3

PROBABILITÉS

Exercice 3.1.

Soit n un entier $\geqslant 1$. Une urne contient n boules blanches et une boule rose. On effectue une succession de tirages d'une boule à chaque fois. Si on tire la boule rose, on la remet dans l'urne avant le tirage suivant ; si on tire une boule blanche, on ne la remet pas dans l'urne. Soit X le nombre de tirages juste nécessaires pour qu'il n'y ait plus aucune boule blanche dans l'urne.
L'expérience est modélisée par un espace probabilisé $\big(\Omega, \mathcal{A}, P\big)$.

1. Calculer $P(X = n)$.

2. a) Pour tout entier $n \geqslant 2$, montrer que : $\ln(n+1) - \ln 2 \leqslant \sum\limits_{k=2}^{n} \frac{1}{k} \leqslant \ln n$.

En déduire un équivalent simple de $\sum\limits_{j=2}^{n} \frac{1}{j}$ quand n tend vers l'infini.

b) Trouver de même un équivalent simple de $\sum\limits_{j=2}^{n} \frac{\ln j}{j}$ quand n tend vers l'infini.

c) En déduire un équivalent simple de $\sum\limits_{j=2}^{n} \big(\frac{1}{j} \sum\limits_{i=2}^{j-1} \frac{1}{i}\big)$ quand n tend vers l'infini.

3. Calculer $P(X = n+1)$, puis en donner un équivalent simple quand n tend vers l'infini.

4. Calculer $P(X = n+2)$, puis en donner un équivalent simple quand n tend vers l'infini.

Solution :

1. Soit B_k $(k \geqslant 1)$ l'événement « la $k^{\text{ème}}$ boule tirée est blanche ».
Alors $[X = n] = B_1 \cap B_2 \cap \cdots \cap B_n$. Donc, par la formule des probabilités composées on a :

$$P(X = n) = P(B_1 \cap B_2 \cap \cdots \cap B_n) = P(B_1)P_{B_1}(B_2)P_{B_1 \cap B_2}(B_3) \ldots$$
$$= \frac{n}{n+1} \times \frac{n-1}{n} \times \cdots \times \frac{1}{2} = \frac{1}{n+1}$$

2. a) Comme $t \mapsto \frac{1}{t}$ décroît sur $\mathbb{R}_+^*$, on a : $\frac{1}{k+1} \leqslant \int_k^{k+1} \frac{dt}{t}$ et $\int_k^{k+1} \frac{dt}{t} \leqslant \frac{1}{k}$. En sommant la première inégalité de $k = 1$ à $k = n-1$, on obtient $\sum\limits_{k=2}^{n} \frac{1}{k} \leqslant \ln n$; en sommant la seconde inégalité de $k = 2$ à $k = n$ on obtient, $\ln(n+1) - \ln 2 \leqslant \sum\limits_{k=2}^{n} \frac{1}{k}$. Comme $\ln n \sim \ln(n+1) - \ln 2$ on a donc :

$$\sum_{j=2}^{n} \frac{1}{j} \sim \ln n$$

b) La fonction $x \mapsto \dfrac{\ln x}{x}$ décroît sur $[e, +\infty[$ et donc sur $[3, +\infty[$. On obtient donc de même :

$$\int_4^{n+1} \frac{\ln t}{t} dt \leqslant \sum_{k=4}^{n} \frac{\ln k}{k} \leqslant \int_3^{n} \frac{\ln t}{t} dt = \frac{1}{2}(\ln n)^2 - \frac{1}{2}(\ln 3)^2$$

D'où en calculant aussi l'intégrale de gauche et comme $\ln(n+1) \sim \ln n$:

$$\sum_{j=2}^{n} \frac{\ln j}{j} \sim \frac{1}{2}(\ln n)^2$$

c) Par sommation d'inégalités du 2. a) ($k = i$ et $n = j - 1$), on a :

$$\sum_{j=2}^{n} \frac{1}{j}(\ln j - \ln 2) \leqslant \sum_{j=2}^{n} \left(\frac{1}{j} \sum_{i=2}^{j-1} \frac{1}{i}\right) \leqslant \sum_{j=2}^{n} \frac{1}{j} \ln(j-1) \leqslant \sum_{j=2}^{n} \frac{1}{j} \ln j$$

D'après les questions 2. a et 2. b, comme $\ln n = o\big(\frac{1}{2}(\ln n)^2\big)$, on a :

$$\sum_{j=2}^{n} \frac{1}{j}(\ln j - \ln 2) = \big(\sum_{j=2}^{n} \frac{1}{j} \ln j\big) - \ln 2\big(\sum_{j=2}^{n} \frac{1}{j}\big) \sim \frac{1}{2}(\ln n)^2$$

d'où :

$$\sum_{j=3}^{n} \left(\frac{1}{j} \sum_{i=2}^{j-1} \frac{1}{i}\right) \sim \frac{1}{2}(\ln n)^2$$

3. L'événement $[X = n+1]$ est réalisé lorsque la boule rose est tirée une seule fois lors des $n+1$ premiers tirages, mais pas au $(n+1)$-ième rang (car sinon on réalise $X = n$), soit :

$$(X = n+1) = \bigcup_{k=1}^{n} [B_1 \cap \cdots \cap B_{k-1} \cap \overline{B_k} \cap B_{k+1} \cap \cdots \cap B_{n+1}]$$

Les évènements de la réunion sont incompatibles, donc :

$$P(X = n+1) = \sum_{k=1}^{n} P(B_1 \cap \cdots \cap B_{k-1} \cap \overline{B_k} \cap B_{k+1} \cap \cdots \cap B_{n+1})$$

Par la formule des probabilités composées, pour tout $k \in \{1, \ldots, n\}$, on a :
$P(B_1 \cap \cdots \cap B_{k-1} \cap \overline{B_k} \cap B_{k+1} \cap \cdots \cap B_{n+1})$

$$= P(B_1) P_{B_1}(B_2) \ldots P_{B_1 \cap \ldots \cap B_{k-1}}(\overline{B}_k) P_{B_1 \cap \ldots \cap B_{k-1} \cap \overline{B_k}}(B_{k+1}) \ldots P_{B_1 \cap \ldots \overline{B_k} \cap \ldots B_n}(B_{n+1})$$

$$= \frac{n}{n+1} \times \frac{n-1}{n} \times \frac{n-k+2}{n-k+3} \times \frac{1}{n-k+2} \times \frac{n-k+1}{n-k+2} \times \cdots \times \frac{1}{2}$$

Donc, par simplifications en cascade :

$$P(X = n+1) = \sum_{k=1}^{n} \frac{1}{(n+1)(n-k+2)} = \frac{1}{n+1} \sum_{j=2}^{n+1} \frac{1}{j}.$$

Donc, d'après la question 2. a) :

$$P(X = n+1) \sim \frac{\ln(n+1)}{n+1} \sim \frac{\ln n}{n}.$$

4. L'événement $[X = n+2]$ est réalisé lorsque la boule rose est tirée deux fois lors des $n+2$ premiers tirages, mais pas au $(n+2)^{\text{ème}}$ rang (car sinon on réalise $X < n+2$). De même qu'à la question 3, on a donc :

$$\begin{aligned}
P(X = n+2) &= \sum_{1 \leq k < l \leq n+1} P(B_1 \cap \cdots \cap \overline{B_k} \cap \cdots \cap \overline{B_l} \cap \cdots \cap B_{n+2}) \\
&= \sum_{1 \leq k < l \leq n+1} \frac{1}{(n+1)(n+2-k)(n+3-l)} \\
&= \frac{1}{n+1} \sum_{2 \leq i \leq j \leq n+1} \frac{1}{ij} = \sum_{k=2}^{n+1} \frac{1}{k^2} + \sum_{j=2}^{n} \big(\frac{1}{j} \sum_{i=2}^{j-1} \frac{1}{i}\big)
\end{aligned}$$

Donc, puisque le premier terme est de limite finie, donc négligeable devant le second :

$$P(X = n+2) \sim \frac{(\ln n)^2}{2n}$$

Exercice 3.2.

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ admettant une variance et dont l'espérance $E(X) = \lambda$ est un paramètre réel inconnu.

Pour n entier $\geqslant 1$, soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon *i.i.d.* de la loi de X. On pose $\overline{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$.

On note S_λ l'ensemble des statistiques $U_n = g_n(X_1, X_2, \ldots, X_n)$, où g_n est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, qui sont des estimateurs sans biais de λ et qui admettent une variance notée V.

On admet la propriété $\mathcal{P}$ suivante :

« Pour tout U_n de S_λ, on a : $\operatorname{cov}(\overline{X}_n, U_n - \overline{X}_n) = 0$. »

On dit qu'un élément Z_n de S_λ est un *estimateur optimal* dans S_λ si pour tout U_n de S_λ, on a : $V(Z_n) \leqslant V(U_n)$.

1. Montrer que $\overline{X}_n$ est un estimateur optimal dans S_λ.

2. Soit Z_n un estimateur optimal dans S_λ. Pour α réel et U_n de S_λ, on pose :

$$W_n(\alpha) = \alpha U_n + (1-\alpha) Z_n$$

a) Montrer que pour tout α réel, $W_n(\alpha)$ est élément de S_λ.

b) Calculer $V(W_n(\alpha))$. En déduire que $\operatorname{Cov}(Z_n, U_n - Z_n) = 0$.

c) Montrer que $Z_n = \overline{X}_n$ presque sûrement.

3. On suppose que X suit la loi de Poisson de paramètre $\lambda > 0$, et on admet que la propriété $(\mathcal{P})$ est vérifiée.

Pour tout $n \geqslant 2$, on pose : $T_n = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \overline{X}_n)^2$

a) Montrer que T_n est un estimateur sans biais de λ.

b) On admet sans démonstration l'existence de $V(T_n)$.

Montrer que $\operatorname{Cov}(T_n, \overline{X}_n) = \frac{\lambda}{n}$.

Solution :

1. $\overline{X}_n$ s'écrit comme une fonction de $X_1, X_2, \ldots, X_n$: c'est un estimateur. De plus, $\overline{X}_n$ admet une espérance égale à λ et une variance, ce qui entraîne que $\overline{X}_n$ appartient à S_λ.

Par la propriété $(\mathcal{P})$, on a : $\operatorname{cov}(\overline{X}_n, U_n) = V(\overline{X}_n)$.
Par l'inégalité de Cauchy-Schwarz, : $|\operatorname{cov}(\overline{X}_n, U_n)|^2 \leqslant V(\overline{X}_n)V(U_n)$, donc $V(\overline{X}_n) \leqslant V(U_n)$, ce qui montre que $\overline{X}_n$ est optimal.

2. a) Question évidente, par linéarité de l'opérateur espérance et le fait que si U et V ont un moment d'orde deux, alors (U, V) a une covariance.

b) Bien évidemment :

$$\begin{aligned} V(W_n(\alpha)) &= \alpha^2 V(U_n) + (1-\alpha)^2 V(Z_n) + 2\alpha(1-\alpha)\operatorname{cov}(U_n, Z_n) \\ &= \alpha^2 V(U_n - Z_n) - 2\alpha(V(Z_n) - \operatorname{cov}(U_n, Z_n)) + V(Z_n) \end{aligned}$$

Or Z_n étant optimal, il vient, pour tout α réel :

$$\alpha^2 V(U_n - Z_n) - 2\alpha(V(Z_n) - \operatorname{cov}(U_n, Z_n)) \geqslant 0$$

Ce trinôme du second degré en α restant positif ou nul, son discriminant est négatif, soit :
$(V(Z_n) - \operatorname{Cov}(U_n, Z_n))^2 \leqslant 0$, donc $V(Z_n) = \operatorname{Cov}(U_n, Z_n)$, ou

$$\operatorname{Cov}(Z_n, U_n - Z_n) = 0.$$

c) On écrit : $Z_n = \overline{X}_n + (Z_n - \overline{X}_n)$ ce qui entraîne que

$$V(Z_n) = V(\overline{X}_n) + V(Z_n - \overline{X}_n) + 2\operatorname{Cov}(\overline{X}_n, Z_n - \overline{X}_n).$$

Comme Z_n et $\overline{X}_n$ sont optimaux, il vient $V(Z_n - \overline{X}_n) + 2\operatorname{Cov}(\overline{X}_n, Z_n - \overline{X}_n) = 0$ et toujours par optimalité de $\overline{X}_n$, $\operatorname{Cov}(\overline{X}_n, Z_n - \overline{X}_n) = 0$.

Finalement, $V(Z_n - \overline{X}_n) = 0$ et $Z_n - \overline{X}_n = C$ p.s., donc $Z_n - \overline{X}_n = 0$ p.s.

3. a) En écrivant $X_i - \overline{X}_n = (X_i - \lambda + \lambda - \overline{X}_n)$, il vient :

$$\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \sum_{i=1}^{n}(X_i - \lambda)^2 + n(\overline{X}_n - \lambda)^2 - 2n(\overline{X}_n - \lambda)^2 = \sum_{i=1}^{n}(X_i - \lambda)^2 - n(\overline{X}_n - \lambda)^2$$

En prenant les espérances :

$$E\Big(\sum_{i=1}^{n}(X_i - \overline{X}_n)^2\Big) = \sum_{i=1}^{n} V(X_i) - nV(\overline{X}_n) = (n-1)\lambda$$

Par suite $E(T_n) = \lambda$ et T_n est un estimateur sans biais de λ.

b) Comme T_n admet une variance, il appartient à S_λ et par la propriété $(\mathcal{P})$ et $T_n \neq \overline{X}_n$, il vient $\mathrm{Cov}(\overline{X}_n, T_n - \overline{X}_n) = 0$. Donc :

$$\mathrm{Cov}(T_n, \overline{X}_n) = \frac{\lambda}{n}$$

Exercice 3.3.

On observe le passage de véhicules à un carrefour. Pour tout réel $t \geqslant 0$, on note X_t le nombre de véhicules qui passent entre les dates 0 et t, avec la convention $X_0 = 0$. On suppose que l'on définit ainsi une famille de variables aléatoires $(X_t)_{t\in\mathbb{R}_+}$ définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$ possédant les propriétés suivantes :

• $E(X_t)$ existe pour tout $t \geqslant 0$ et la fonction $m(t) = E(X_t)$ est continûment dérivable sur $\mathbb{R}^+$, sa dérivée étant notée $\lambda(t)$.

• Le nombre de passages $Y_{t_0,h}$ se produisant entre les instants t_0 et $t_0 + h$, avec $t_0 \geqslant 0$ et $h > 0$, est une variable aléatoire dont la loi est définie par :

$$\forall\, k \in \mathbb{N}, P(Y_{t_0,h} = k) = \frac{\alpha^k \mathrm{e}^{-\alpha}}{k!} \text{ où } \alpha = \int_{t_0}^{t_0+h} \lambda(u)du = m(t_0 + h) - m(t_0)$$

(on notera donc que α dépend de t_0 et de h)

1. Soit T_1 la variable aléatoire représentant la date de passage du premier véhicule observé. Calculer $P(T_1 \geqslant t)$. En déduire une densité de T_1 en utilisant les fonctions λ et m.

2. a) On suppose que l'on connaît la date t_0 de passage du premier véhicule. Soit L_1 la longueur de l'intervalle de temps séparant les passages du premier et du deuxième véhicule. Calculer une densité de L_1 en fonction de t_0, λ et m.

b) On suppose dans cette question que pour $t \geqslant 0$, $\lambda(t) = at + b$, a et b étant des constantes positives. Déterminer une densité f_1 de L_1 dans ce cas.

3. On suppose désormais que pour tout $t \geqslant 0$: $\lambda(t) = \lambda_0$, constante positive. Pour $k \in \mathbb{N}^*$, soit T_k la date de passage du k-ième véhicule.

a) Calculer $P(T_k \geqslant t)$ à l'aide de la loi de $Y_{0,t}$. Démontrer qu'une densité f_k de T_k est :

$$f_k(t) = \begin{cases} \dfrac{\lambda_0(\lambda_0 t)^{k-1}\mathrm{e}^{-\lambda_0 t}}{(k-1)!} & \text{pour } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

b) Calculer l'espérance de T_k.

Solution :

1. On a $P(T_1 \geqslant t) = P(X_t = 0) = P(Y_{0,t}) = 0) = \mathrm{e}^{-\alpha} = \mathrm{e}^{-m(t)}$.
D'où $P(T_1 < t) = 1 - \mathrm{e}^{-m(t)}$ et on peut prendre : $f_{T_1}(t) = \lambda(t)\mathrm{e}^{-m(t)}$ (bien sûr pour $t \geqslant 0$).

2. a) On a : $P(L_1 > t) = P(Y_{t_0,t} = 0) = \mathrm{e}^{-(m(t_0+t)-m(t_0))}$.
D'où, $P(L_1 \leq t) = 1 - \mathrm{e}^{-(m(t_0+t)-m(t_0))}$ et,

$$f_{L_1}(t) = m'(t_0 + t)\mathrm{e}^{-(m(t_0+t)-m(t_0))} = \lambda(t_0 + t)\mathrm{e}^{-(m(t_0+t)-m(t_0))}$$

b) Dans ce cas $m(t) = a\dfrac{t^2}{2} + bt$, ce qui donne :

$$f_1(t) = (a(t_0+t)+b)\exp(-(a\frac{(t+t_0)^2}{2} + b(t+t_0) - a\frac{t_0^2}{2} - bt_0))$$

$$f_1(t) = \begin{cases} (a(t+t_0)+b)\exp(-(a\dfrac{t^2}{2} + (at_0+b)t)) & \text{si } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

3. a) Pour $t \geqslant 0$, on a, $P(T_k \geqslant t) = P(Y_{0,t} < k) = \sum\limits_{i=0}^{k-1} \dfrac{\alpha^i \mathrm{e}^{-\alpha}}{i!} = \sum\limits_{i=0}^{k-1} \dfrac{(\lambda_0 t)^i \mathrm{e}^{-\lambda_0 t}}{i!}$

On en déduit : $f_{t_k}(t) = -\lambda_0 \sum\limits_{i=1}^{k-1} \dfrac{(\lambda_0 t)^{i-1}\mathrm{e}^{-\lambda_0 t}}{(i-1)!} + \lambda_0 \sum\limits_{i=0}^{k-1} \dfrac{(\lambda_0 t)^i \mathrm{e}^{-\lambda_0 t}}{i!}$

$$f_{t_k}(t) = \begin{cases} \dfrac{\lambda_0(\lambda_0 t)^{k-1}\mathrm{e}^{-\lambda_0 t}}{(k-1)!} & \text{si } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

b) Enfin :

$$\begin{aligned} E(T_k) &= \int_0^{+\infty} t f_k(t)dt = \frac{1}{(k-1)!}\int_0^{+\infty} (\lambda_0 t)^k \mathrm{e}^{-\lambda_0 t} dt \\ &= \frac{1}{\lambda_0 (k-1)!}\int_0^{+\infty} u^k \mathrm{e}^{-u} du = \frac{k}{\lambda_0} \end{aligned}$$

Exercice 3.4.

Soit $n \in \mathbb{N}^*$. On considère n variables aléatoires mutuellement indépendantes $X_1, \ldots, X_n$, qui suivent une loi normale d'espérance m et de variance σ^2.

On pose $\overline{X}_n = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ et $U_n = \dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i - \overline{X}_n)^2$.

1. Soit $A = (a_{i,j})_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{R})$ définie par : $a_{i,j} = \begin{cases} n-1 & \text{si } i = j \\ -1 & \text{si } i \neq j \end{cases}$. Soit B le vecteur colonne de $\mathcal{M}_{n,1}(\mathbb{R})$ dont tous les coefficients sont égaux à 1.

a) Justifier que la matrice A est diagonalisable.

b) Déterminer les valeurs propres de A ainsi qu'une base des sous-espaces propres associés.

c) Justifier l'existence d'une matrice P de $GL_n(\mathbb{R})$ dont la dernière colonne est proportionnelle à B et telle que ${}^tPAP = D$, où D est une matrice diagonale à déterminer (on ne demande pas la matrice P).

d) On note ${}^tP = (p_{i,j})_{1\leq i,j\leq n}$.

Montrer que pour tout i de $[\![1, n-1]\!]$, $\sum\limits_{j=1}^{n} p_{i,j} = 0$ et pour tout i de $[\![1, n]\!]$, $\sum\limits_{j=1}^{n} p_{i,j}^2 = 1$.

2. On note $M = \dfrac{1}{n}A$ et q l'application de $\mathcal{M}_{n,1}(\mathbb{R})$ dans $\mathbb{R}$ définie par : $\forall\, X \in \mathcal{M}_{n,1}(\mathbb{R}), q(X) = {}^tXMX$.

On pose $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\overline{x}_n = \dfrac{1}{n}\sum\limits_{j=1}^{n} x_j$.

Montrer que $q(X) = \sum\limits_{i=1}^{n}(x_i - \overline{x}_n)^2$, puis $q(X) = \sum\limits_{i=1}^{n-1}(\sum\limits_{j=1}^{n} p_{i,j}x_j)^2$.

3. Pour tout i de $[\![1, n-1]\!]$, on pose $Y_i = \sum\limits_{j=1}^{n} p_{i,j}X_j$.

a) Justifier que $E(Y_i) = 0$ et $V(Y_i) = \sigma^2$.

b) On suppose que les Y_i sont mutuellement indépendantes. Montrer que U_n suit la loi $\Gamma\big(2, \dfrac{n-1}{2}\big)$.

Solution :

1. a) La matrice A est symétrique réelle : elle est diagonalisable dans une base orthonormée.

b) Si l'on note J la matrice dont tous les éléments valent 1, on a $A = nI - J$. Les valeurs propres de J sont 0 de sous-espace propre associé de dimension $n-1$ ($\operatorname{Ker} J$), et n dont le sous-espace propre associé est de dimension 1 engendré par le vecteur B (orthogonal à $\operatorname{Ker} J$).
Les valeurs propres de A sont donc n de sous-espace propre associé $\operatorname{Ker} J$ de dimension $n-1$, et 0 de sous-espace propre associé $\operatorname{Vect}(B)$.

c) La matrice D demandée est la matrice diagonale $\operatorname{diag}(n, \dots, n, 0)$.

d) La matrice P étant orthogonale, la relation ${}^tPP = I$ donne : $\sum\limits_{j=1}^{n} p_{i,j}^2 = 1$. Si l'on note $C_1, \dots, C_{n-1}$ les premières colonnes de la matrice P, elles représentent les coordonnées des vecteurs du noyau de J, ce qui donne la relation : $\sum\limits_{j=1}^{n} p_{i,j} = 0$.

2. L'écriture de l'application q donne :

$$q(X) = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\Big(\sum_{j=1}^{n} x_j\Big)^2 = \sum_{i=1}^{n}(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j)^2$$

La relation $q(X) = \sum\limits_{i=1}^{n-1}(\sum\limits_{j=1}^{n} p_{i,j}x_j)^2$ est obtenue à partir de l'écriture de q dans la base orthonormée de diagonalisation, c'est-à-dire $q(X) = {}^t(PX)D(PX)$.

3. a) On a $E(Y_i) = 0$ par linéarité de l'espérance et la relation $\sum\limits_{j=1}^{n} p_{i,j} = 0$.
Par indépendance des variables aléatoires (X_i), il vient :

$$V(Y_i) = \sum_{j=1}^{n} p_{i,j}^2 V(Y_j) = \sigma^2.$$

Enfin, on sait (encore l'indépendance des X_i) que les Y_i suivent encore des lois normales. Elles suivent donc la loi normale centrée de variance σ^2.

b) Grâce aux questions précédentes, on peut écrire :

$$U_n = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - M_n)^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n-1}\Big(\sum_{j=1}^{n} p_{i,j}X_j\Big)^2 = \sum_{i=1}^{n-1}\Big(\frac{Y_i}{\sigma}\Big)^2$$

Donc, U_n est la somme des carrés de $n-1$ variables aléatoires indépendantes de même loi normale centrée réduite. On sait alors que : $U_n \hookrightarrow \Gamma\Big(2, \dfrac{n-1}{2}\Big)$.

Exercice 3.5.

On considère une fonction F définie sur $[0, +\infty[$ à valeurs réelles, croissante, continue, telle que $F(0) = 0$ et $\lim\limits_{x\to+\infty} F(x) = 1$.

1. On définit la suite $(p_n)_{n\in\mathbb{N}}$ par, pour tout $n \in \mathbb{N}, p_n = F(n+1) - F(n)$.
Montrer que la série de terme général p_n est convergente et déterminer sa somme.

2. a) Soit x un réel fixé dans $[0, 1[$. Montrer que la série de terme général $F(x+n) - F(n)$ est convergente.

b) On considère la fonction φ définie sur $[0, 1[$ par : $\varphi(x) = \sum\limits_{n=0}^{+\infty}(F(x+n) - F(n))$.
Montrer que φ est croissante.

c) Soient x et y deux réels fixés de $[0, 1[$. Montrer que pour tout ε réel fixé strictement positif, il existe N_0 entier positif indépendant de x et y tel que, pour tout $N \geqslant N_0$:

$$\sum_{n=N}^{+\infty} |F(x+n) - F(y+n)| < \varepsilon$$

En déduire que φ est continue sur $[0, 1[$.

3. On considère désormais que F est la fonction de répartition d'une variable aléatoire X à valeurs dans $\mathbb{R}^+$ définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. On considère les variables aléatoires notées $d(X)$ et $\lfloor X \rfloor$ définies par, pour tout $\omega \in \Omega$:

$$d(X)(\omega) = X(\omega) - \lfloor X(\omega) \rfloor \text{ et } \lfloor X \rfloor(\omega) = \lfloor X(\omega) \rfloor$$

où $\lfloor x \rfloor$ désigne la partie entière du réel x.

a) Déterminer la loi de $\lfloor X \rfloor$ à l'aide des (p_n) et la fonction de répartition de $d(X)$.

b) On suppose que X suit une loi exponentielle de paramètre $\lambda > 0$. Expliciter la loi de $\lfloor X \rfloor$ et la fonction de répartition de $d(X)$. Déterminer une densité de $d(X)$, puis calculer les espérances de ces deux variables aléatoires.

Solution :

1. Pour tout $N \in \mathbb{N}^*$, on a $\sum_{n=0}^{N} p_n = F(N+1)$. D'après les hypothèses, on en conclut que la série converge et que sa somme vaut 1.

2. a) Pour tout $n \in \mathbb{N}^*$, on a : $0 \leqslant F(x+n) - F(n) \leqslant p_n$. Cela montre que la série considérée est convergente et que sa somme est inférieure à 1.

b) Pour tout $n \in \mathbb{N}^*$ et tout réels x et y tels que $x \geqslant y$, $F(x+n) - F(n) \geqslant F(y+n) - F(n)$ ce qui montre que $\varphi(x) \geqslant \varphi(y)$ et donc que la fonction φ est croissante.

c) De manière évidente, pour tout $n \in \mathbb{N}^*$, $|F(x+n) - F(y+n)| \leqslant p_n$. D'où,

$$\forall N \in \mathbb{N}, N \geqslant N_0, \sum_{n=N}^{+\infty} |F(x+n) - F(y+n)| \leqslant \sum_{n=N}^{+\infty} p_n$$

Pour tout $\varepsilon > 0$, la convergence de la deuxième série implique qu'il existe $N_0 \in \mathbb{N}^*$ tel que la somme considérée soit inférieure à ε, d'où le résultat.
On écrit alors :

$$|\varphi(x) - \varphi(y)| \leqslant \sum_{n=1}^{N_0} |F(x+n) - F(y+n)| + \sum_{n=N_0+1}^{+\infty} |F(x+n) - F(y+n)|$$

La seconde somme est inférieure à ε, indépendamment de x et y, et par continuité d'une somme finie de fonctions continues, il existe $\delta > 0$ tel que la première somme soit inférieure à ε, pour $|x-y| < \delta$. Ceci montre la continuité de φ sur $0, 1[$.

3. a) Pour tout $n \in \mathbb{N}$, $P(\lfloor X \rfloor = n) = p_n$, et :

$$P(d(X) < x) = \begin{cases} 0 & \text{si } x < 0 \\ \varphi(x) & \text{si } x \in [0,1[\\ 1 & \text{si } x \geqslant 1 \end{cases}$$

b) Sous les hypothèses de cette question, pour tout $n \in \mathbb{N}$:

$$P(\lfloor X \rfloor = n) = \mathrm{e}^{-\lambda n} - \mathrm{e}^{-\lambda(n+1)}$$

Donc :

$$\begin{aligned}\varphi(x) &= \sum_{n=0}^{+\infty} \left[(1 - \mathrm{e}^{-\lambda(x+n)}) - (1 - \mathrm{e}^{-\lambda n})\right] = \sum_{n=0}^{+\infty} (\mathrm{e}^{-\lambda n} - \mathrm{e}^{-\lambda(x+n)}) \\ &= \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda n}(1 - \mathrm{e}^{-\lambda x}) = (1 - \mathrm{e}^{-\lambda x}) \sum_{n=0}^{+\infty} (\mathrm{e}^{-\lambda})^n = \frac{1 - \mathrm{e}^{-\lambda x}}{1 - \mathrm{e}^{-\lambda}}\end{aligned}$$

Une densité de $d(X)$ est donnée par $f(x) = \dfrac{\lambda \mathrm{e}^{-\lambda x}}{1 - \mathrm{e}^{-\lambda}}$, si $x \in [0,1[$, et 0 sinon.

On a :

$$\begin{aligned}E(\lfloor X \rfloor) &= \sum_{n=0}^{+\infty} n(\mathrm{e}^{-\lambda n} - \mathrm{e}^{-\lambda(n+1)}) = \sum_{n=0}^{+\infty} n\mathrm{e}^{-\lambda n} - \sum_{n=0}^{+\infty} (n+1-1)\mathrm{e}^{-\lambda(n+1)}) \\ &= \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda(n+1)} = \frac{\mathrm{e}^{-\lambda}}{1 - \mathrm{e}^{-\lambda}}\end{aligned}$$

et

$$E(d(X)) = E(X) - E(\lfloor X \rfloor) = \frac{1}{\lambda} - \frac{\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}} = \frac{1-(\lambda+1)\mathrm{e}^{-\lambda}}{\lambda(1-\mathrm{e}^{-\lambda})}$$

Exercice 3.6.

Soit X une variable aléatoire réelle de densité g et de fonction de répartition G. On suppose que g est une fonction paire définie sur $\mathbb{R}$.

1. Exprimer, pour tout réel x, $G(-x)$ en fonction de $G(x)$.

2. On définit la fonction $\widetilde{g}$ par : $\widetilde{g}(t) = \begin{cases} 0 & \text{si } t \leqslant 0 \\ 2g(t) & \text{si } t > 0 \end{cases}$.

Vérifier que $\widetilde{g}$ est une densité de probabilité. On note $\widetilde{X}$ une variable aléatoire de densité $\widetilde{g}$.

3. Exprimer la fonction de répartition $\widetilde{G}$ de $\widetilde{X}$ en fonction de G.
Donner en particulier, pour $0 < a < b$, $\widetilde{G}(b) - \widetilde{G}(a)$ en fonction de $G(b)$ et de $G(a)$.

4. Pour tout m de $\left[\frac{1}{2}, 1\right]$ et tout réel t, on pose :

$$f_m(t) = 2(1-m)g(t) + (2m-1)\widetilde{g}(t).$$

a) Vérifier que f_m est une densité de probabilité et exprimer $f_m(t)$ en fonction de m et de $g(t)$.

b) Soit Y_m une variable aléatoire de densité f_m. Déterminer la fonction de répartition F_m de Y_m, en fonction de m et de G.

c) Déterminer la loi de probabilité de la variable aléatoire $\lfloor |Y_m| \rfloor$ (partie entière de la valeur absolue de Y_m) en fonction de G, puis de $\widetilde{G}$.

5. On considère une suite $(X_k)_{k\in\mathbb{N}^*}$ de variables aléatoires indépendantes suivant toutes la même loi que Y_m.
On pose, pour tout entier naturel n non nul, $Z_n = (X_1, \ldots, X_n)$ et $T_n = \sum\limits_{i=1}^{n} \mathbf{1}_{(X_i \geq 0)}$, c'est-à-dire que T_n est le nombre de variables du n-uplet Z_n qui sont positives. On pose enfin $S_n = \dfrac{1}{n}T_n$.

a) Donner la loi de T_n, son espérance, sa variance.

b) Montrer que S_n est un estimateur convergent de m.

6. Soit φ une fonction de $[\![0, n]\!]$ dans $\mathbb{R}$, telle que : $\forall\, m \in \left[\frac{1}{2}, 1\right[$, $E(\varphi(T_n)) = 0$. Montrer que $\varphi = 0$.

Solution :

1. On a $G(x) = \displaystyle\int_{-\infty}^{x} g(t)dt$ et donc, $G(-x) = \displaystyle\int_{-\infty}^{-x} g(t)dt$. Avec le changement de variable $u = -t$, on obtient $G(-x) = \displaystyle\int_{+\infty}^{x} g(u)(-du) = \int_{x}^{+\infty} g(u)du$. Conclusion :

$$G(-x) = 1 - G(x).$$

2. La fonction $\widetilde{g}$ est évidemment continue sauf éventuellement en un nombre fini de points et positive. De plus $\displaystyle\int_{-\infty}^{+\infty} \widetilde{g}(t)dt = 2\int_{0}^{+\infty} g(t)dt = 1.$

3. Si $x < 0$, $\widetilde{G}(x) = 0$.
Si $x \geqslant 0$, $\widetilde{G}(x) = \displaystyle\int_{-\infty}^{x} \widetilde{g}(t)dt = 2\int_{0}^{x} g(t)dt = \int_{-x}^{x} g(t)dt = G(x) - G(-x)$

$= 2G(x) - 1.$

On a donc, pour $0 < a < b$, $\widetilde{G}(b) - \widetilde{G}(a) = 2(G(b) - G(a))$.

4. a) La fonction f_m est évidemment continue sauf éventuellement en un nombre fini de points et positive.
De plus $\displaystyle\int_{-\infty}^{+\infty} f_m(t)dt = 2(1-m)\int_{-\infty}^{+\infty} g(t)dt + (2m-1)\int_{-\infty}^{+\infty} \widetilde{g}(t)dt$, soit :

$$\int_{-\infty}^{+\infty} f_m(t)dt = 2(1-m) + (2m-1) = 1.$$

Si $t \leqslant 0$, on a $f_m(t) = 2(1-m)g(t)$.
Si $t > 0$, on a $f_m(t) = 2(1-m)g(t) + 2(2m-1)g(t) = 2mg(t)$.

b) Si $x \leqslant 0$, $F_m(x) = \displaystyle\int_{-\infty}^{x} 2(1-m)g(t)dt = 2(1-m)G(x)$.

Si $x > 0, F_m(x) = \displaystyle\int_{-\infty}^{0} 2(1-m)g(t)dt + \int_0^x 2mg(t)dt = 1 - m + 2m(G(x) - \frac{1}{2})$.

c) Soit $k \in \mathbb{N}$. On a :

$$\begin{aligned} P(\lfloor |Y_m| \rfloor = k) &= P(k \leqslant |Y_m| < k+1) \\ &= P(k \leqslant Y_m < k+1) + P(-k-1 < Y_m \leqslant -k). \end{aligned}$$

Soit $P(\lfloor |Y_m| \rfloor = k) = F_m(k+1) - F_m(k) + F_m(-k) - F_m(-k-1)$. En remplaçant F_m à l'aide de G, on trouve :

$$P([|Y_m|] = k) = 2(G(k+1) - G(k)) = \widetilde{G}(k+1) - \widetilde{G}(k).$$

5. a) La variable aléatoire T_n suit une loi binomiale dont le paramètre est

$$p = P(X_i \geqslant 0) = \int_0^{+\infty} f_m(t)dt = m.$$

On en déduit que $E(T_n) = nm$ et $V(T_n) = nm(1-m)$.

b) On a donc $E(S_n) = m$ et $V(S_n) = \dfrac{m(1-m)}{n}$. Ceci prouve, via l'inégalité de Bienaymé-Tchebychef, que S_n converge en probabilité vers m et est donc un estimateur convergent de m.

6. On a : $E(\varphi(T_n)) = 0 \Longleftrightarrow \displaystyle\sum_{k=0}^{n} \varphi(k)\tbinom{n}{k} m^k (1-m)^{n-k} = 0$

$$\Longleftrightarrow \sum_{k=0}^{n} \varphi(k)\tbinom{n}{k}\Big(\frac{m}{1-m}\Big)^k = 0.$$

Considérons le polynôme P défini par $P = \displaystyle\sum_{k=0}^{n} \varphi(k)\tbinom{n}{k} X^k$. Ce polynôme s'annule pour tout les réels $\alpha_m = \dfrac{m}{1-m}$ obtenus lorsque m parcourt l'intervalle $[\frac{1}{2}, 1[$. Il possède donc une infinité de racines, c'est donc le poynôme nul.

Conclusion : $\forall\, k \in [\![0, n]\!], \varphi(k) = 0$ et φ est la fonction nulle.

Exercice 3.7.

1. Soit X et Y deux variables aléatoires réelles à densité définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et f_Y une densité de Y. Pour tout z réel, justifier la convergence de l'intégrale

$$\int_{-\infty}^{+\infty} P(X \leqslant z - y)\, f_Y(y)dy$$

On admet que, si X et Y sont indépendantes, alors :

$$\forall z \in \mathbb{R}, P(X + Y \leqslant z) = \int_{-\infty}^{+\infty} P(X \leqslant z - y)\, f_Y(y)dy$$

Soit $\alpha \in \mathbb{R}_+^*$ fixé et $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes telle que, pour tout $k \in \mathbb{N}^*$, X_k suit la loi exponentielle de paramètre $k\alpha$.

Pour n entier de $\mathbb{N}^*$, on note $S_n = \sum_{k=1}^{n} X_k$, Φ_n la fonction de répartition de S_n et $T_n = \dfrac{S_n}{n}$.

2. Montrer que : $(\forall\, n \in \mathbb{N}^*), (\forall z \in \mathbb{R}_+), \Phi_n(z) = \big(1 - \mathrm{e}^{-\alpha z}\big)^n$.

(*On pourra employer (en le justifiant) le changement de variable $u = \mathrm{e}^{-\alpha x}$.*)

3. a) Calculer l'espérance $E(T_n)$ de la variable aléatoire T_n. A-t-elle une limite pour n tendant vers $+\infty$? Déterminer un équivalent de $E(T_n)$ lorsque n tend vers $+\infty$.

b) Calculer la variance de T_n. A-t-elle une limite pour n tendant vers $+\infty$?

4.Étudier la convergence en loi de la suite $(T_n)_{n\in\mathbb{N}^*}$.

Solution :

1. Par comparaison : $\forall z \in \mathbb{R}, 0 \leqslant P(X \leqslant z-y)f_Y(y) \leqslant f_Y(y)$ qui est une fonction d'intégrale convergente sur $\mathbb{R}$. La convergence en résulte.

2. La variable $S_1 = X_1$ suit la loi $\mathcal{E}(\alpha)$ et par récurrence, pour $z \geqslant 0$ comme $S_n(\Omega) = \mathbb{R}^+$:
$P(S_{n+1} \leqslant z) = P(S_n + X_{n+1} \leqslant z) = \displaystyle\int_0^{+\infty} P\big(S_n \leqslant z-x\big)\,(n+1)\alpha \mathrm{e}^{-(n+1)\alpha x}dx$
Comme $P(S_n \leqslant z-x) = \begin{cases} 0 & \text{si } x \geqslant z \\ \big(1-\mathrm{e}^{-\alpha(z-x)}\big)^n & \text{si } x \leqslant z \end{cases}$, en posant $u = \mathrm{e}^{-\alpha x}$ il vient :

$$\begin{aligned} P(S_{n+1} \leqslant z) &= (n+1)\int_0^z \big(1-\mathrm{e}^{-\alpha(z-x)}\big)^n\,\alpha \mathrm{e}^{-(n+1)\alpha x}dx \\ &= (n+1)\int_{\mathrm{e}^{-\alpha z}}^1 \big(1-\frac{\mathrm{e}^{-\alpha z}}{u}\big)^n u^n du = (n+1)\int_{\mathrm{e}^{-\alpha z}}^1 \big(u-\mathrm{e}^{-\alpha z}\big)^n du \\ &= \big(1-\mathrm{e}^{-\alpha z}\big)^{n+1} \end{aligned}$$

3. a) Vu que $E(\mathcal{E}(\lambda)) = \dfrac{1}{\lambda}$, on a par linéarité $E(T_n) = \dfrac{1}{n\alpha}\displaystyle\sum_{k=1}^n \frac{1}{k}$ et on en déduit classiquement (par comparaison série–intégrale) :

$$E(T_n) \underset{(\infty)}{\sim} \frac{\ln n}{n\alpha}$$

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

b) Comme $V\big(\mathcal{E}(\lambda)\big) = \dfrac{1}{\lambda^2}$, on a par indépendance des X_k :

$$V(T_n) = \frac{1}{n^2}\sum_{k=1}^n V(X_k) = \frac{1}{\alpha^2 n^2}\sum_{k=1}^n \frac{1}{k^2}$$

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

4. La fonction de répartition F_n de T_n est nulle sur $\mathbb{R}^-$ et, pour $x \geqslant 0$,

$$F_n(x) = P(S_n \leq n\,x) = \Phi_n(n\,x) = \big(1-\mathrm{e}^{-\alpha n\,x}\big)^n$$

À $x > 0$ fixé, un développement limité donne :
$F_n(x) = \exp\big[n\ln\big(1-\mathrm{e}^{-\alpha nx}\big)\big] = \exp\big[-n\mathrm{e}^{-\alpha nx} + o\big(n\mathrm{e}^{-\alpha n\,x}\big)\big] \to \mathrm{e}^0 = 1$
Donc la suite $(T_n)_{n\in\mathbb{N}^*}$ tend en loi vers la variable constante égale à 0.

Exercice 3.8.

Une boîte contient n jetons numérotés de 1 à n. On tire un jeton au hasard, on note son numéro et on le remet dans la boîte. Si le numéro de ce jeton est i, alors on tire au hasard et sans remise i jetons de la boîte que l'on distribue au hasard dans trois urnes U_1, U_2 et U_3 (vides au départ). Pour tout k de $[\![1,3]\!]$, on note X_k la variable aléatoire désignant le nombre de jetons de l'urne U_k après cette opération.

1. Soit X la variable aléatoire donnant le numéro du jeton que l'on a tiré au départ dans la boîte. Quelle est la loi de X ? Déterminer la loi conjointe du couple (X_k, X), où $k \in [\![1,3]\!]$.

2. Calculer pour tout $k \in [\![1,3]\!]$, l'espérance de X_k.

3.a) Trouver la loi du vecteur aléatoire (X_1, X_2, X).

b) En déduire la loi du vecteur aléatoire (X_1, X_2, X_3).

4. On définit pour tout $k \in [\![1,3]\!]$, la variable aléatoire Y_k par : $Y_k = \dfrac{X_k}{X}$.

Calculer les espérances des variables aléatoires Y_k et $(Y_k)^2$. Donner un équivalent de la variance de Y_k, lorsque n tend vers l'infini.

Solution :

1. X suit la loi uniforme sur $\{1,\dots,n\}$. La variable X_k est à valeur $\{0,\dots,n\}$. Soit $(i,j)\in\{0,\dots,n\}\times\{1,\dots,n\}$. On commence par remarquer que $P\left((X_k=i)\cap(X=j)\right)=0$ si $i>j$.
Pour $i\leqslant j$, on a :

$$P((X_k=i)\cap(X=j))=P_{(X=j)}(X_k=i)P(X=j)=\frac{1}{n}P_{(X=j)}(X_k=i)$$

La loi de X_k sachant $X=j$ est la loi binomiale $\mathcal{B}(j,1/3)$. D'où :

$$P_{(X=j)}(X_k=i)=\binom{j}{i}\left(\frac{1}{3}\right)^i\left(\frac{2}{3}\right)^{j-i}.$$

D'où :

$$P((X_k=i)\cap(X=j))=\begin{cases}0 & \text{si } i>j\\ \dfrac{1}{n}\dbinom{j}{i}\dfrac{2^{j-i}}{3^j} & \text{si } 0\leqslant i\leqslant j\end{cases}$$

2. On remarque que $X_1+X_2+X_3=X$. Or $E(X_1)=E(X_2)=E(X_3)$, d'où

$$E(X_k)=\frac{E(X)}{3}=\frac{n+1}{6}$$

3.a) Soit $(i,j,k)\in\{0,\dots,n\}^2\times\{1,\dots,n\}$.
On observe que $P((X_1=i)\cap(X_2=j)\cap(X=k))=0$ si $i+j>k$.
Si $i+j\leqslant k$, on a :

$$P_{(X=k)}((X_1=i)\cap(X_2=j))=\frac{1}{3^k}\binom{k}{i}\binom{k-i}{j}$$

D'où :

$$P((X_1=i)\cap(X_2=j)\cap(X=k))=\begin{cases}0 & \text{si } i+j>k\\ \dfrac{1}{n}\dfrac{k!}{i!j!(k-i-j)!}\left(\dfrac{1}{3}\right)^k & \text{si } i+j\leqslant k\end{cases}$$

b) Soit $(i,j,k)\in\{0,\dots,n\}^3$.
Si $i+j+k>n$ on a $P((X_1=i)\cap(X_2=j)\cap(X_3=k))=0$.
Lorsque $i+j+k\leqslant n$, on voit que :

$$\begin{aligned}p_{i,j,k}&=P((X_1=i)\cap(X_2=j)\cap(X_3=k))\\&=P((X_1=i)\cap(X_2=j)\cap(X=i+j+k))\\&=\frac{1}{n}\frac{(i+j+k)!}{i!j!k!}\left(\frac{1}{3}\right)^{i+j+k}\end{aligned}$$

4. La variable Y_k est à valeur dans $\Omega=\{a/b;(a,b)\in\{0,\dots,n\}\times\{1,\dots,n\}\}$. On observe que $Y_1+Y_2+Y_3=1$. Par ailleurs et par symétrie, on a :

$$E(Y_1)=E(Y_2)=E(Y_3),\text{ d'où } E(Y_k)=1/3.$$

Utilisons la formule de l'espérance totale par rapport au système complet d'événements $(X=j)$ pour $j=1,\dots,n$. Il vient :

$$E(Y_k^2)=\sum_{j=1}^n E_{(X=j)}(Y_k^2)P(X=j)=\frac{1}{n}\sum_{j=1}^n\tfrac{1}{j^2}E_{(X=j)}(X_k^2)$$

Or nous avons remarqué que la loi de X_k sachant $X=j$ est la loi binomiale $\mathcal{B}(j,1/3)$ d'espérance $\dfrac{j}{3}$ et de variance $\dfrac{2j}{9}$, d'où :

$$E(Y_k^2)=\frac{1}{n}\sum_{j=1}^n\frac{1}{j^2}\left(\frac{2j}{9}+\left(\frac{j}{3}\right)^2\right)=\frac{1}{n}\sum_{j=1}^n\frac{1}{j}\left(\frac{2}{9}+j\frac{1}{9}\right)=\frac{2}{9n}\sum_{j=1}^n\frac{1}{j}+\frac{1}{9}.$$

Avec la formule de Huygens, on obtient :

$$V(Y_k)=E(Y_k^2)-E(Y_k)^2=\frac{2}{9n}\sum_{j=1}^n\frac{1}{j}$$

On démontre, par comparaison série-intégrale, que l'on a $\sum_{j=1}^{n} \frac{1}{j} \sim \ln n$, et on obtient :

$$V(Y_k) \underset{(n\to\infty)}{\sim} \frac{2\ln n}{9n}$$

Exercice 3.9.

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, toutes de même loi uniforme sur le segment $[0,1]$.

Pour $n \in \mathbb{N}$ et $\omega \in \Omega$, on réordonne par ordre croissant les réels

$$X_1(\omega), X_2(\omega), \dots, X_{2n+1}(\omega)$$

et on note $M_n(\omega)$ le terme médian, c'est-à-dire le $(n+1)^{\text{ème}}$ de ces termes dans l'ordre croissant.

1. Pour tout réel x de $[0,1]$, exprimer $P(M_n \leqslant x)$ à l'aide d'une somme.

2. En considérant les variables aléatoires $X'_k = 1 - X_k$ pour $k \in \mathbb{N}^*$, montrer que l'espérance de la variable aléatoire M_n est égale à $\frac{1}{2}$.

3. Prouver l'égalité : $E(M_n) = \int_0^1 P(M_n > x)\, dx$.

4. Pour tous entiers naturels k et m tels que $0 \leqslant k \leqslant m$ calculer l'intégrale : $I_{k,m} = \int_0^1 x^k(1-x)^{m-k}\, dx$.

5. Retrouver ainsi, par un calcul, la valeur de l'espérance de M_n.

Solution :

1. Soit x élément de $[0,1]$ fixé. On définit la variable aléatoire Z_x par :

$$\forall\, \omega \in \Omega \qquad Z_x(\omega) = \text{card}\{i \in [\![1, 2n+1]\!] \qquad X_i(\omega) \leqslant x\}$$

La variable Z_x suit la loi binomiale de paramètres $(2n+1, x)$. On a donc :

$$\forall\, k \in [\![0, 2n+1]\!], P(Z_x = k) = \binom{2n+1}{k} x^k (1-x)^{2n+1-k}$$

Ainsi : $P(M_n \leqslant x) = P(Z_x \geqslant n+1) = \sum_{k=n+1}^{2n+1} P(Z_x = k)$

En résumé :

$$P(M_n \leqslant x) = \begin{cases} \sum_{k=n+1}^{2n+1} \binom{2n+1}{k} x^k (1-x)^{2n+1-k} & \text{si } x \in [0,1] \\ 0 & \text{si } x < 0 \\ 1 & \text{si } x > 1 \end{cases}$$

Sous cette forme, il apparaît que la fonction de répartition de M_n est continue sur $\mathbb{R}$, dérivable à dérivée continue sur $\mathbb{R} \setminus \{0,1\}$. M_n est une variable à densité. Elle admet une espérance car son univers image est inclus dans $[0,1]$.

2. Si l'on range les valeurs $X_1(\omega), \dots, X_{2n+1}(\omega)$ dans l'ordre croissant sous la forme :

$$Y_1(\omega) \leqslant \cdots \leqslant Y_n(\omega) \leqslant Y_{n+1}(\omega) \leqslant Y_{n+2}(\omega) \leqslant \cdots \leqslant Y_{2n+1}(\omega)$$

on obtient : $M_n(\omega) = Y_{n+1}(\omega)$.

On a :

$$1 - Y_{2n+1}(\omega) \leqslant \cdots \leqslant 1 - Y_{n+1}(\omega) \leqslant 1 - Y_n(\omega) \leqslant \cdots \leqslant 1 - Y_1(\omega)$$

On définit alors $M'_n(\omega)$, terme médian de $X'_1(\omega)$, ... , $X'_{2n+1}(\omega)$. Il vient :

$$\forall\, \omega \in \Omega, M'_n(\omega) = 1 - Y_{n+1}(\omega) = 1 - M_n(\omega)$$

d'où l'on déduit que $M_n + M'_n = 1$ et donc $E(M_n) + E(M'_n) = 1$.

Par ailleurs, X_k et $1 - X_k$ ont la même loi, donc il en va de même de M_n et M'_n. Leurs espérances sont donc égales et toutes deux égales à $1/2$.

3. Soit f_n une densité de M_n. Elle est nulle à l'extérieur de l'intervalle $[0,1]$. Une primitive de f_n est la fonction de répartition F_n de M_n trouvée dans la première question. Mais on peut prendre pour primitive $F_n(x)-1$, *i.e.* $-P(M_n > x)$. Il vient, par intégration par parties :

$$E(M_n) = \int_0^1 x f_n(x)dx = \big[-xP(M_n > x)\big]_0^1 + \int_0^1 P(M_n > x)dx$$

et le résultat demandé se déduit du fait que $P(M_n > 1) = 0$.

4. Supposons tout d'abord $k < m$. En intégrant par parties :

$$I_{k,m} = \int_0^1 x^k(1-x)^{m-k}dx = \frac{m-k}{k+1}\int_0^1 x^{k+1}(1-x)^{m-(k+1)}dx = \frac{m-k}{k+1}I_{k+1,m}$$

D'où :

$$I_{k,m} = \frac{m-k}{k+1}\times\frac{m-(k+1)}{k+2}\times\cdots\times\frac{1}{m}I_{k+(m-k),m}, \text{ donc :}$$

$$I_{k,m} = \frac{(m-k)!k!}{m!}I_{m,m}$$

relation valable lorsque $k = m$.

Comme, en outre, $I_{m,m} = \displaystyle\int_0^1 x^m dx = \frac{1}{m+1}$ on a, finalement :

$$I_{k,m} = \frac{1}{m+1}\times\frac{(m-k)!k!}{m!} = \frac{1}{m+1}\binom{m}{k}^{-1}$$

5. On a :

$$E(M_n) = \int_0^1 [1 - P(M_n \leqslant x)]dx = 1 - \int_0^1 \Big[\sum_{k=n+1}^{2n+1}\tbinom{2n+1}{k}x^k(1-x)^{2n+1-k}\Big]dx$$

donc : $E(M_n) = 1 - \displaystyle\sum_{k=n+1}^{2n+1}\tbinom{2n+1}{k}I_{k,2n+1} = 1 - \sum_{k=n+1}^{2n+1}\tbinom{2n+1}{k}\tbinom{2n+1}{k}^{-1}\frac{1}{2n+2}$, d'où :

$$E(M_n) = 1 - \sum_{k=n+1}^{2n+1}\frac{1}{2n+2} = \frac{1}{2}$$

Exercice 3.10.

Soit n un entier $\geqslant 2$ et $x_1, x_2, \ldots, x_n$ des nombres réels donnés non tous égaux. On pose : $M = \max(x_1, x_2, \ldots, x_n)$ et $m = \min(x_1, x_2, \ldots, x_n)$.

1. Soit Y une variable aléatoire à densité, définie sur un espace probabilisé $\big(\Omega, \mathcal{A}, P\big)$, qui suit la loi uniforme sur le segment $[m, M]$. Pour tout $i \in [\![1,n]\!]$, on considère l'événement $A_i = [Y \leqslant x_i]$.

a) Déterminer l'événement $\displaystyle\bigcup_{i=1}^n A_i$; en déduire sa probabilité.

b) Pour tout $i \in [\![1,n]\!]$, calculer $P(A_i)$. Pour tout $(i,j) \in [\![1,n]\!]^2$ avec $i \neq j$, calculer $P(A_i \cap A_j)$. Pour toute partie I de $[\![1,n]\!]$, exprimer $P\big(\displaystyle\bigcap_{i\in I} A_i\big)$ à l'aide de $\displaystyle\min_{i\in I} x_i$.

c) Montrer que :

$$\max(x_1, x_2, \ldots, x_n) = \sum_{i=1}^n (-1)^{i-1}\sum_{1\leq k_1<\cdots<k_i\leq n}\min(x_{k_1}, x_{k_2}, \ldots, x_{k_i})$$

2. Dans cette question, les nombres $x_1, x_2, \ldots, x_n$ sont supposés strictement positifs.

a) Montrer que $\displaystyle\lim_{N\to+\infty}(x_1^N + \cdots + x_n^N)^{\frac{1}{N}} = M$.

b) En déduire l'existence et la valeur de $\displaystyle\lim_{N\to+\infty}(x_1^{-N} + \cdots + x_n^{-N})^{-\frac{1}{N}}$.

Solution :

1. a) On a $\displaystyle\bigcup_{i=1}^n A_i = \Omega$ de probabilité 1.

b) De même $P(A_i) = \dfrac{x_i - m}{M - m}$;

$P(A_i \cap A_j) = P(Y \leqslant \min(x_i, x_j)) = \dfrac{\min(x_i, x_j) - m}{M - m}$

Plus généralement :

$$P\big(\bigcap_{i\in I} A_i\big) = \frac{\min\limits_{i\in I} x_i - m}{M - m}$$

c) D'après la formule du crible on a :

$$\begin{aligned}
P\big(\bigcup_{i\in I} A_i\big) &= \sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} P\big(\bigcap_{j=1}^{i} A_{k_j}\big)\\
&= \sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} \frac{\min(x_{k_1}, \ldots, x_{k_i}) - m}{M - m}
\end{aligned}$$

$$\begin{aligned}
P\big(\bigcup_{i\in I} A_i\big) &= \frac{1}{M-m}\Big(\sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} \min(x_{k_1}, \ldots, x_{k_i})\Big)\\
&\qquad - \frac{m}{M-m} \sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} 1\\
&= \frac{1}{M-m}\Big(\sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} \min(x_{k_1}, \ldots, x_{k_i})\Big)\\
&\qquad - \frac{m}{M-m} \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i}\\
&= \frac{1}{M-m}\Big(\sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} \min(x_{k_1}, \ldots, x_{k_i})\Big) - \frac{m}{M-m}
\end{aligned}$$

(la dernière égalité étant obtenue par la formule du binôme), d'où :

$$\sum_{i=1}^{n} (-1)^{i-1} \sum_{1\leq k_1<\cdots<k_i\leq n} \min(x_{k_1}, \ldots, x_{k_i}) - m = M - m$$

et le résultat annoncé.

2. a) On a : $\dfrac{1}{M}(x_1^N + \cdots + x_n^N)^{\frac{1}{N}} = \exp\Big[\dfrac{1}{N}\ln\Big(\big(\dfrac{x_1}{M}\big)^N + \cdots + \big(\dfrac{x_n}{M}\big)^N\Big)\Big]$

Or, pour tout i on a :

$$\lim_{N\to+\infty} \big(\frac{x_i}{M}\big)^N = \begin{cases} 0 & \text{si } x_i < M\\ 1 & \text{si } x_i = M\end{cases}$$

Donc : $\lim\limits_{N\to+\infty} \dfrac{1}{N}\ln\Big(\big(\dfrac{x_1}{M}\big)^N + \cdots + \big(\dfrac{x_n}{M}\big)^N\Big) = 0$ (car la quantité placée sous le logarithme a une limite appartenant à $\mathbb{N}^*$), d'où le résultat annoncé.

b) On a : $(x_1^{-N} + \cdots + x_n^{-N})^{-\frac{1}{N}} = \dfrac{1}{\big[(\frac{1}{x_1})^N + \cdots + (\frac{1}{x_n})^N\big]^{\frac{1}{N}}}$

D'après la question 2.a), le dénominateur tend vers $\max\big(\dfrac{1}{x_1}, \ldots, \dfrac{1}{x_n}\big)$ donc :

$$\lim_{N\to+\infty} (x_1^{-N} + \cdots + x_n^{-N})^{-\frac{1}{N}} = \frac{1}{\max\big(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\big)} = m$$

Exercice 3.11.

Soit J la matrice de $\mathcal{M}_n(\mathbb{C})$ définie par $J = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0\\ 0 & 0 & 1 & \ldots & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & 1\\ 1 & 0 & \ldots & 0 & 0\end{pmatrix}$,

c'est-à-dire par : $(J)_{p,q} = \begin{cases} 1 & \text{si } q - p = 1\\ 1 & \text{si } p = n, q = 1\\ 0 & \text{sinon}\end{cases}$, où $n \geqslant 2$.

1. Calculer J^n. En déduire les valeurs propres réelles ou complexes possibles de J.

2. Déterminer les valeurs propres de J, en exhibant pour chacune un vecteur propre (réel ou complexe) associé. En déduire que la matrice J est diagonalisable sur $\mathbb{C}$.

3. Dans cette question n désigne un **entier naturel impair supérieur ou égal à 3**. On note $A_0, A_1, \dots, A_{n-1}$ les n points du plan, d'affixes respectives $z_k = \mathrm{e}^{2ik\pi/n}$, pour $k \in [\![0, n-1]\!]$.
- au temps $t = 0$, une puce se trouve en A_0.
- si à l'instant t, elle se trouve en A_k, pour $k \in [\![1, n-2]\!]$, elle se trouvera à l'instant $t+1$, soit en A_{k-1} soit en A_{k+1}, ceci avec équiprobabilité.
- si à l'instant t, elle se trouve en A_0, elle se trouvera à l'instant $t+1$, soit en A_{n-1} soit en A_1, ceci avec équiprobabilité.
- si à l'instant t, elle se trouve en A_{n-1}, elle se trouvera à l'instant $t+1$, soit en A_{n-2} soit en A_0, ceci avec équiprobabilité.

Pour tout m de $\mathbb{N}$, on note X_m la variable aléatoire définie par :

$$\forall\, k \in [\![0, n-1]\!], (X_m = k) = \text{« la puce est en } A_k \text{ à l'instant } m \text{ »}.$$

On pose, pour tout m de $\mathbb{N}$, $U_m = \begin{pmatrix} P(X_m = 0) \\ P(X_m = 1) \\ \vdots \\ P(X_m = n-1) \end{pmatrix}$.

a) Déterminer U_0

b) Déterminer une matrice A de $\mathcal{M}_n(\mathbb{R})$ telle que pour tout m : $U_{m+1} = AU_m$.

c) Exprimer A à l'aide de puissances de la matrice J.

d) En déduire les valeurs propres de la matrice A. Montrer que la matrice A est diagonalisable et que l'on peut choisir la matrice de passage diagonalisante P orthogonale. Déterminer un vecteur propre associé à la valeur propre de module maximal.

e) Déterminer la limite en loi de la suite de variables aléatoires $(X_m)_{m\in\mathbb{N}}$.

Solution :

1. L'endomorphisme j canoniquement associé à J est l'endomorphisme de la permutation circulaire $(e_1, e_2, \dots, e_n) \mapsto (e_n, e_1, \dots, e_{n-1})$. Donc $j^n = Id$ et $J^n = I$.
Les valeurs propres de J sont incluses dans l'ensemble des racines du polynôme $X^n - 1$, soit $\{\mathrm{e}^{2ik\pi/n}, k \in [\![0, n-1]\!]\}$.

2. On résout le système d'équation $JX = \lambda X$, où $\lambda = \mathrm{e}^{2ik\pi/n}$ et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Cela donne :

$$\left\{\begin{array}{c} x_2 = \lambda x_1 \\ x_3 = \lambda x_2 \\ \vdots \\ x_n = \lambda x_{n-1} \\ x_1 = \lambda x_n \end{array}\right., \text{ soit : } X = x_1 \begin{pmatrix} 1, \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{n-1} \end{pmatrix}$$

Ainsi, sur $\mathbb{C}$, J admet n valeurs propres distinctes et est diagonalisable. Chaque sous-espace propre est de dimension 1.

3. a) $U_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

b) c) On a par la formule des probabilités totales (en supposant les événements $(X_m = k)$ de probabilités non nulles) :

$$P(X_{m+1}=j)=\sum_{k=0}^{n-1}P_{(X_m=k)}(X_{m+1}=j)P(X_m=k)$$

Toutes les probabilités conditionnelles sont nulles sauf deux qui valent $1/2$. On obtient ainsi

$$A=\frac{1}{2}\begin{pmatrix}0&1&0&\dots&1\\1&0&1&\dots&0\\\vdots&\ddots&\ddots&\ddots&\vdots\\0&\ddots&\ddots&\ddots&1\\1&0&\dots&1&0\end{pmatrix}=\frac{1}{2}(J+J^{n-1})=\frac{1}{2}(J+J^{-1})$$

Le résultat restant valable même en présence d'événements quasi-impossibles.

d) Les valeurs propres de A sont $\frac{1}{2}\big(\mathrm{e}^{2ik\pi/n}+\mathrm{e}^{-2ik\pi/n}\big)=\cos(2k\pi/n)$, pour $k\in[\![0,n-1]\!]$, les vecteurs propres associés étant ceux obtenus en 2.

La matrice A étant symétrique réelle, elle est diagonalisable. Elle est semblable, avec une matrice de passage orthogonale, à la matrice diagonale

$$\mathrm{diag}(1,\cos(2\pi/n),\cos(4\pi/n),\dots,\cos((2n-2)\pi/n)).$$

La valeur propre 1 est dominante, les autres valeurs propres étant en valeur absolue strictement inférieures à 1 (ceci parce que n est impair, donc -1 n'est pas valeur propre de A).

Le vecteur propre unitaire associé à $\lambda=1$ est $X=\dfrac{1}{\sqrt{n}}\begin{pmatrix}1\\\vdots\\1\end{pmatrix}$.

e) On a $A=PD^tP$, la première colonne de P étant constituée du vecteur colonne X.
Pour tout m de $\mathbb{N}$, $A^m=PD^m\,{}^tP$.
De plus, $\lim\limits_{m\to+\infty}D^m=\mathrm{diag}(1,0,\dots,0)$. Ainsi :

$$\begin{aligned}\lim_{m\to+\infty}U_m&=\lim_{m\to+\infty}A^mU_0=P\,\mathrm{diag}(1,0,\dots,0)^tPU_0\\&=\begin{pmatrix}1/\sqrt{n}&0&\dots&0\\\vdots&\vdots&&\vdots\\1/\sqrt{n}&0&\dots&0\end{pmatrix}\begin{pmatrix}1/\sqrt{n}&\dots&1/\sqrt{n}\\ *&\dots&*\\\vdots&&\vdots\\ \star&\star&\star\end{pmatrix}U_0=\begin{pmatrix}1/n\\\vdots\\1/n\end{pmatrix}\end{aligned}$$

Ainsi, la suite (X_m) tend en loi vers la loi uniforme sur $[\![0,n-1]\!]$

Exercice 3.12.

Soient n et m deux entiers naturels non nuls. Une urne contient n boules rouges et m boules bleues. Les boules rouges sont numérotées de 1 à n. Les boules sont tirées (sans remise) au hasard et une à une de l'urne jusqu'à ce qu'une boule bleue soit tirée. On note alors les numéros des boules rouges qui ont été tirées. Soit X le plus grand de ces numéros et Y le plus petit. Si la première boule tirée est bleue, on pose $X=Y=0$. On note également T le nombre de boules rouges tirées. L'expérience est modélisée à l'aide d'un espace probabilisé $(\Omega,\mathcal{B},P)$.

1. Soient A,B,C trois événements. Montrer que :

$$P(\overline{A}\cap B\cap C)=P(\overline{A})-P(\overline{A}\cap\overline{B})-P(\overline{A}\cap\overline{C})+P(\overline{A}\cap\overline{B}\cap\overline{C}).$$

2. Étude de T.

 a) Déterminer l'ensemble des valeurs prises par T.

 b) Calculer $P(T=0),P(T=1)$.

 c) Soit $k\in[\![2,n]\!]$. Calculer $P(T=k)$.

3. Soit R une partie de l'ensemble des boules rouges de cardinal r. Montrer que la probabilité q_r qu'aucune boule rouge de R n'ait été tirée au cours du jeu vaut $\dfrac{m}{m+r}$.

4. Soient $i,j\in[\![0,n]\!]$. On cherche à déterminer $p_{i,j}=P[(X=i)\cap(Y=j)]$.

 a) Calculer $p_{0,0}$.

b) On suppose que $i = j \neq 0$. Montrer que $p_{i,j} = \dfrac{m}{(n+m)(n+m-1)}$.

c) On suppose que $i > j > 0$. On note $t = i - j$. Montrer que

$$p_{i,j} = \frac{2m}{(n+m-t-1)(n+m-t)(n+m-t+1)}.$$

Solution :

1. Par la formule d'inclusion-exclusion :

$$P(\overline{A} \cap B \cap C) = P(\overline{A}) - P(\overline{A} \cap (\overline{B} \cup \overline{C})) = P(\overline{A}) - P((\overline{A} \cap \overline{B}) \cup (\overline{A} \cap \overline{C}))$$
$$= P(\overline{A}) - P(\overline{A} \cap \overline{B}) - P(\overline{A} \cap \overline{C}) + P(\overline{A} \cap \overline{B} \cap \overline{C})$$

2. a) Clairement $T(\Omega) = [\![0, n]\!]$.

b) $\star$ Le nombre de boules rouges tirées vaut 0 si et seulement si la première boule tirée est bleue. Ainsi, $P(T = 0) = \dfrac{m}{n+m}$.

$\star$ Le nombre de boules rouges tirées vaut 1 si et seulement si la première boule tirée est rouge et la deuxième est bleue. Ainsi,

$$P(T = 1) = \frac{n}{n+m} \times \frac{m}{n+m-1} = \frac{nm}{(n+m)(n+m-1)}.$$

c) En reprenant l'argumentaire précédent et pour $k \geqslant 1$:

$$P(T = k) = \frac{n}{n+m} \times \frac{n-1}{n+m-1} \times \cdots \times \frac{n-k+1}{n+m-k+1} \times \frac{m}{n+m-k}$$
$$= \frac{m.n!(n+m-k-1)!}{(n-k)!(n+m)!}$$

3. La probabilité qu'aucune boule de R n'ait été tirée au cours du jeu est égale à la probabilité que la première boule tirée parmi la réunion des boules bleues et de R soit une boule bleue. Ainsi, par symétrie, cette probabilité vaut $q_r = \dfrac{m}{m+r}$.

4. a) Lorsque $i = j = 0$, aucune boule rouge n'a été tirée. Ainsi, d'après la question précédente, $p_{0,0} = \dfrac{m}{n+m}$.

b) Lorsque $i = j > 0$, les boules rouges ayant des numéros deux à deux distincts, seule la boule numérotée j a été tirée. Ainsi, la première boule tirée est la boule rouge numéro j et la deuxième boule tirée est une boule bleue. Ainsi,

$$p_{j,j} = \frac{1}{n+m} \times \frac{m}{n+m-1} = \frac{m}{(n+m)(n+m-1)}.$$

c) Lorsque $i > j$, les boules rouges numérotées i et j sont tirées et aucune boule rouge dont le numéro appartient à $[\![1, i-1]\!] \cup [\![j+1, n]\!]$ n'est tirée.

Ainsi, en utilisant la formule obtenue dans la première question (avec $\overline{A} =$ « ne tirer aucune boule rouge de numéro appartenant à $[\![1, i-1]\!] \cup [\![j+1, n]\!]$ », $B =$ « tirer la boule rouge numéro i », $C =$ « tirer la boule rouge numéro j » et en appliquant trois fois le résultat de la question 3.

$$p_{i,j} = q_{n+m-t-1} - 2q_{n+m-t} + q_{n+m-t+1}$$
$$= m\frac{(n+m-t)^2 + n + m - t - 2((n+m-t)^2 - 1) + (n+m-t)^2 - (n+m-t)}{(n+m-t-1)(n+m-t)(n+m-t+1)}$$
$$= \frac{2m}{(n+m-t-1)(n+m-t)(n+m-t+1)}$$

Exercice 3.13.

Soit $(\Omega, \mathcal{A}, P)$, un espace probabilisé sur lequel sont définies toutes les variables aléatoires de cet exercice. Soit $(Y_k)_{k \geq 1}$, une suite de variables aléatoires indépendantes et de même loi exponentielle de paramètre $b > 0$. Soit N une variable aléatoire suivant la loi géométrique de paramètre s, où $0 < s < 1$, indépendante des variables aléatoires Y_k.

On pose $Z = \sup(Y_1, Y_2, \ldots, Y_N)$. On admet que Z est une variable aléatoire à densité.

1. a) Déterminer une densité de Z.

b) Montrer que Z admet une espérance $E(Z)$ et montrer que $E(Z) = \dfrac{-\ln s}{b(1-s)}$.

2. Soit g la fonction définie sur $\mathbb{R}^+$ par : $g(t) = \begin{cases} 1 & \text{si } t = 0 \\ \dfrac{te^{-t}}{1-e^{-t}} & \text{si } t > 0 \end{cases}$

a) Montrer que g est continue et bornée sur $\mathbb{R}^+$.

b) Montrer que : $\forall\, t \geqslant 0, \forall\, n \in \mathbb{N},\ g(t) = g(t)e^{-(n+1)t} + \sum\limits_{k=0}^{n} t.e^{-(k+1)t}$.

3. a) Justifier que pour tout k de $\mathbb{N}$, $\displaystyle\int_0^{+\infty} t.e^{-(k+1)t}dt$ est convergente et la calculer.

b) Montrer que $\displaystyle\int_0^{+\infty} g(t)dt$ est convergente et égale à $\sum\limits_{k=1}^{+\infty} \dfrac{1}{k^2}$.

c) A l'aide du changement de variable $t = -\ln s$, calculer l'intégrale $\displaystyle\int_0^1 E(Z)ds$.

Solution :

1. a) Pour tout $j \geqslant 1$, $P_{(N=j)}(Z \leqslant t) = P_{(N=j)}(\bigcap\limits_{i=1}^{j} (Y_i \leqslant t))$ et par indépendance des variables en jeu :

$$P_{(N=j)}(Z \leqslant t) = \prod_{i=1}^{j} P(Y_i \leqslant t) = (1 - e^{-bt})^j$$

La famille $([N = j])_{j\in\mathbb{N}^*}$ forme un système complet d'événements. Ainsi, pour tout $t > 0$:

$P(Z \leqslant t) = \sum\limits_{j=1}^{\infty} P_{(N=j)}(Z \leqslant t)P(N = j) = \sum\limits_{j=1}^{\infty} (1 - e^{-bt})^j s(1 - s)^{j-1}$

$$P(Z \leqslant t) = \frac{s(1 - e^{-bt})}{1 - (1 - e^{-bt})(1 - s)}$$

car $0 \leqslant (1 - e^{-bt})(1 - s) < 1$. Ainsi, par dérivation, une densité de Z est donnée par :

$$f_Z(t) = \begin{cases} 0 & \text{si } t \leqslant 0 \\ \dfrac{sbe^{-bt}}{(1 - (1 - e^{-bt})(1 - s))^2} & \text{si } t > 0 \end{cases}$$

b) Soit $A > 0$. En intégrant par parties, il vient :

$$\int_0^A tf_Z(t)dt = \big[t(F_Z(t) - 1)\big]_0^A - \int_0^A (F_Z(t) - 1)dt$$

$\displaystyle\int_0^A tf_Z(t)dt = \big[t(F_Z(t) - 1)\big]_0^A - \big[\frac{1}{(1-s)b} \ln(1 - (1 - s)(1 - e^{-bt}))\big]_0^A$

En faisant tendre A vers $+\infty$, il vient :

$$E(Z) = -\frac{\ln s}{(1 - s)b}$$

2. a) Comme $1 - e^{-t} \underset{(0)}{\sim} t$, on a $\lim\limits_{t\to 0} g(t) = 1 = g(0)$ et g est continue en 0, donc sur $\mathbb{R}^+$. De plus $\lim\limits_{t\to+\infty} g(t) = 0$ entraîne que la fonction g est bornée sur $\mathbb{R}^+$.

b) Il suffit de remarquer que pour tout $t \geqslant 0$, $\sum\limits_{k=0}^{n} e^{-(k+1)t} = e^{-t} \times \dfrac{1 - e^{(n+1)t}}{1 - e^{-t}}$.

3. a) Le changement de variable C^1 bijectif, $u = (k + 1)t$ donne :

$$\int_0^{+\infty} te^{-(k+1)t}dt = \frac{1}{(k+1)^2}\int_0^{+\infty} ue^{-u}du = \frac{1}{(k+1)^2}$$

ce qui donne l'existence et le calcul de l'intégrale proposée.

b) On utilise la question 2. b)

$$\int_0^{+\infty} g(t)dt = \sum_{k=0}^{n} \int_0^{+\infty} t\mathrm{e}^{-(k+1)t}dt + \int_0^{+\infty} g(t)\mathrm{e}^{-(n+1)t}dt$$

Or : $|g(t)\mathrm{e}^{-(n+1)t)}| \leqslant M\mathrm{e}^{-(n+1)t)}$ donne $\left|\int_0^{+\infty} g(t)\mathrm{e}^{-(n+1)t)}dt\right| \leqslant \dfrac{M}{n+1}$,

quantité qui tend vers 0 lorsque n tend vers $+\infty$. En résumé :

$$\int_0^{+\infty} g(t)dt = \sum_{k=0}^{\infty} \frac{1}{(k+1)^2}$$

c) On utilise le changement de variable suggéré ($-t = \ln s$), qui est de classe C^1, bijectif de $]0,1[$ sur $]0,+\infty]$. Il vient :

$$\int_0^1 E(Z)ds = \int_0^1 \frac{-\ln s}{b(1-s)}ds = \frac{1}{b}\int_0^{+\infty} \frac{t\mathrm{e}^{-t}}{1-\mathrm{e}^{-t}}dt = \frac{1}{b}\sum_{k=0}^{\infty} \frac{1}{(k+1)^2}$$

Exercice 3.14.

On considère une suite de variables aléatoires réelles mutuellement indépendantes $N, X_1, \dots, X_n, \dots$ définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On suppose que N est une variable aléatoire réelle à valeurs dans $\mathbb{N}^*$ possédant un moment d'ordre 2 et que les variables aléatoires (X_i), $i \in \mathbb{N}^*$, suivent la même loi que X, où X est une variable aléatoire réelle à valeurs dans $\mathbb{N}$ et possédant un moment d'ordre 2.

On note Y la variable aléatoire définie par $Y = \sum\limits_{i=1}^{N} X_i$ c'est-à-dire :

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

1. Déterminer l'espérance $E(Y)$ en fonction de $E(X)$ et de $E(N)$.

2. En utilisant la formule de l'espérance totale, déterminer $E(Y^2)$ en fonction de $E(X)$, $V(X)$, $E(N)$ et $E(N^2)$.

3. En déduire $V(Y)$ en fonction de $E(X)$, $V(X)$, $E(N)$ et $V(N)$.

4. Une urne contient n jetons numérotés de 1 à n et on dispose d'une pièce de monnaie qui donne le côté pile avec la probabilité p, où $0 < p < 1$. Un joueur tire un jeton dans l'urne et lance ensuite la pièce de monnaie autant de fois que le numéro indiqué par le jeton.
Calculer la moyenne et la variance de la variable aléatoire comptabilisant le nombre de pile obtenu.

Solution :

1. La famille $(Y = k)_{k \geq 0}$ constitue un système complet d'événements. Soit n un élément fixé de $N(\Omega)$. On a :

$$E_{(N=n)}(Y) = \sum_{k=0}^{+\infty} kP_{(N=n)}(Y = k)$$

à condition que cette série converge.
Or :

$$\begin{aligned} P_{(N=n)}(Y = k) &= \frac{P((N = n) \cap (Y = k))}{P(N = n)} \\ &= \frac{P((N = n) \cap (X_1 + \cdots + X_n = k))}{P(N = n)} \\ &= P(X_1 + \cdots + X_n = k) \end{aligned}$$

On en déduit la convergence de la série précédente, avec :

$$E_{(N=n)}(Y) = \sum_{k=0}^{+\infty} kP(X_1 + \cdots + X_n = k) = E(X_1 + \cdots + X_n) = nE(X)$$

puis :

$$E(Y) = \sum_{n\in N(\Omega)} E_{(N=n)}(Y)P(N = n) = \sum_{n\in N(\Omega)} nE(X)P(N = n)$$

$= E(X) \sum_{n\in N(\Omega)} nP(N = n)$

Cette dernière série converge et sa somme vaut $E(N)$. On a, finalement :

$$E(Y) = E(N)E(X)$$

2. De façon analogue :

$$E_{(N=n)}(Y^2) = \sum_{k=0}^{+\infty} kP((X_1 + \cdots + X_n)^2 = k) = E((X_1 + \cdots + X_n)^2)$$

donc :

$E_{(N=n)}(Y^2) = V(X_1 + \cdots + X_n) + [E(X_1 + \cdots + X_n)]^2 = nV(X) + (nE(X))^2$

du fait de l'indépendance des variables. Ainsi :

$$\begin{aligned} E(Y^2) &= \sum_{n\in N(\Omega)} E_{(N=n)}(Y^2)P(N = n) \\ &= \sum_{n\in N(\Omega)} nV(X)P(N = n) + \sum_{n\in N(\Omega)} n^2(E(X))^2 P(N = n) \end{aligned}$$

et, finalement : $E(Y^2) = V(X)E(N) + (E(X))^2E(N^2)$.

3. Maintenant :

$V(Y) = E(Y^2) - (E(Y))^2 = V(X)E(N) + (E(X))^2E(N^2) - (E(X))^2(E(N))^2$

d'où : $V(Y) = V(X)E(N) + (E(X))^2V(N)$.

4. La variable N suit la loi uniforme discrète sur $[\![1, n]\!]$. Les variables X_i suivent la loi de Bernoulli de paramètre p. La variable Y représente en fait le nombre de pile obtenu. On a :

$$E(X) = p, E(N) = \frac{n+1}{2}, V(X) = p(1-p) \text{ et } V(N) = \frac{n^2-1}{12}$$

On en déduit : $E(Y) = \dfrac{n+1}{2}p$

puis :

$$V(Y) = p(1-p)\frac{n+1}{2} + p^2\frac{n^2-1}{12} = p\frac{n+1}{12}[(n-7)p + 6]$$

Exercice 3.15.

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ qui suit la loi de Poisson de paramètre réel inconnu $\lambda > 0$.

Pour n entier $\geqslant 1$, soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon *i.i.d.* de la loi de X. On pose :

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

1. a) Rappeler la loi de S_n, ainsi que son espérance et sa variance.

b) Montrer que $\overline{X}_n$ est un estimateur sans biais et convergent de λ.

2. Montrer que $\exp(-\overline{X}_n)$ est un estimateur asymptotiquement sans biais de $\exp(-\lambda)$.

Dans la suite de l'exercice, s désigne un entier naturel.

3. a) Déterminer la loi conditionnelle de X_1 sachant $[S_n = s]$.

b) Soit T_n la variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$ par $T_n = \left(1 - \frac{1}{n}\right)^{S_n}$

Montrer que T_n est un estimateur sans biais de $\exp(-\lambda)$.

c) Calculer la variance de T_n. Montrer que la suite $(T_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers $\exp(-\lambda)$.

4. Soit $(x_1, x_2, \ldots, x_n)$ un n-uplet de $\mathbb{N}^n$.

a) Calculer $P([X_1 = x_1] \cap \cdots \cap [X_n = x_n] \cap [S_n = s])$.

b) Montrer que la probabilité conditionnelle $P_{[S_n=s]}([X_1 = x_1] \cap \cdots \cap [X_n = x_n])$ est indépendante de la valeur de λ.

Solution :

1. a) Par stabilité de la loi de Poisson, il vient $S_n \hookrightarrow \mathcal{P}(n\lambda)$ et $E(S_n) = n\lambda$, $V(S_n) = n\lambda$.

b) Comme $\overline{X}_n$ est une fonction de $(X_1, X_2, \ldots, X_n)$, c'est un estimateur. De plus $E(\overline{X}_n) = \lambda$, il est sans biais et $V(\overline{X}_n) = \dfrac{\lambda}{n}$, il est convergent.

2. a) Par le théorème de transfert :

$$E(\exp(-\overline{X}_n)) = \sum_{k=0}^{\infty} \mathrm{e}^{-k/n}\mathrm{e}^{-n\lambda}\frac{(n\lambda)^k}{k!} = \mathrm{e}^{-n\lambda}\sum_{k=0}^{\infty}\frac{1}{k!}(\mathrm{e}^{-1/n}n\lambda)^k \text{ Soit :}$$

$$E(\exp(-\overline{X}_n)) = \mathrm{e}^{-n\lambda}\exp(\mathrm{e}^{-1/n}n\lambda) = \mathrm{e}^{-n\lambda(1-\mathrm{e}^{-1/n})}$$

Ainsi $\overline{X}_n$ est un estimateur asymptotiquement sans biais de $\exp(-\lambda)$, puisque $\lim\limits_{n\to\infty} n(1-\mathrm{e}^{-1/n}) = -1$ donne $\lim\limits_{n\to+\infty} E(\overline{X}_n) = \exp(-\lambda)$.

3. a) Si $k \notin [\![0, s]\!]$, alors $P_{[S_n=s]}(X_1 = k) = 0$.

Si $k \in [\![0, s]\!]$, alors $P_{[S_n=s]}(X_1 = k) = \dfrac{P([S_n = s] \cap [X_1 = k])}{P([S_n = s])}$. Donc,

$$\begin{aligned}
P_{[S_n=s]}(X_1 = k) &= \frac{P([X_1 = k] \cap [X_2 + \cdots + X_n = s-k])}{P([S_n = s])}\\
&= \frac{P(X_1 = k)P(X_2 + \cdots + X_n = s-k)}{P(S_n = s)}\\
&= \frac{\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}\mathrm{e}^{-(n-1)\lambda}\frac{((n-1)\lambda)^{s-k}}{(s-k)!}}{\mathrm{e}^{-n\lambda}\frac{(n\lambda)^s}{s!}}
\end{aligned}$$

$$P_{[S_n=s]}(X_1 = k) = \binom{s}{k}\Big(\frac{1}{n}\Big)^k\Big(1-\frac{1}{n}\Big)^{s-k}$$

en remarquant que $X_2 + \cdots + X_n$ suit la loi de Poisson de paramètre $(n-1)\lambda$, et après simplifications. Ainsi, la loi conditionnelle de X_1 sachant $[S_n = s]$ est la loi binomiale $\mathcal{B}(s, 1/n)$.

b) Par le théorème du transfert :

$$E(T_n) = \sum_{s=0}^{\infty}\Big(1-\frac{1}{n}\Big)^s\mathrm{e}^{-n\lambda}\frac{(n\lambda)^s}{s!} = \mathrm{e}^{-n\lambda}\sum_{s=0}^{\infty}\frac{((n-1)\lambda)^s}{s!} = \mathrm{e}^{-n\lambda}\times\mathrm{e}^{(n-1)\lambda}$$

$$E(T_n) = \mathrm{e}^{-\lambda}$$

ce qui montre que T_n est un estimateur sans biais de $\exp(-\lambda)$.

c) De même :

$$\begin{aligned}
E(T_n^2) &= \sum_{s=0}^{\infty}\Big(1-\frac{1}{n}\Big)^{2s}\mathrm{e}^{-n\lambda}\frac{(n\lambda)^s}{s!} = \mathrm{e}^{-n\lambda}\sum_{s=0}^{\infty}\frac{1}{s!}((1-\frac{1}{n})^2 n\lambda)^s\\
&= \mathrm{e}^{-n\lambda}\mathrm{e}^{n\lambda(1-\frac{1}{n})^2} = \mathrm{e}^{-2\lambda+\frac{\lambda}{n}}
\end{aligned}$$

et donc

$$V(T_n) = \mathrm{e}^{-2\lambda}(\mathrm{e}^{\lambda/n} - 1)$$

Par l'inégalité de Tchébichev, pour tout $\varepsilon > 0$:

$$P(|T_n - \mathrm{e}^{-\lambda}| \geqslant \varepsilon) \leqslant \frac{V(T_n)}{\varepsilon^2} = \frac{\mathrm{e}^{-2\lambda}}{\varepsilon^2}\big(\mathrm{e}^{\lambda/n} - 1\big)$$

quantité qui tend vers 0 lorsque n tend vers $+\infty$. La suite $(T_n)_{n\in\mathbb{N}^*}$ converge donc en probabilité vers la variable certaine égale à $\exp(-\lambda)$.

4. a) On a

$$([X_1 = x_1] \cap [X_2 = x_2] \cap \cdots \cap [X_n = x_n] \cap [S_n = s]) =$$
$$([X_1 = x_1] \cap [X_2 = x_2] \cap \cdots \cap [X_{n-1} = x_{n-1}] \cap [X_n = s - \sum_{i=1}^{n-1} x_i])$$

Par indépendance des X_i, on en déduit :

$$\delta = P([X_1 = x_1] \cap [X_2 = x_2] \cap \cdots \cap [X_n = x_n] \cap [S_n = s])$$
$$= \begin{cases} 0 & \text{si } s \neq \sum\limits_{i=1}^{n} x_i \\ \mathrm{e}^{-n\lambda} \dfrac{\lambda^s}{x_1! \dots x_n!} & \text{si } s = \sum\limits_{i=1}^{n} x_i \end{cases}$$

b) Par définition des probabilités conditionnelles :

$$P_{[S_n=s]}([X_1 = x_1] \cap \cdots \cap [X_n = x_n]) = \begin{cases} 0 & \text{si } s \neq \sum\limits_{i=1}^{n} x_i \\ \dfrac{s!}{x_1! \dots x_n!} \left(\dfrac{1}{n}\right)^s & \text{si } s = \sum\limits_{i=1}^{n} x_i \end{cases}$$

Exercice 3.16.

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé sur lequel sont définies toutes les variables aléatoires de l'exercice.

Soit N une variable aléatoire suivant la loi de Poisson de paramètre $\lambda > 0$.

1. a) Montrer que, pour toute fonction g définie sur $\mathbb{N}$ telle que les espérances existent, on a :

$$E\big[N\,g(N)\big] = \lambda E\big[g(N+1)\big]$$

b) Calculer $E\big(\dfrac{1}{N+1}\big)$.

2. Soit T une variable aléatoire à valeurs dans $\mathbb{N}$ telle qu'il existe un réel λ vérifiant, pour toute fonction g telle que les espérances existent,

$$E\big[Tg(T)\big] = \lambda E\big[g(T+1)\big]$$

La variable aléatoire T suit-elle une loi de Poisson ?

3. Soit $(X_k)_{k\in\mathbb{N}}$ une suite de variables aléatoires mutuellement indépendantes, de même loi, à valeurs dans $\mathbb{N}$. On définit une variable aléatoire S par :

$$S = \sum_{k=1}^{N} X_k \text{ c'est-à-dire } \forall\, \omega \in \Omega : S(\omega) = \begin{cases} \sum\limits_{k=1}^{N(\omega)} X_k(\omega) & \text{si } N(\omega) \geqslant 1 \\ 0 & \text{sinon} \end{cases}$$

Montrer que, pour toute fonction g telle que les espérances existent, on a :

$$E\big[Sg(S)\big] = \lambda E\big[X_0\, g(S + X_0)\big]$$

Solution :

Toutes les espérances sont supposées exister, donc les séries écrites sont toutes supposées absolument convergentes et les calculs effectués sont légitimes.

1. a) On a donc :

$$E(Ng(N)) = \sum_{n=0}^{+\infty} ng(n)\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \lambda \sum_{n=1}^{+\infty} g(n)\mathrm{e}^{-\lambda}\frac{\lambda^{n-1}}{(n-1)!}$$
$$= \lambda \sum_{n=0}^{+\infty} g(n+1)\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \lambda E\big(g(N+1)\big)$$

b) La fonction $g(x) = 1/x$ n'est pas définie en 0, valeur prise par N : on ne peut donc pas appliquer le résultat précédent !

Un calcul direct – et correct – donne :

$$E\Big(\frac{1}{N+1}\Big) = \sum_{n=0}^{+\infty} \frac{1}{n+1}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^{n+1}}{(n+1)!} = \frac{\mathrm{e}^{-\lambda}}{\lambda}\left(\mathrm{e}^{\lambda} - 1\right)$$

Soit

$$E\Big(\frac{1}{N+1}\Big) = \frac{1 - \mathrm{e}^{-\lambda}}{\lambda}$$

2. Soit $p_n = P(T = n)$ pour $n \in \mathbb{N}$. La relation donne :

$$\sum_{n=0}^{+\infty} ng(n)p_n = \lambda \sum_{n=0}^{+\infty} g(n+1)p_n = \lambda \sum_{n=1}^{+\infty} g(n)p_{n-1}$$

Soit $n_0 \in \mathbb{N}$ fixé et $g : \mathbb{N} \to \mathbb{R}$ définie par $g(n_0) = 1$ et $g(n) = 0$ pour $n \neq n_0$. Alors, il ne reste dans les sommes que $n_0 p_{n_0} = \lambda p_{n_0 - 1}$, d'où, par récurrence, $p_n = p_0 \dfrac{\lambda^n}{n!}$, puis $\sum\limits_{n=0}^{+\infty} p_n = 1$ donne $p_0 = \mathrm{e}^{-\lambda}$ et :

$$T \hookrightarrow \mathcal{P}(\lambda).$$

3. On a, en notant $S_n = \sum\limits_{k=1}^{n} X_k$:

$$\begin{aligned}
E\big(Sg(S)\big) &= \sum_{n=1}^{+\infty} E_{(N=n)}\big(Sg(S)\big)P(N=n) = \sum_{n=1}^{+\infty} E\big(S_n g(S_n)\big)\, P(N=n) \\
&= \sum_{n=1}^{+\infty} E\Big(\big(\sum_{k=1}^{n} X_k\big) g(S_n)\Big) P(N=n) \\
&= \sum_{n=1}^{+\infty} \Big(nE\big(X_n g(S_n)\big)\Big)P(N=n) = \sum_{n=1}^{+\infty} E\big(X_n g(S_n)\big)\frac{\mathrm{e}^{-\lambda}\lambda^n}{(n-1)!} \\
&= \lambda \sum_{n=0}^{+\infty} E\big(X_{n+1} g(S_{n+1})\big)\frac{\mathrm{e}^{-\lambda}\lambda^n}{n!}
\end{aligned}$$

par symétrie et en décalant les indices. Ainsi :

$$\begin{aligned}
E(X_{n+1}g(S_{n+1})) &= \sum_{k=0}^{+\infty} kg(S_n + k)P(X_{n+1} = k) \\
&= \sum_{k=0}^{+\infty} kg(S_n + k)P(X_0 = k) = E(X_0 g(S_n + X_0))
\end{aligned}$$

et :

$$\begin{aligned}
E(Sg(S)) &= \lambda \sum_{n=0}^{+\infty} E(X_0 g(S_n + X_0))\frac{\mathrm{e}^{-\lambda}\lambda^n}{n!} \\
&= \lambda \sum_{n=0}^{+\infty} E_{(N=n)}(X_0 g(S + X_0))P(N=n)
\end{aligned}$$

$$E(Sg(S)) = \lambda E(X_0 g(S + X_0))$$

Exercice 3.17.

Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ suivant la loi normale centrée réduite.

1. Déterminer, pour tout entier n de $\mathbb{N}$, le moment $E(X^{2n})$.

2. a) Déterminer la loi de X^2.

b) Soit m un entier supérieur ou égal à 2, et $(X_1, \ldots, X_m)$, un m-échantillon i.i.d de la loi de X. Déterminer la loi de $Y = \sum\limits_{i=1}^{m} X_i^2$.

c) En déduire le moment $E(Y^n)$, pour $n \in \mathbb{N}^*$.

3. Soit $a_1, \ldots, a_m$ des réels. On admet que pour tout $n \geqslant 1$:

$$(a_1^2 + a_2^2 + \cdots + a_m^2)^n = \sum_{\substack{i_1, i_2, \ldots i_m \in \mathbb{N} \\ i_1 + i_2 + \cdots + i_m = n}} \frac{n!}{i_1! i_2! \ldots i_m!} a_1^{2i_1} a_2^{2i_2} \cdots a_m^{2i_m}$$

En déduire que :

$$E(Y^n) = \sum_{\substack{i_1, i_2, \ldots i_m \in \mathbb{N} \\ i_1 + i_2 + \cdots + i_m = n}} \frac{n!}{2^n} \binom{2i_1}{i_1}\binom{2i_2}{i_2} \cdots \binom{2i_1}{i_m}$$

4. Montrer que :

$$\sum_{\substack{i_1, i_2, \ldots i_m \in \mathbb{N} \\ i_1 + i_2 + \cdots + i_m = n}} \frac{n!}{2^n} \binom{2i_1}{i_1}\binom{2i_2}{i_2} \cdots \binom{2i_1}{i_m} = \frac{2^n \Gamma(m/2 + n)}{\Gamma(m/2)}$$

Solution :

Nous noterons comme d'habitude Φ la fonction de répartition de X et φ la densité obtenue par dérivation de Φ sur $\mathbb{R}$, donc telle que :

$$\forall x \in \mathbb{R}, \varphi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$$

1. Sous réserve de convergence, on a :

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

La fonction à intégrer est continue et comme $\lim_{x \to +\infty} x^2.x^{2n}\mathrm{e}^{-x^2/2} = 0$, la convergence résulte de la règle de Riemann.

Pour $n \geqslant 1$, on écrit : $E(X^{2n}) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} x^{2n-1} \times x\, \mathrm{e}^{-x^2/2} dx$.

On effectue alors l'intégration par parties ainsi préparée :

$$\begin{cases} u(x) = x^{2n-1} \implies u'(x) = (2n-1)x^{2n-2} \\ v'(x) = x\, \mathrm{e}^{-x^2/2} \impliedby v(x) = -\mathrm{e}^{-x^2/2} \end{cases}$$

Comme $u(0)v(0) = 0$ et $\lim_{x \to +\infty} u(x)v(x) = 0$, on peut réaliser cette intégration par parties directement avec les bornes données et :

$$E(X^{2n}) = (2n-1) \times \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} x^{2n-2}\, \mathrm{e}^{-x^2/2} dx = (2n-1)E(X^{2n-2})$$

Comme $E(X^0) = 1$, il vient, par le principe de récurrence :

$$E(X^{2n}) = (2n-1)(2n-3) \cdots 1 = \frac{(2n)!}{2^n n!}$$

Cette dernière formule étant valable aussi pour $n = 0$.

2. a) La variable aléatoire X^2 est à valeurs positives. Pour tout $x \geqslant 0$, on a :

$$P(X^2 \leqslant x) = P(-\sqrt{x} \leqslant X \leqslant \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}) = 2\Phi(\sqrt{x}) - 1$$

Par dérivation, une densité de X^2 est donc :

$$f_{X^2}(x) = \begin{cases} \frac{1}{\sqrt{x}} \varphi(\sqrt{x}) = \frac{1}{\sqrt{2\pi x}} \mathrm{e}^{-x/2} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

Ce qui prouve que X^2 suit la loi $\Gamma(2, 1/2)$.

b) Par stabilité de la loi Gamma, Y suit la loi $\Gamma(2, m/2)$, de densité :

$$f_Y(x) = \begin{cases} \frac{1}{\Gamma(m/2)2^{m/2}} \times y^{\frac{m}{2}-1} \mathrm{e}^{-\frac{y}{2}} & \text{si } y \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$

c) Donc, en utilisant encore le théorème du transfert (la convergence est banale)

$$E(Y^n) = \frac{1}{2^{m/2}\Gamma(m/2)} \int_0^{+\infty} y^{n+\frac{m}{2}-1} \mathrm{e}^{-y/2} dy$$

Le changement de variable $y = 2x$, donne alors :

$$\begin{aligned} E(Y^n) &= \frac{1}{2^{m/2}\Gamma(m/2)} \int_0^{+\infty} 2^{n+\frac{m}{2}-1} x^{n+\frac{m}{2}-1} \mathrm{e}^{-x} 2dx \\ &= \frac{2^n}{\Gamma(m/2)} \int_0^{+\infty} x^{n+\frac{m}{2}-1} \mathrm{e}^{-x} dx. \end{aligned}$$

$$E(Y^n) = \frac{2^n \Gamma(\frac{m}{2} + n)}{\Gamma(\frac{m}{2})}$$

3. Par le résultat admis, on a donc en substituant les X_k aux a_k :

$$E(Y^n) = \sum_{\substack{i_1,i_2,\dots i_m\in\mathbb{N}\\ i_1+i_2+\cdots+i_m=n}} \frac{n!}{i_1!i_2!\dots i_m!}E(X_1^{2i_1})\cdots E(X_m^{2i_m})$$

et par la première question

$$\begin{aligned} E(Y^n) &= \sum_{\substack{i_1,i_2,\dots i_m\in\mathbb{N}\\ i_1+i_2+\cdots+i_m=n}} \frac{n!}{i_1!i_2!\dots i_m!}\times\frac{(2i_1)!}{2^{i_1}i_1!}\times\cdots\times\frac{(2i_m)!}{2^{i_m}i_m!} \\ &= \frac{n!}{2^n}\sum_{\substack{i_1,i_2,\dots i_m\in\mathbb{N}\\ i_1+i_2+\cdots+i_m=n}} \binom{2i_1}{i_1}\cdots\binom{2i_m}{i_m} \end{aligned}$$

et donc :

$$\frac{n!}{2^n}\sum_{\substack{i_1,i_2,\dots i_m\in\mathbb{N}\\ i_1+i_2+\cdots+i_m=n}} \binom{2i_1}{i_1}\binom{2i_2}{i_2}\cdots\binom{2i_1}{i_m} = \frac{2^n\Gamma(m/2+n)}{\Gamma(m/2)}$$

Exercice 3.18.

1. Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$ à valeurs dans $\mathbb{N}$.
   a) Montrer que pour tout x de $[-1,1]$, la série $\sum P(X=n)x^n$ converge. On note $G_X(x)$ sa somme.
   b) Montrer que $G_X(x)=E(x^X)$.

On admet que si une série $\sum\limits_{n\geqslant 0} a_nx^n$ est telle que pour tout x de $[-1,1]$, on a : $G_X(x)=\sum\limits_{n=0}^{\infty}a_nx^n$, alors pour tout n de $\mathbb{N}$, $a_n=P(X=n)$.

2. Soit (X,Y) un couple de variables aléatoires définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$ à valeurs dans $\mathbb{N}$.
   a) Dire pourquoi pour tout (x,y) de $[-1,1]^2$, la série
$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}P(X=n\cap Y=m)x^ny^m$$
converge. On note $G_{(X,Y)}(x,y)$ sa somme.
   b) Montrer que $G_{(X,Y)}(x,y)=E(x^Xy^Y)$.

On suppose désormais que pour tout (x,y) de $[-1,1]^2$,
$$G_{(X,Y)}(x,y)=\frac{py}{\ln(1-p)}\times\frac{\ln(1-pxy)}{1-(1-p)y},$$
où p est un réel de $]0,1[$.

3. a) Déterminer $G_X(x)$ pour tout x de $[-1,1]$.
   b) En utilisant une formule de Taylor, montrer que pour tout x de $[-1,1]$: $\ln(1-px)=-\sum\limits_{n=1}^{\infty}\frac{p^n}{n}x^n$.

(On pourra étudier les variations de la fonction $\psi : x\mapsto\frac{u-x}{1-x}$ sur l'intervalle $[0,u]$ ou $[u,0]$, lorsque $|u|<1$)

   c) En déduire la loi marginale de X.

4. a) Montrer que les variables aléatoires X et $Y-X$ vérifient, pour tout (x,y) de $[-1,1]^2$:
$$G_{(X,Y-X)}(x,y)=G_X(x)G_{Y-X}(y).$$
   b) Déterminer la loi de $Y-X$.

Solution :

1. a) Pour tout n de $\mathbb{N}$, pour tout x de $[-1,1]$: $|P(X=n)||x^n|\leqslant P(X=n)$ et $\sum\limits_{n=0}^{\infty}P(X=n)=1$. Par la règle de majoration, la série proposéee est absolument convergente et la fonction G_X est donc bien définie sur $[-1,1]$.
   b) Par définition de l'espérance et par application du théorème de transfert :
$$G_X(x)=E(x^X).$$

2. Pour tout (n,m) de $\mathbb{N}^2$, pour tout (x,y) de $[-1,1]^2$, on a :

$$|P(X=n\cap Y=m)x^ny^m| \leqslant P(X=n\cap Y=m)$$

Or : $\sum\limits_{n=0}^{\infty}\sum\limits_{m=0}^{\infty} P(X=n\cap Y=m) = P(\Omega) = 1$.

Par application du théorème de Fubini, la fonction $G_{(X,Y)}$ est donc bien définie sur $[-1,1]^2$.

b) Par définition de l'espérance et théorème du transfert, on a :

$$G_{(X,Y)}(x,y) = E(x^Xy^Y).$$

3. a) On a $G_X(x) = E(x^X) = E(x^X1^Y) = G_{(X,Y)}(x,1) = \dfrac{\ln(1-px)}{\ln(1-p)}$.

b) La formule de Taylor avec reste intégral pour $u \mapsto \ln(1-u)$, avec $-1 < u < 1$, donne, pour tout n de $\mathbb{N}$:

$$\ln(1-u) = -\sum_{k=1}^{n}\frac{u^k}{k} + \int_0^u \frac{(x-u)^n}{(1-x)^{n+1}}dx$$

la fonction $\psi : x \to \dfrac{u-x}{1-x}$ est monotone telle que $\psi(0) = u$ et $\psi(u) = 0$. Ainsi, en distinguant les deux cas $u \in \,]-1,0[$ et $u \in [0,1[$, on a, $[0,u]$ désignant le segment d'extrémités 0 et u, même pour u négatif :

$$\forall\, x \in [0,u], \left|\frac{x-u}{1-x}\right| \leqslant |u|$$

Par conséquent, pour u tel que $-1 < u < 1$:

$$\left|\int_0^u \frac{(x-u)^n}{(1-x)^{n+1}}dx\right| \leqslant \left|\int_0^u \left|\frac{x-u}{1-x}\right|^n \frac{dx}{1-x}\right| \leqslant |u|^n\left|\int_0^u \frac{dx}{1-x}\right| \underset{n\to\infty}{\longrightarrow} 0$$

En revenant à la formule de Taylor, on peut donc passer à la limite lorsque n tend vers l'infini, et :

$$\forall\, u \in \,]-1,1[, \ln(1-u) = -\sum_{k=1}^{\infty}\frac{u^k}{k}$$

En particulier :

$$\forall\, x \in [0,1], \ln(1-px) = -\sum_{n=1}^{\infty}\frac{p^n}{n}\,x^n$$

c) Ainsi, pour $x \in [-1,1]$:

$$G_X(x) = \frac{1}{\ln(1-p)}\ln(1-px) = \sum_{n=1}^{\infty} -\frac{p^n}{n\ln(1-p)}\,x^n$$

Ainsi :

$$P(X=0) = 0, \forall\, n \geqslant 1, P(X=n) = -\frac{p^n}{n\ln(1-p)}$$

4. a) On a :

$$\begin{aligned} G_{(X,Y-X)}(x,y) &= E(x^Xy^{Y-X}) = E((\tfrac{x}{y})^Xy^Y) = G_{(X,Y)}(\tfrac{x}{y},y) \\ &= \frac{py}{1-(1-p)y}\times\frac{\ln(1-px)}{\ln(1-p)} \end{aligned}$$

Or on a vu que $G_X(x) = \dfrac{1}{\ln(1-p)}\ln(1-px)$ et de la même façon :

$$\begin{aligned} G_{Y-X}(y) &= E(y^{Y-X}) = E(1^Xy^{Y-X}) = G_{(X,Y-X)}(1,y) \\ &= \frac{py}{1-(1-p)y}\times\frac{\ln(1-p)}{\ln(1-p)} = \frac{py}{1-(1-p)y} \end{aligned}$$

Donc

$$G_{(X,Y-X)}(x,y) = G_X(x)G_{Y-X}(y).$$

b) Par le résultat admis, la loi de $Y-X$ se déduit de la connaissance de G_{Y-X} :

Comme $|(1-p)y| < 1$, on peut écrire, grâce à l'identité geométrique :

$$\begin{aligned} G_{Y-X}(y) &= \frac{py}{1-(1-p)y} = py\sum_{n=0}^{\infty}\big((1-p)y\big)^n = \sum_{n=0}^{\infty} p(1-p)^ny^{n+1} \\ &= \sum_{n=1}^{\infty} p(1-p)^{n-1}y^n \end{aligned}$$

Donc : $P(Y-X=0) = 0$ et $\forall\, n \geqslant 1, P(Y-X=n) = p(1-p)^{n-1}$:

$$Y-X \hookrightarrow \mathcal{G}(p)$$