb{N}^*$.

On pose : $\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k$, et on désigne par m le nombre de réalisations égales à 2.

On suppose que : $\bar{x} > \max(1; 4(1 - \frac{m}{N}))$, et on note : $E = \{k \in \llbracket 1, N \rrbracket \mid x_k \geq 2\}$.

1. Montrer que $\sum_{k \in E} x_k = N\bar{x} - m$.

2. Soit L et H les deux fonctions définies sur I par

$$L(p) = P\left(\bigcap_{k=1}^N [X_k = 2x_k]\right) \text{ et } H = \ln \circ L$$

a) Donner, pour tout p de I , l'expression de $L(p)$.

b) Montrer que H est de classe C^1 sur I .

c) Montrer que H admet un maximum sur I en un point noté \hat{p} .

Solution :

1. On a : $N\bar{x} = \sum_{k=1}^N x_k = \sum_{k \in E} x_k + \sum_{k \notin E} x_k = \sum_{k \in E} x_k + m \times 1$, soit :

$$\sum_{k \in E} x_k = N\bar{x} - m$$

2. a) On a :

$$\begin{aligned} L(p) &= \prod_{k=1}^N P([X_k = 2x_k]) = \prod_{k \notin E} P([X_k = 2]) \times \prod_{k \in E} P([X_k = 2x_k]) \\ &= (1 - 2pq)^m \times (2pq)^{2(N-m)} (1 - 2pq)^{\sum x_k - 2(N-m)} \\ L(p) &= (2pq)^{2(N-m)} (1 - 2pq)^{N\bar{x} - 2(N-m)} \end{aligned}$$

b) En développant :

$$H(p) = 2(N - m) \ln(2p(1 - p)) + (N\bar{x} - 2(N - m)) \ln(1 - 2p(1 - p))$$

Or $p \in]1/2, 1[\implies 0 < p(1 - p) < \frac{1}{4}$ et H est clairement de classe C^1 .

$$c) H'(p) = 2(N - m) \frac{1 - 2p}{p(1 - p)} + (N\bar{x} - 2(N - m)) \frac{-2 + 4p}{1 - 2p(1 - p)}$$

$$\text{Soit : } H'(p) = \frac{2(1 - 2p)}{pq(1 - 2pq)} ((N - m)(1 - 2pq) + (2(N - m) - N\bar{x})pq)$$

Ou encore :

$$H'(p) = \frac{2(1-2p)}{pq(1-2pq)}((N-m) - N\bar{x}p(1-p))$$

Comme $p > \frac{1}{2}$, le signe de $H'(p)$ est l'opposé de celui de $p^2 - p + \frac{N-m}{N\bar{x}}$.

Ce trinôme du second degré a deux racines de somme 1 et de produit $N-m$ positif. Ainsi ces deux racines sont entre 0 et 1, et on retient la racine supérieure à $1/2$:

$$\hat{p} = \frac{1}{2} + \frac{1}{2} \sqrt{1 - \frac{4(N-m)}{N\bar{x}}}$$

qui donne le tableau de variations ($H'(p)$ est positif entre les racines) :

p	$1/2$		\hat{p}		1
H	$-N\bar{x} \ln 2$	\nearrow	$H(\hat{p})$	\searrow	$-\infty$

Exercice 3.19.

Soit n un nombre entier supérieur ou égal à 1. On a invité n personnes à une conférence mais certains invités ne pourront pas venir ; on appelle « auditoire » l'ensemble des personnes qui viennent. On note p_k la probabilité que cet auditoire soit formé de k personnes fixées et on suppose que p_k ne dépend que de k , et pas des personnes composant cet auditoire. On suppose aussi que, pour tout $k \in \llbracket 0, n \rrbracket$, un auditoire de k personnes est k fois plus probable qu'un auditoire d'une seule personne *i.e.* $p_k = kp_1$.

1. Combien y a-t-il d'auditoires différents possibles ? Combien comportent k personnes ?

2. Montrer que $p_1 = \frac{1}{n2^{n-1}}$. Déterminer la loi du nombre X de personnes qui viennent.

3. Quelle est la probabilité qu'un invité donné soit bien présent ?

4. Montrer que, avec cette modélisation, les événements « l'invité a est présent » et « l'invité b est présent » ne sont pas indépendants.

5. Les invités qui viendront ont prévenu le conférencier qui a réservé une salle qui comporte exactement le bon nombre de sièges. Une personne qui n'était pas invitée décide de venir aussi ; elle a autant de chances de trouver un siège que les personnes invitées.

Déterminer la probabilité q_n que cette personne reste debout et déterminer un équivalent simple de q_n quand n tend vers $+\infty$.

Solution :

1. Comme $\text{card } \mathcal{P}(E) = 2^{\text{card } E}$, il y a 2^n auditoires possibles, dont $\binom{n}{k}$ auditoires formés de k personnes, pour $k \in \llbracket 0, n \rrbracket$.

2. Il y a $\binom{n}{k}$ auditoires de k personnes, chacun de probabilité $p_k = kp_1$, donc par partition de l'univers des possibles :

$$1 = \sum_{k=0}^n \binom{n}{k} p_k = p_1 \sum_{k=0}^n k \binom{n}{k} = p_1 \sum_{k=1}^n k \binom{n}{k}$$

(Notons que l'on a $p_0 = 0$, il est donc quasi-impossible que le conférencier se retrouve seul !)

Pour $k \geq 1$, on a $k \binom{n}{k} = n \binom{n-1}{k-1}$ et donc $\sum_{k=1}^n k \binom{n}{k} = n \sum_{k=1}^n \binom{n-1}{k-1} = n2^{n-1}$ et donc :

$$p_1 = \frac{1}{n2^{n-1}}$$

Ainsi :

$$X(\Omega) = \llbracket 1, n \rrbracket \text{ et } \forall k \in \llbracket 1, n \rrbracket, P(X = k) = \binom{n}{k} p_k = \frac{k}{n2^{n-1}} \binom{n}{k}$$

3. La probabilité qu'un invité donné « I » soit présent est la somme des probabilités de tous les auditoires où figure I .

Il y a $\binom{n-1}{k-1}$ auditoires de k personnes contenant I et ainsi la probabilité P_I cherchée vaut :

$$\begin{aligned} P_I &= \sum_{k=1}^n \binom{n-1}{k-1} p_k = p_1 \sum_{k=1}^n k \binom{n-1}{k-1} = p_1 \sum_{k=1}^n (k-1+1) \binom{n-1}{k-1} \\ &= p_1 \left(\sum_{k=2}^n (n-1) \binom{n-2}{k-2} + \sum_{k=1}^{n-1} \binom{n-1}{k-1} \right) = p_1 ((n-1)2^{n-2} + 2^{n-1}) \\ p_I &= \frac{(n-1)2^{n-2} + 2^{n-1}}{n2^{n-1}} = \frac{n+1}{2n} \end{aligned}$$

4. De même la probabilité que deux individus « a » et « b » donnés soient présents est donnée par

$$\begin{aligned} P_{a,b} &= \sum_{k=2}^n \binom{n-2}{k-2} p_k = p_1 \left(\sum_{k=0}^{n-2} k \binom{n-2}{k} + 2 \sum_{k=0}^{n-2} \binom{n-2}{k} \right) \\ &= \frac{(n-2)2^{n-3} + 2 \cdot 2^{n-2}}{n2^{n-1}} = \frac{n+2}{4n} \end{aligned}$$

et :

$$P_{a,b} - P_a P_b = \frac{n+2}{4n} - \left(\frac{n+1}{2n} \right)^2 = -\frac{1}{4n^2} \neq 0$$

Ce qui prouve la non-indépendance des événements.

5. La probabilité que la personne en question reste debout, sachant que l'auditoire comporte k personnes, est $\frac{1}{k+1}$. Par la formule des probabilités totales, il vient alors :

$$\begin{aligned} q_n &= p_1 \sum_{k=0}^n \frac{1}{k+1} \times k \binom{n}{k} = p_1 \left(\sum_{k=0}^n \binom{n}{k} - \sum_{k=0}^n \frac{1}{k+1} \binom{n}{k} \right) \\ &= p_1 \left(2^n - \frac{1}{n+1} \sum_{k=0}^n \binom{n+1}{k+1} \right) = \frac{1}{n2^{n-1}} \left(2^n - \frac{2^{n+1}-1}{n+1} \right) \\ q_n &= \frac{2}{n} - \frac{4}{n(n+1)} + \frac{1}{n(n+1)2^{n-1}} \underset{(\infty)}{\sim} \frac{2}{n} \end{aligned}$$

Exercice 3.20.

1. Soit t un réel tel que $t \notin \{-1, 1\}$. Déterminer deux réels $A(t)$ et $B(t)$ tels que pour tout réel $u \geq 0$:

$$\frac{1}{(1+u)(1+t^2u)} = \frac{A(t)}{1+u} + \frac{B(t)}{1+t^2u}$$

2. Soit g la fonction définie sur $\mathbb{R}^* \setminus \{-1, 1\}$ par : $g(t) = \frac{2}{\pi^2} \times \frac{\ln |t|}{t^2 - 1}$.

a) Montrer que g se prolonge par continuité en -1 et 1 .

b) Montrer que l'intégrale $\int_0^{+\infty} \frac{\ln t}{t^2 - 1} dt$ est convergente.

3. Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{A}, P) dont une densité f est définie sur \mathbb{R} par :

$$f(x) = \frac{1}{\pi(1+x^2)}$$

Soit Y une variable aléatoire indépendante de X , de même loi que X . On pose $T = \frac{X}{Y}$, et on admet qu'une densité f_T de T est donnée, pour tout t réel non nul, par :

$$f_T(t) = \int_{-\infty}^{+\infty} |y| f(y) f(ty) dy$$

a) Montrer que pour tout réel t non nul :

$$f_T(t) = \frac{2}{\pi^2} \int_0^{+\infty} \frac{y}{(1+y^2)(1+t^2y^2)} dy$$

b) En déduire l'expression de f_T .

c) En déduire la valeur de l'intégrale $\int_0^{+\infty} \frac{\ln t}{t^2 - 1} dt$.

Solution :

1. Par réduction au même dénominateur et identification, il vient :

$$A(t) = -\frac{1}{t^2 - 1} \quad ; \quad B(t) = \frac{t^2}{t^2 - 1}$$

2. a) La fonction g est paire et $\ln t \underset{(1)}{\sim} t - 1 \implies g(t) \underset{(1)}{\sim} \frac{2}{\pi^2} \times \frac{t-1}{t^2-1} = \frac{2}{\pi^2} \times \frac{1}{t+1}$, d'où $\lim_{t \rightarrow 1} g(t) = \frac{1}{\pi^2}$ et donc $\lim_{t \rightarrow -1} g(t) = \frac{1}{\pi^2}$.

On pose donc : $g(-1) = g(1) = \frac{1}{\pi^2}$.

b) La fonction h à intégrer est bien entendu implicitement supposée prolongée par continuité en 1. Cette fonction est positive et :

★ $\lim_{t \rightarrow +\infty} t^{3/2} h(t) = 0$ (négligeabilité classique) et la règle de Riemann

montre que $\int_1^{+\infty} h(t) dt$ converge.

★ $h(t) \underset{(0^+)}{\sim} -\ln t$ et la convergence connue de l'intégrale $\int_0^1 \ln t dt$ donne celle de $\int_0^1 h(t) dt$.

L'intégrale proposée converge.

3. a) Le résultat admis donne, par parité :

$$f_T(t) = \int_{-\infty}^{+\infty} |y| \times \frac{1}{\pi} \times \frac{1}{1+y^2} \times \frac{1}{\pi} \times \frac{1}{1+t^2 y^2} dy = \frac{1}{\pi^2} \int_0^{+\infty} \frac{2y}{(1+y^2)(1+t^2 y^2)} dy$$

b) Le changement de variable $u = y^2$ est de classe \mathcal{C}^1 et strictement monotone, donc légitime et donne :

$$f_T(t) = \frac{1}{\pi^2} \int_0^{+\infty} \frac{du}{(1+u)(1+t^2 u^2)}$$

Or, par le résultat de la question 1. et pour $t \notin \{-1, 1\}$:

$$\int_0^A \frac{du}{(1+u)(1+t^2 u^2)} = \frac{1}{t^2 - 1} \left[\ln \left(\frac{1+t^2 u}{1+u} \right) \right]_0^A = \frac{1}{t^2 - 1} \ln \left(\frac{1+t^2 A}{1+A} \right)$$

et donc pour $t \neq 0$: $\int_0^{+\infty} \frac{du}{(1+u)(1+t^2 u^2)} = \frac{1}{t^2 - 1} \ln(t^2)$, soit :

$$\forall t \notin \{-1, 0, 1\}, f_T(t) = \frac{2}{\pi^2} \times \frac{\ln |t|}{t^2 - 1}$$

et on peut décider de poser $f_T(-1) = f_T(1) = \frac{1}{\pi^2}$.

c) Comme T est une variable paire, on a : $\int_0^{+\infty} f(t) dt = \frac{1}{2}$ et ainsi :

$$\int_0^{+\infty} \frac{\ln t}{t^2 - 1} dt = \frac{\pi^2}{4}$$

Exercice 3.21.

On suppose que toutes les variables aléatoires considérées sont définies sur le même espace probabilisé (Ω, \mathcal{A}, P) . La lettre a désigne un réel strictement positif donné.

1. Soit f la fonction définie sur \mathbb{R} par :

$$f(x) = \begin{cases} (a+1)x^a & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

a) Vérifier que f est une densité de probabilité.

b) Soit X une variable aléatoire admettant f comme densité. Calculer l'espérance $E(X)$ de X .

2. Soit $n \in \mathbb{N}^*$. On considère la fonction g_n définie par :

$$g_n(x) = \begin{cases} \frac{(a+1)^n x^a}{\Gamma(n)} (-\ln x)^{n-1} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

a) Montrer l'existence de l'intégrale $\int_0^1 x^a (-\ln x)^{n-1} dx$.

b) À l'aide du changement de variable $u = -\ln x$, montrer que g_n est une densité de probabilité.

Soit T_n une variable aléatoire réelle de densité g_n .

c) Vérifier que T_n admet une espérance et calculer $E(T_n)$.

3. Soit X_1 et X_2 deux variables aléatoires indépendantes, suivant la même loi que X . On pose $Z = \frac{X_1}{X_2}$.

a) Donner une densité des variables aléatoires Y_1 et Y_2 définies par $Y_1 = \ln X_1$ et $Y_2 = -\ln X_2$.

b) En déduire une densité de la variable $T = \ln Z$.

c) Soit h la fonction définie par :
$$h(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{a+1}{2}x^a & \text{si } 0 < x < 1 \\ \frac{a+1}{2}x^{-(a+2)} & \text{si } x \geq 1 \end{cases}$$

Montrer que h est une densité de Z .

4. Soit $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant toutes la même loi que X . On pose $W_n = \prod_{k=1}^n X_k$.

Montrer que la variable aléatoire W_n admet g_n comme densité.

Solution :

1. a) La fonction f est positive, continue par morceaux et

$$\int_{\mathbb{R}} f = \int_0^1 (a+1)x^a dx = [x^{a+1}]_0^1 = 1$$

f est une densité de probabilité

b) $E(X) = \int_0^1 (a+1)x^{a+1} dx = \frac{a+1}{a+2}$.

2. a) La fonction à intégrer est continue sur $]0, 1]$ et comme $a > 0$, elle est prolongeable par continuité en 0 en posant $g_n(0) = 0$ (limite classique), l'intégrale est « faussement impropre » ...

b) ... et le changement de variable $u = -\ln x$ est de classe \mathcal{C}^1 strictement monotone, donc légitime et donne :

$$\int_0^1 x^a (-\ln x)^{n-1} dx = \int_{+\infty}^0 (e^{-u})^a u^{n-1} (-e^{-u}) du = \int_0^{+\infty} u^{n-1} e^{-(a+1)u} du$$

La connaissance des lois Gamma montre alors que g_n , qui est positive et continue par morceaux est bien une densité de probabilité.

c) T_n est une variable aléatoire bornée, elle admet donc une espérance et :

$$\begin{aligned} E(T_n) &= \int_0^1 \frac{(a+1)x^{a+1}}{\Gamma(n)} (-\ln x)^{n-1} dx = \int_0^{+\infty} \frac{(a+1)^n}{\Gamma(n)} e^{-(a+2)u} u^{n-1} du \\ &= \frac{(a+1)^n}{(a+2)^n} \int_0^{+\infty} \frac{(a+2)^n}{\Gamma(n)} e^{-(a+2)u} u^{n-1} du \end{aligned}$$

et en reconnaissant encore une intégrale de densité d'une loi Γ :

$$E(T_n) = \frac{(a+1)^n}{(a+2)^n}$$

3. a) $\star Y_1$ prend ses valeurs dans \mathbb{R}^- , et pour $y < 0$:

$$F_{Y_1}(y) = P(Y_1 \leq y) = P(\ln X_1 \leq y) = P(X_1 \leq e^y) = F_{X_1}(e^y) = (e^y)^{a+1}$$

On peut donc prendre pour densité de Y_1 la fonction f_1 définie par :

$$f_1(y) = (a+1)e^{(a+1)y} \text{ si } y \leq 0 \text{ et } f_1(y) = 0 \text{ si } y > 0$$

\star On obtient symétriquement et en notant f_2 une densité de Y_2 :

$$f_2(y) = (a+1)e^{-(a+1)y} \text{ si } y \geq 0 \text{ et } f_2(y) = 0 \text{ si } y < 0$$

b) X_1 et X_2 sont indépendantes, donc Y_1 et Y_2 aussi et, par convolution, une densité f_T de T est définie par :

$$f_T(x) = \int_{-\infty}^{+\infty} f_1(x-t)f_2(t)dt$$

et les expressions de f_1 et f_2 obligent à distinguer deux cas :

\star Si $x \leq 0$,

$$\begin{aligned} f_T(x) &= \int_0^{+\infty} (a+1)e^{(a+1)(x-t)}(a+1)e^{-(a+1)t}dt \\ &= (a+1)^2 e^{(a+1)x} \int_0^{+\infty} e^{-2(a+1)t}dt = \frac{1}{2}(a+1)e^{(a+1)x} \end{aligned}$$

\star Si $x > 0$,

$$\begin{aligned} f_T(x) &= \int_x^{+\infty} (a+1)e^{(a+1)(x-t)}(a+1)e^{-(a+1)t}dt \\ &= (a+1)^2 e^{(a+1)x} \int_x^{+\infty} e^{-2(a+1)t}dt = \frac{1}{2}(a+1)e^{-(a+1)x} \end{aligned}$$

c) Z prend ses valeurs dans \mathbb{R}^+ et pour $x > 0$,

$$F_Z(x) = P(Z \leq x) = P(T \leq \ln x) = F_T(\ln x)$$

On peut donc prendre, toujours pour $x > 0$, $h(x) = \frac{1}{x}f_T(\ln x)$, soit :

$$h(x) = \frac{1}{2}(a+1)x^a, \text{ si } 0 < x < 1 \text{ et } h(x) = \frac{1}{2}(a+1)x^{-(a+2)} \text{ si } x \geq 1$$

4. W_n prend ses valeurs dans $[0, 1]$ et pour $x \in]0, 1[$:

$$\begin{aligned} F_{W_n}(x) &= P(W_n \leq x) = P(\ln W_n \leq \ln x) = P(-\ln W_n \geq -\ln x) \\ &= 1 - P\left(\sum_{k=1}^n -\ln X_k \leq -\ln x\right) \end{aligned}$$

Chaque variable aléatoire $-\ln(X_k)$ suit la loi exponentielle $\mathcal{E}(a+1)$, donc, par indépendance, leur somme suit la loi Gamma de paramètres $\frac{1}{a+1}$ et n .

Par dérivation de la fonction de répartition d'une telle loi, on peut alors prendre :

$$\forall x \in]0, 1[, f_{W_n}(x) = \frac{1}{x} \frac{(a+1)^n}{\Gamma(n)} e^{(a+1) \ln x} (-\ln x)^{n-1}$$

Exercice 3.22.

Un sondage consiste à proposer l'affirmation « A » à certaines personnes d'une population donnée. Le sujet abordé étant délicat, le stratagème suivant est mis en place afin de mettre en confiance les personnes sondées pour qu'elles ne mentent pas.

L'enquêteur dispose d'un paquet de 20 cartes, numérotées de 1 à 20, qu'il remet à la personne sondée. Celle-ci tire une carte au hasard et ne la montre pas à l'enquêteur.

La règle est alors la suivante :

- si la carte porte le numéro 1, la personne sondée répond « vrai » si elle est d'accord avec l'affirmation « A » et « faux » sinon,
- si la carte porte un autre numéro, la personne sondée répond « vrai » si elle n'est pas d'accord avec l'affirmation « A » et « faux » sinon.

Le but de l'enquête est d'évaluer la proportion p de gens de cette population qui sont réellement d'accord avec l'affirmation « A ».

1. On interroge une personne selon la règle précédente et on considère l'événement suivant, noté V : « la personne répond « vrai » ». On note θ la probabilité de l'événement V .

Exprimer θ en fonction de p , puis en déduire p en fonction de θ .

2. Certaines considérations théoriques laissent penser que $p = 17/18$.

a) Calculer θ .

b) Calculer la probabilité pour qu'une personne ayant répondu « vrai » soit d'accord avec l'affirmation « A ».

On revient au cas général où on ne connaît ni p , ni θ .

3. On considère un échantillon aléatoire, de taille n , extrait de la population considérée et on note S_n le nombre de réponses « vrai » obtenues. On suppose n assez grand pour pouvoir considérer que cet échantillonnage est assimilable à un tirage avec remise.

a) Donner la loi de S_n ainsi que son espérance et sa variance.

b) Montrer que $\frac{S_n}{n}$ est un estimateur sans biais et convergent de θ .

4. Dans cette question, on suppose que l'on a réalisé un échantillon de 100 personnes et on constate que 23 personnes ont répondu « vrai ».

a) Donner une estimation ponctuelle de θ et de p .

b) Donner un intervalle de confiance au niveau de confiance 0.95 de θ puis de p .

On rappelle que si Φ désigne la fonction de répartition d'une variable aléatoire X suivant la loi normale centrée réduite, alors $\Phi(1.96) = 0.975$.

Solution :

1. Notons T_1 l'événement « la personne tire la carte 1 », $(T_1, \overline{T_1})$ est un système complet, et :

$$\theta = P(V) = P(T_1)P_{T_1}(V) + P(\overline{T_1})P_{\overline{T_1}}(V) = \frac{1}{20}p + \frac{19}{20}(1-p)$$

$$\theta = \frac{19-18p}{20}, p = \frac{19-20\theta}{18}$$

2. a) $p = \frac{17}{18}$ donne $\theta = \frac{1}{10}$.

b) Notons D l'événement « la personne est d'accord avec l'affirmation A », on a :

$$P_V(D) = \frac{P(D)P_D(V)}{P(V)} = \frac{\frac{1}{20}p}{\theta} = \frac{17}{36}$$

3. a) S_n suit la loi binomiale $\mathcal{B}(n, \theta)$ d'espérance $n\theta$ et de variance $n\theta(1-\theta)$.

b) Ainsi $E(S_n/n) = \theta$ et $V(S_n/n) = \frac{1}{n}\theta(1-\theta) \xrightarrow[n \rightarrow \infty]{} 0$, donc :

(S_n/n) est un estimateur sans biais convergent de θ

4. a) Puisque 23 personnes ont dit « vrai », une estimation ponctuelle de θ est $\hat{\theta} = 0,23$ et une estimation ponctuelle de p est alors $\hat{p} = \frac{19-20\hat{\theta}}{18} = 0,8$.

b) On a $t_{0,95} \simeq 1,96$ et $[0,23 - \frac{t_{0,95}}{2\sqrt{100}}; 0,23 + \frac{t_{0,95}}{2\sqrt{100}}]$ est un intervalle de confiance symétrique de θ au niveau de confiance de 0,95. Ce qui donne ici pour intervalle de confiance de θ :

$$[0,132 ; 0,328]$$

Avec $p = \frac{19-20\hat{\theta}}{18}$, un intervalle de confiance de p , au même niveau de confiance, est donc :

$$[0,6911 ; 0,9089]$$

Exercice 3.23.

Soit g la fonction définie sur \mathbb{R}^* par

$$\forall t \neq 0, g(t) = \frac{e^t - 1}{t}$$

1. a) Montrer que g est prolongeable par continuité en 0 et que ce prolongement, encore noté g , admet un développement limité à tout ordre au voisinage de 0.

b) Tracer la représentation graphique de g .

2. Soit X une variable aléatoire qui suit la loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$, et $Y = \frac{1}{X + 1}$.

a) Déterminer la loi de la variable aléatoire Y .

b) Démontrer que, pour tout r de \mathbb{N} , Y^r admet une espérance $E(Y^r)$.

c) Calculer l'espérance de Y . Exprimer la variance de Y en fonction de la somme

$$S(\lambda) = \sum_{n=0}^{+\infty} \frac{\lambda^{n+1}}{(n+1)^2 n!}.$$

3. a) Démontrer l'encadrement, valable pour $N \in \mathbb{N}^*$ et $t \in [0, \lambda]$:

$$0 \leq g(t) - \sum_{n=0}^{N-1} \frac{t^n}{(n+1)!} \leq \frac{e^\lambda \lambda^N}{(N+1)!}$$

b) En déduire l'expression de la somme de la série de terme général $\frac{\lambda^{n+1}}{(n+1)^2 n!}$, $n \in \mathbb{N}$, à l'aide d'une intégrale.

c) Expliciter la variance $V(Y)$ à l'aide de la fonction g .

Solution :

1. a) \star On sait que $e^t - 1 \underset{(0)}{\sim} t$, donc $\lim_{t \rightarrow 0} g(t) = 1$ et g est prolongeable par continuité en 0, en posant $g(0) = 1$.

\star On sait que $e^t = \sum_{k=0}^{n+1} \frac{t^k}{k!} + o(t^{n+1})$, donc $g(t) = \frac{e^t - 1}{t} = \sum_{k=1}^{n+1} \frac{t^{k-1}}{k!} + o(t^n)$ et g admet bien un développement limité à l'ordre n , au voisinage de 0.

b) Pas de problème, la représentation graphique de g a même allure que la représentation graphique de la fonction \exp .

2. a) On a $Y(\Omega) = \{\frac{1}{n+1}, n \in \mathbb{N}\}$ et $\forall n \in \mathbb{N}, P(Y = \frac{1}{n+1}) = \frac{\lambda^n}{n!} e^{-\lambda}$.

b) Y est une variable aléatoire bornée, elle admet donc des moments de tous ordres, ce que l'on peut vérifier aisément sur la série définissant $E(Y^r)$.

$$c) \star E(Y) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{(n+1)!} = \frac{e^{-\lambda}}{\lambda} \sum_{n=0}^{\infty} \frac{\lambda^{n+1}}{(n+1)!} = \frac{e^{-\lambda}}{\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{n!}$$

$$E(Y) = \frac{e^{-\lambda}}{\lambda} (e^\lambda - 1)$$

$$\star E(Y^2) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{(n+1)^2 n!} = \frac{e^{-\lambda}}{\lambda} S(\lambda), \text{ d'où :}$$

$$V(Y) = \frac{e^{-\lambda}}{\lambda} S(\lambda) - \left(\frac{1 - e^{-\lambda}}{\lambda} \right)^2$$

3. a) D'après la formule de Taylor-Lagrange et la connaissance du développement de la fonction exponentielle, on a :

$$g(t) - \sum_{n=0}^{N-1} \frac{t^n}{(n+1)!} = \frac{t^N}{(N+1)!} e^c, \text{ pour un certain } c \text{ de } [0, t]$$

Soit :

$$\forall N \in \mathbb{N}^*, \forall t \in [0, \lambda], 0 \leq g(t) - \sum_{n=0}^{N-1} \frac{t^n}{(n+1)!} \leq \frac{\lambda^N}{(N+1)!} e^\lambda$$

b) En intégrant sur le segment $[0, \lambda]$:

$$0 \leq \int_0^\lambda g(t) dt - \sum_{n=0}^{N-1} \frac{\lambda^{n+1}}{(n+1)^2 n!} \leq \frac{\lambda^{N+1}}{(N+1)!} e^\lambda$$

Comme $\lim_{N \rightarrow \infty} \frac{\lambda^{N+1}}{(N+1)!} e^\lambda = 0$, il vient, par encadrement : $S(\lambda) = \int_0^\lambda g(t) dt$

$$c) \text{ Et donc : } V(Y) = \frac{e^{-\lambda}}{\lambda} \int_0^\lambda g(t) dt - [g(-\lambda)]^2$$

Exercice 3.24.

Un joueur prend pour cible un mur muni d'un repère orthonormé (O, i, j) . On note X et Y les variables aléatoires désignant respectivement l'abscisse et l'ordonnée du point d'impact du tir sur le mur, et on suppose que X et Y sont indépendantes, de même loi normale centrée réduite. On note Φ la fonction de répartition de X .

1. Donner une densité de $|X|$. Montrer que la variable aléatoire $Z = |X| + |Y|$ admet comme densité la fonction h définie par :

$$h(t) = \begin{cases} \frac{2}{\pi} e^{-t^2/4} \int_{-t/2}^{t/2} e^{-v^2} dv & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases}$$

2. a) Déterminer la dérivée de la fonction φ définie sur \mathbb{R} par :

$$\varphi(t) = \int_{-t/2}^{t/2} e^{-v^2} dv$$

b) Étudier les variations de φ et tracer l'allure de sa représentation graphique dans un repère orthonormé du plan.

3. On peint sur le mur le carré plein de sommets $(1, 0)$, $(0, 1)$, $(-1, 0)$ et $(0, -1)$. On note p la probabilité pour que l'impact soit dans le carré. Déterminer p en fonction de Φ .

Solution :

1. La variable aléatoire $|X|$ prend ses valeurs dans \mathbb{R}^+ et pour tout $x \geq 0$:

$$P(|X| \leq x) = P(-x \leq X \leq x) = \Phi(x) - \Phi(-x) = 2\Phi(x) - 1$$

Par dérivation légitime, une densité f de $|X|$ est :

$$f(x) = \begin{cases} \frac{2}{\sqrt{2\pi}} e^{-x^2/2} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

X et Y étant indépendantes, il en est de même de $|X|$ et $|Y|$ et une densité h de Z s'obtient par convolution (Z est une variable aléatoire clairement à valeurs dans \mathbb{R}^+ , on prend donc h nulle sur \mathbb{R}_-^*) :

$$\forall t \geq 0, h(t) = \int_{-\infty}^{+\infty} f(u)f(t-u)du = \int_0^t \frac{4}{2\pi} e^{-u^2/2} e^{-(t-u)^2/2} du$$

Le changement de variable $v = u - \frac{t}{2}$ donne alors :

$$\forall t \geq 0, h(t) = \frac{2e^{-t^2/4}}{\pi} \int_{-t/2}^{t/2} e^{-v^2} dv$$

2. a) $\varphi'(t) = \frac{1}{2} e^{-(t/2)^2} - (-\frac{1}{2}) e^{-(-t/2)^2} = e^{-t^2/4}$.

b) Ainsi la fonction φ est strictement croissante sur \mathbb{R} , clairement impaire et de limite en $+\infty$ valant $2\sqrt{\pi}$. Le tracé s'en déduit sans peine.

3. Le carré en question est vu comme un losange et est l'ensemble des points du plan dont les coordonnées vérifient $|x| + |y| \leq 1$, ce qui correspond donc à l'événement $Z \leq 1$.

$$\begin{aligned} P(Z \leq 1) &= \int_0^1 h(t) dt = \int_0^1 \left(\frac{2}{\pi} e^{-t^2/4} \int_{-t/2}^{t/2} e^{-v^2} dv \right) dt = \frac{1}{\pi} \int_0^1 2\varphi'(t)\varphi(t) dt \\ &= \frac{1}{\pi} [\varphi^2(t)]_0^1 = \frac{1}{\pi} \int_{-1/2}^{1/2} e^{-v^2} dv \end{aligned}$$

Enfin, le changement de variable $v = \frac{w}{\sqrt{2}}$ donne :

$$P(Z \leq 1) = \frac{1}{\pi\sqrt{2}} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} e^{-w^2/2} dw = \frac{1}{\sqrt{\pi}} \left(2\Phi\left(\frac{1}{\sqrt{2}}\right) - 1 \right)$$

Exercice 3.25.

Soit X une variable aléatoire à densité à valeurs dans \mathbb{R} et f une densité de X . On note Y la variable aléatoire égale à la partie entière de X .

1. a) Pour $k \in \mathbb{Z}$, expliciter $P(Y = k)$ au moyen d'une intégrale.

b) On suppose que f est nulle sur \mathbb{R}^- . Montrer que Y admet une espérance si et seulement si X en admet une et que, dans ce cas :

$$E(Y) \leq E(X) \leq E(Y) + 1$$

c) Expliciter la loi de Y et son espérance dans le cas où X suit la loi exponentielle de paramètre λ .

2. Dans cette question X suit la loi normale centrée réduite.

a) Expliciter la loi de Y à l'aide de la fonction de répartition Φ de X .

b) Pour $k \in \mathbb{N}^*$, comparer $P(Y = -k)$ et $P(Y = k - 1)$.

c) Montrer que Y admet une espérance et calculer celle-ci.

Solution :

1. a) On a $(Y = k) = (k \leq X < k + 1)$ et comme une variable aléatoire à densité ne charge pas les points :

$$P(Y = k) = \int_k^{k+1} f(t) dt$$

b) Si f est nulle sur \mathbb{R}^- , on a $Y(\Omega) \subset \mathbb{N}$ et pour tout k de \mathbb{N} :

$$\rightarrow kP(Y = k) = k \int_k^{k+1} f(t) dt \leq \int_k^{k+1} tf(t) dt$$

donc l'existence de $E(X)$, *i.e.* de $\int_0^{+\infty} tf(t) dt$, prouve la convergence de la série de terme général $kP(Y = k)$, donc de l'espérance de Y et, par prolongement des inégalités à la limite :

$$E(Y) = \sum_{k=0}^{\infty} kP(Y = k) \leq \int_0^{+\infty} tf(t) dt = E(X)$$

$$\rightarrow \text{De même } (k+1)P(Y = k) = (k+1) \int_k^{k+1} f(t) dt \geq \int_k^{k+1} tf(t) dt.$$

Si $E(Y)$ existe, alors $E(Y + 1)$ aussi et la suite $n \mapsto \int_0^{n+1} tf(t) dt$, qui est croissante, est majorée, donc a une limite ℓ lorsque n tend vers l'infini.

Ainsi la fonction $x \mapsto \int_0^x tf(t) dt$ est une fonction croissante majorée (l'intégrale sur $[0, x]$ est majorée par l'intégrale sur $[0, \lfloor x \rfloor + 1]$, elle-même majorée par ℓ), donc a une limite en $+\infty$ et $E(X)$ existe et $E(Y + 1) \geq E(X)$.

$$E(Y) \leq E(X) \leq E(Y) + 1$$

c) Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors $P(Y = k) = \int_k^{k+1} \lambda e^{-\lambda t} dt = e^{-\lambda k} - e^{-\lambda(k+1)}$, et :

$$E(Y) = (1 - e^{-\lambda}) \sum_{k=0}^{\infty} k(e^{-\lambda})^k = (1 - e^{-\lambda}) \times \frac{e^{-\lambda}}{(1 - e^{-\lambda})^2} = \frac{e^{-\lambda}}{1 - e^{-\lambda}}$$

2. a) $P(Y = k) = \Phi(k + 1) - \Phi(k)$.

b) Pour tout k , on a : $P(Y = -k) = \Phi(-k + 1) - \Phi(-k)$. Or $\Phi(-x) = 1 - \Phi(x)$, donc :

$$P(Y = -k) = (1 - \Phi(k - 1)) - (1 - \Phi(k)) = \Phi(k) - \Phi(k - 1)$$

c'est-à-dire :

$$P(Y = -k) = P(Y = k - 1)$$

c) Pour $k > 0$, on a $kP(Y = k) = \frac{k}{\sqrt{2\pi}} \int_k^{k+1} e^{-t^2/2} dt \leq \frac{1}{\sqrt{2\pi}} k \cdot e^{-k^2/2}$

Donc $k^3 P(Y = k) \xrightarrow[k \rightarrow \infty]{} 0$ et $\sum_{k=1}^{\infty} kP(Y = k)$ converge.

De la même façon (cf. b)), $\sum_{k=1}^{\infty} -kP(Y = -k)$ converge et finalement Y admet une espérance.

On peut écrire maintenant, en regroupant :

$$E(Y) = \sum_{k=-\infty}^{+\infty} kP(Y = k) = \sum_{k=1}^{\infty} k(P(Y = k) - P(Y = -k))$$

Soit :

$$\begin{aligned} E(Y) &= \sum_{k=1}^{\infty} [kP(Y = k) - kP(Y = k - 1)] \\ &= \sum_{k=1}^{\infty} kP(Y = k) - \sum_{k=1}^{\infty} (k - 1)P(Y = k - 1) - \sum_{k=1}^{\infty} P(Y = k - 1) \\ &= -P(Y \geq 0) = -P(X \geq 0) \\ &E(Y) = -\frac{1}{2} \end{aligned}$$

Exercice 3.26.

Soient X et Y deux variables aléatoires définies sur le même espace probabilisé (Ω, \mathcal{A}, P) , indépendantes et possédant comme densité la fonction f définie sur \mathbb{R} par :

$$f(x) = \frac{1}{\pi(1+x^2)}$$

1. Montrer que la variable aléatoire X_1 définie par :

$$X_1 = \begin{cases} \ln |X| & \text{si } X \neq 0 \\ 0 & \text{sinon} \end{cases}$$

est une variable aléatoire à densité et déterminer une densité de X_1 .

2. Prouver que la variable aléatoire $Z = \ln |XY|$ admet comme densité la fonction g définie par :

$$g(x) = \begin{cases} \frac{4xe^x}{\pi^2(e^{2x}-1)} & \text{si } x \in \mathbb{R}^* \\ \frac{2}{\pi^2} & \text{si } x = 0 \end{cases}$$

3. Montrer que $T = |XY|$ est une variable à densité et donner une densité de T .

4. En déduire la convergence et la valeur des intégrales :

$$I = \int_0^{+\infty} \frac{\ln x}{x^2-1} dx, \quad \text{et} \quad J = \int_0^1 \frac{\ln x}{x^2-1} dx$$

Solution :

1. X_1 prend ses valeurs dans \mathbb{R} et pour tout x de \mathbb{R} :

$$\begin{aligned} F_{X_1}(x) &= P(X_1 \leq x) = P(\ln |X| \leq x) = P(|X| \leq e^x) = P(-e^x \leq X \leq e^x) \\ &= \int_{-e^x}^{e^x} \frac{dt}{\pi(1+t^2)} = \int_0^{e^x} \frac{2}{\pi(1+t^2)} dt \end{aligned}$$

La fonction F_1 est de classe \mathcal{C}^1 sur \mathbb{R} , ce qui prouve que X_1 est une variable à densité et, par dérivation d'une composée, on peut prendre pour densité la fonction f_{X_1} définie par :

$$\forall x \in \mathbb{R}, f(x) = f_{X_1}(x) = \frac{2}{\pi(1+e^{2x})} e^x$$

2. On a $Z = \ln |X| + \ln |Y|$ et les variables aléatoires $X_1 = \ln |X|$ et $Y_1 = \ln |Y|$ sont indépendantes et de même loi. Ainsi Z est une variable à densité et une densité g de Z s'obtient par (auto)-convolution :

$$\begin{aligned} \forall x \in \mathbb{R}, g(x) &= \int_{-\infty}^{+\infty} f(u)f(x-u) du = \frac{4e^x}{\pi^2} \int_{-\infty}^{+\infty} \frac{du}{(1+e^{2u})(1+e^{2(x-u)})} \\ &= \frac{4e^x}{\pi^2} \int_{-\infty}^{+\infty} \frac{e^{2u} du}{(1+e^{2u})(e^{2x}+e^{2u})} \end{aligned}$$

Par le changement de variable légitime $v = e^{2u}$ il vient :

$$\forall x \in \mathbb{R}, g(x) = \frac{2e^x}{\pi^2} \int_0^{+\infty} \frac{dv}{(1+v)(e^{2x}+v)}$$

Pour $x \neq 0$, la décomposition de la fraction rationnelle donne :

$$\begin{aligned} g(x) &= \frac{2e^x}{\pi^2(e^{2x}-1)} \int_0^{+\infty} \left(\frac{1}{1+v} - \frac{1}{e^{2x}+v} \right) dv \\ g(x) &= \frac{2e^x}{\pi^2(e^{2x}-1)} \left[\ln \left(\frac{1+v}{e^{2x}+v} \right) \right]_0^{+\infty} = \frac{2e^x}{\pi^2(e^{2x}-1)} \ln(e^{2x}) \end{aligned}$$

et on peut donner à g n'importe quelle valeur en 0, par exemple $\frac{2}{\pi^2}$ (qui rend g continue sur \mathbb{R}), soit :

$$g(x) = \begin{cases} \frac{4xe^x}{\pi^2(e^{2x}-1)} & \text{si } x \in \mathbb{R}^* \\ \frac{2}{\pi^2} & \text{si } x = 0 \end{cases}$$

3. $T = e^Z$ prend ses valeurs dans \mathbb{R}_+^* et pour $t > 0$, on a :

$$F_T(t) = P(T \leq t) = P(Z \leq \ln t) = \int_{-\infty}^{\ln t} \frac{4ue^u}{\pi^2(e^{2u}-1)} du$$

et le changement de variable autorisé $u = \ln y$ donne le résultat :

$$\forall t > 0, F_T(t) = \int_0^t \frac{4 \ln y}{\pi^2(y^2-1)} dy$$

(la fonction à intégrer se prolonge par continuité en 1, il n'y a donc pas de problème)

Par dérivation, une densité φ de T est :

$$\varphi(t) = \begin{cases} \frac{4}{\pi^2} \times \frac{\ln t}{t^2-1} & \text{si } t > 0 \text{ et } t \neq 1 \\ \frac{2}{\pi^2} & \text{si } t = 1 \text{ (pour la continuité...)} \\ 0 & \text{sinon} \end{cases}$$

4. L'intégrale de φ sur \mathbb{R} valant 1, on a :

$$I = \int_0^{+\infty} \frac{\ln x}{x^2-1} dx = \frac{\pi^2}{4}$$

Le changement de variable $t = \frac{1}{u}$ montre que $\int_0^1 \frac{\ln x}{x^2 - 1} dx = \int_1^{+\infty} \frac{\ln x}{x^2 - 1} dx$
et ainsi :

$$J = \int_0^1 \frac{\ln x}{x^2 - 1} dx = \frac{\pi^2}{8}$$