3

PROBABILITÉS

Exercice 3.01.

Soit $\lambda > 0$ un réel fixé. On considère une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}^*}$ définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$. On suppose que les variables X_n sont mutuellement indépendantes et suivent chacune une loi exponentielle de paramètre λ. Soient $p \in \,]0,1[$ et $q = 1-p$.

1. Soit $n \in \mathbb{N}^*$. Montrer que $X_1 + X_2 + \cdots + X_n$ est une variable aléatoire à densité dont on donnera une densité.

2. Montrer que :

$$\forall\, x \in \mathbb{R}, P(X_1 + X_2 + \cdots + X_n \leqslant x) = \mathbf{1}_{\mathbb{R}^+}(x)\,(1 - \mathrm{e}^{-\lambda x}\sum_{k=0}^{n-1}\frac{(\lambda x)^k}{k!})$$

où $\mathbf{1}_{\mathbb{R}^+}$ est la fonction indicatrice de $\mathbb{R}^+$, *i.e.* la fonction définie par :

$$\mathbf{1}_{\mathbb{R}^+} : x \mapsto \begin{cases} 1 & \text{si } x \in \mathbb{R}^+ \\ 0 & \text{sinon.} \end{cases}$$

3. Soit N une variable aléatoire réelle suivant la loi géométrique de paramètre p.

On pose $S = \sum_{n=1}^{N} X_n$, *i.e.* pour tout $\omega \in \Omega$ $S(\omega) = \sum_{n=1}^{N(\omega)} X_n(\omega)$.

a) Soit F_S la fonction de répartition de S. Montrer que

$$\forall\, x \in \mathbb{R}, F_S(x) = \mathbf{1}_{\mathbb{R}^+}(x)(1 - \mathrm{e}^{-\lambda x}\sum_{n=1}^{+\infty}\sum_{k=0}^{n-1}\frac{(\lambda x)^k}{k!}pq^{n-1}).$$

b) En déduire la loi de S (on admettra qu'il est possible de permuter l'ordre des sommations dans la formule précédente).

c) Exprimer l'espérance $E(S)$ de S en fonction de $E(X_1)$ et de $E(N)$.

Solution :

1. Chaque variable X_i suit la loi exponentielle de paramètre λ, autrement dit la loi Gamma $\Gamma(\lambda, 1)$. Comme les variables X_i sont indépendantes, on sait, d'après le théorème de stabilité des lois Gamma, que $X_1 + X_2 + \ldots + X_n$ suit la loi $\Gamma(\lambda, n)$. Une densité f_n de $X_1 + X_2 + \ldots + X_n$ est alors donnée par :

$f_n(x) = 0$ si $x < 0$ et $f_n(x) = \dfrac{\lambda^n x^{n-1} \mathrm{e}^{-\lambda x}}{\Gamma(n)} = \dfrac{\lambda^n x^{n-1} \mathrm{e}^{-\lambda x}}{(n-1)!}$ si $x \geqslant 0$.

2. On suppose d'abord que $x < 0$. On a alors $P(X_1 + X_2 + \ldots + X_n \leqslant x < 0) = 0$.
Considérons ensuite un réel $x \geqslant 0$. On a alors
$P(X_1 + X_2 + \ldots + X_n \leqslant x) = \displaystyle\int_0^x f_n(t)dt = \int_0^x \frac{\lambda}{(n-1)!} (\lambda x)^{n-1} \mathrm{e}^{-\lambda x} dx.$

On applique la formule de Taylor avec reste intégral à l'ordre $n-1$ à la fonction exponentielle entre 0 et λx.

On trouve $\mathrm{e}^{\lambda x} = \displaystyle\sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} + \int_0^{\lambda x} \frac{(\lambda x - t)^{n-1}}{(n-1)!} \mathrm{e}^t dt.$

On effectue dans l'intégrale le changement de variable $u = x - \dfrac{t}{\lambda}$. On obtient :

$$\mathrm{e}^{\lambda x} - \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} = \mathrm{e}^{\lambda x} \int_0^x \frac{\lambda}{(n-1)!} (\lambda u)^{n-1} \mathrm{e}^{-\lambda u} dt.$$

On en déduit que $P(X_1 + X_2 + \ldots + X_n \leqslant x) = 1 - \mathrm{e}^{-\lambda x} \displaystyle\sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}$. Ceci donne la formule attendue.

3. a) Comme S est à valeurs dans $\mathbb{R}^+$, pour tout $x < 0$, $F_S(x) = P(S \leqslant x) = 0$.

Considérons à présent $x \geqslant 0$. En appliquant la formule des probabilités totales au système complet d'événements $(N = n)_{n \in \mathbb{N}^*}$, on trouve :

$$\begin{aligned} F_S(x) = P(S \leqslant x) &= \sum_{n=1}^{+\infty} P_{[N=n]}(S \leqslant x) P(N = n) \\ &= \sum_{n=1}^{+\infty} P(X_1 + X_2 + \ldots + X_n \leqslant x) pq^{n-1} \\ &= \sum_{n=1}^{+\infty} (1 - \mathrm{e}^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}) pq^{n-1} \end{aligned}$$

$$= \sum_{n=1}^{+\infty} pq^{n-1} - \mathrm{e}^{-\lambda x} \sum_{n=1}^{+\infty} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} pq^{n-1}.$$

En remarquant que $\sum_{n=1}^{+\infty} pq^{n-1} = 1$, on obtient la formule souhaitée.

b) On applique à la série double le théorème de permutation :

$$\sum_{n=1}^{+\infty} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} pq^{n-1} = \sum_{k=0}^{+\infty} \frac{(\lambda x)^k}{k!} \sum_{n=k+1}^{+\infty} pq^{n-1} = \sum_{k=0}^{+\infty} \frac{(\lambda x)^k}{k!} p \frac{q^k}{1-q}$$

$$= \sum_{k=0}^{+\infty} \frac{(\lambda x q)^k}{k!} = \mathrm{e}^{\lambda x q}.$$

Ainsi, pour tout $x \in \mathbb{R}$, $F_S(x) = \mathbf{1}_{\mathbb{R}^+}(x)(1 - \mathrm{e}^{-\lambda x}\mathrm{e}^{\lambda x q}) = \mathbf{1}_{\mathbb{R}^+}(x)(1 - \mathrm{e}^{-\lambda x p})$.

La variable S suit donc la loi exponentielle de paramètre λp.

c) On en déduit que $E(S) = \dfrac{1}{\lambda p} = \dfrac{1}{\lambda} \times \dfrac{1}{p} = E(X_1)E(N)$.

Exercice 3.02.

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

On admet que les formules de calcul de la covariance d'un couple de variables à densités sont identiques à celles du calcul de la covariance d'un couple de variables discrètes.

Soient $X_1, \ldots, X_n$, n variables aléatoires. On dit que le vecteur $(X_1, \ldots, X_n)$ est *gaussien* si pour tout $(a_1, \ldots, a_n)$ réels, la variable aléatoire $a_1X_1 + a_2X_2 + \cdots + a_nX_n$ suit une loi normale (on considérera que la variable nulle suit une loi normale).

1. a) Soit $(X_1, \ldots, X_n)$ un vecteur gaussien. Montrer que chaque variable X_i suit une loi normale.

b) Montrer que si $X_1 \ldots, X_n$ sont des variables aléatoires indépendantes suivant des lois normales, alors $(X_1, \ldots, X_n)$ est un vecteur gaussien.

2. Dans cette question X suit la loi normale centrée réduite, et Y indépendante de X, suit la loi discrète définie par $P(Y = 1) = P(Y = -1) = \dfrac{1}{2}$.

a) Déterminer la loi de XY.

b) Le vecteur (X, XY) est-il gaussien ?

3. Dans cette question (X, Y, Z) est un vecteur gaussien de matrice de variance/covariance

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 6 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

et de vecteur espérance $E\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

a) Déterminer la loi de $U = X + Y$ et $V = X + Z$.

b) Écrire la matrice de variance/covariance de (U, V).

c) Déterminer une matrice $P \in \mathcal{M}_2(\mathbb{R})$ telle que les variables aléatoires U' et V' définies par $\begin{pmatrix} U' \\ V' \end{pmatrix} = P\begin{pmatrix} U \\ V \end{pmatrix}$ soient non corrélées.

Solution :

1. a) Il suffit de prendre $a_j = 0$ pour $j \neq i$ et $a_i = 1$.

b) C'est un théorème du cours sur la stabilité des lois normales.

2. a) Calculons :

$$\begin{aligned} P(XY \leqslant x) &= P((XY \leqslant X) \cap [Y = 1]) + P((XY \leqslant X) \cap [Y = -1]) \\ &= \frac{1}{2}P(X \leqslant x) + \frac{1}{2}P(X \geqslant -x) = P(X \leqslant x) = \Phi(x) \end{aligned}$$

Ainsi XY suit la loi normale centrée réduite.

b) En revanche le vecteur (X, XY) n'est pas gaussien.

En effet $X + XY = (1 + Y)X$ vérifie :

$P(X + XY = 0) = P(X(1 + Y) = 0) = P(Y = -1) = \frac{1}{2}$ (l'événement $X = 0$ est quasi-impossible), donc $X+XY$ n'a pas une fonction de répartition continue !

3. Grâce au vecteur espérance et à la matrice de variance-covariance, il vient :

a) $E(X + Y) = 1, V(X + Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X, Y) = 5$, donc $X + Y \hookrightarrow \mathcal{N}(1, 5)$

$E(X + Z) = 2, V(X + Z) = V(X) + V(Z) + 2\,\mathrm{Cov}(X, Z) = 2$, donc $X + Z \hookrightarrow \mathcal{N}(2, 2)$

b) De plus :

$\mathrm{Cov}(X + Y, X + Z) = E(X^2 + XY + XZ + YZ) - E(X + Y)E(X + Z) = 2$

Le couple $(X+Y, X+Z)$ est gaussien car le triplet (X, Y, Z) l'est. La matrice de variance/covariance de $(X + Y, X + Z)$ est $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$.

c) Cette matrice symétrique réelle admet 1 et 6 comme valeurs propres, une base orthonormée de vecteurs propres étant formée des vecteurs $\frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ et $\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Avec $P = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, il vient $A = PD{}^tP$ avec $D = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$.

La matrice de variance-covariance de $P\begin{pmatrix} X+Y \\ X+Z \end{pmatrix}$ est la matrice diagonale D.

Exercice 3.03.

Dans tout le problème, N désigne un entier naturel non nul.

Soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, et suivant toutes la loi uniforme sur $[\![1, N]\!]$.
On note T_n et Z_n les deux variables aléatoires définies, pour tout entier naturel n supérieur ou égal à 1, par :

$$T_n = \sup(U_1, U_2, \ldots, U_n) \text{ et } Z_n = \inf(U_1, U_2, \ldots, U_n)$$

On pose $S_n = T_n + Z_n - 1$.

On pose enfin, pour tout entier naturel n non nul, $a_n(N) = \begin{cases} \sum\limits_{k=1}^{N-1} \left(\frac{k}{N}\right)^n & \text{si } N \geqslant 2 \\ 0 & \text{si } N = 1 \end{cases}$.

1. Soit Y une variable aléatoire à valeurs dans $[\![1, N]\!]$. Établir la relation suivante :

$$E(Y) = \sum_{k=0}^{N-1} P([Y > k])$$

Dans la suite de l'exercice, on cherche à estimer le paramètre N.

2. a) Calculer, pour tout entier naturel k appartenant à $[\![1, N]\!]$, $P([T_n \leqslant k])$.

b) En déduire la loi de T_n.

c) Calculer $E(T_n)$ en fonction de N et de $a_n(N)$.

3. a) Calculer, pour tout entier naturel k appartenant à $[\![0, N-1]\!]$, $P([Z_n > k])$.

b) En déduire $E(Z_n)$ en fonction de $a_n(N)$.

4. a) $T_n + a_n(N)$ est-il un estimateur sans biais de N ?

b) Montrer que T_n est un estimateur asymptotiquement sans biais de N.

c) Montrer que S_n est un estimateur sans biais de N.

Solution :

1. On a $E(Y) = \sum\limits_{k=1}^{N} kP([Y = k]) = \sum\limits_{k=1}^{N} k\big[P([Y > k-1]) - P([Y > k])\big]$

$$= \sum_{k=1}^{N} kP([Y > k-1]) - \sum_{k=1}^{N} kP([Y > k]).$$

En effectuant un glissement d'indice dans la première somme :

$$E(Y) = \sum_{k=0}^{N-1} (k+1)P([Y > k]) - \sum_{k=1}^{N} kP([Y > k])$$

En rajoutant dans la seconde somme le terme correspondant à $k = 0$ qui est nul :

$$E(Y) = \sum_{k=0}^{N-1} (k+1)P([Y > k]) - \sum_{k=0}^{N} kP([Y > k])$$

En rassemblant dans une seule somme les termes communs :

$$E(Y) = \sum_{k=0}^{N-1} P([Y > k]) - NP([Y > N])$$

Comme $P([Y > N]) = 0$, il reste finalement : $E(Y) = \sum\limits_{k=0}^{N-1} P([Y > k])$.

2. a) On a : $[T_n \leqslant k] = [X_1 \leqslant k] \cap \cdots \cap [X_n \leqslant k]$. Comme les variables sont indépendantes, on obtient : $\forall\, k \in [\![1, N]\!], \quad P([T_n \leqslant k]) = \Big(\dfrac{k}{N}\Big)^n$

b) On a, pour tout k de $[\![1, N]\!]$, $P([T_n = k]) = P([T_n \leqslant k]) - P([T_n \leqslant k-1])$. Comme la formule de la question précédente est encore valable pour $k = 0$, on obtient finalement :

$\forall\, k \in [\![1, N]\!], \quad P([T_n = k]) = \big(\dfrac{k}{N}\big)^n - \big(\dfrac{k-1}{N}\big)^n$

c) On a donc : $E(T_n) = \sum\limits_{k=1}^{N} k\Big[\big(\dfrac{k}{N}\big)^n - \big(\dfrac{k-1}{N}\big)^n\Big]$.

En séparant en deux sommes :

$$E(T_n) = \sum_{k=1}^{N} k\big(\frac{k}{N}\big)^n - \sum_{k=1}^{N} k\big(\frac{k-1}{N}\big)^n$$

En effectuant un décalage d'indice dans la seconde somme :

$$E(T_n) = \sum_{k=1}^{N} k\big(\frac{k}{N}\big)^n - \sum_{k=0}^{N-1} (k+1)\big(\frac{k}{N}\big)^n$$

On peut rajouter dans la première somme le terme correspondant à $k = 0$ qui est nul :

$$E(T_n) = \sum_{k=0}^{N} k\big(\frac{k}{N}\big)^n - \sum_{k=0}^{N-1} (k+1)\big(\frac{k}{N}\big)^n$$

En rassemblant les termes communs : $E(T_n) = N - \sum\limits_{k=0}^{N-1} \big(\dfrac{k}{N}\big)^n$. Finalement :

$$E(T_n) = N - a_n(N)$$

3. a) On a $P([Z_n > k]) = P([X_1 > k] \cap \ldots \cap [X_n > k])$.

Comme les variables X_i sont indépendantes, on a : $P([Z_n > k]) = \big(\dfrac{N-k}{N}\big)^n$

b) En utilisant la question préliminaire : $E(Z_n) = \sum\limits_{k=0}^{N-1} \big(\frac{N-k}{N}\big)^n$
En effectuant le changement d'indice $j = N - k$:
$$E(Z_n) = \sum_{j=1}^{N} \big(\frac{j}{N}\big)^n$$
Finalement : $E(Z_n) = 1 + a_n(N)$.

4. a) On a, par linéarité de l'opérateur espérance, $E(T_n + a_n(N)) = N$ mais comme $T_n + a_n(N)$ dépend de N, ce n'est pas un estimateur de N.

b) On a $a_n(N) = \sum\limits_{k=1}^{N-1} \big(\frac{k}{N}\big)^n$. Comme $0 \leqslant \frac{k}{N} < 1$ et comme la limite de la somme est égale à la somme des limites, on a $\lim\limits_{n\to+\infty} d_n(N) = 0$, d'où : $\lim\limits_{n\to+\infty} E(T_n) = N$.
Conclusion : T_n est un estimateur asymptotiquemnt sans biais de N.

c) On a, par linéarité, $E(S_n) = E(T_n) + E(Z_n) - 1$.
En utilisant les résultats précédents, on a donc :
$$E(S_n) = N$$

Exercice 3.04.

On considère un arrêt d'une ligne de bus. Le passage des bus à cet arrêt est prévu à chaque heure juste : 12h00, 13h00, 14h00 etc. Mais la circulation est telle que que les bus peuvent avoir du retard.

Dans cet exercice, on mesure le temps à partir de midi, considéré comme l'instant 0, et l'unité de temps est l'heure.
Le retard du $k^{\text{ème}}$ bus est une variable aléatoire Z_k. On suppose les variables Z_k indépendantes de même loi, à valeurs dans $[0, 1]$. La loi commune des Z_k a pour densité f, pour fonction de répartition F, pour espérance μ et pour variance σ^2.

1. Un voyageur arrive à l'instant $x \in [0, 1[$. On note T_x la variable aléatoire du temps d'attente de ce voyageur pour l'arrivée d'un bus.

a) Montrer que
$$P(T_x \leqslant t) = \begin{cases} F(t+x) - F(x) & \text{si } 0 \leqslant t < 1-x \\ 1 - F(x) + F(x)\,F(t+x-1) & \text{si } 1-x \leq t < 2-x \\ 1 & \text{si } 2-x \leq t \end{cases}$$

b) Préciser une densité de T_x.

2. Montrer que :
$$E(T_x) = \int_0^{1-x} tf(t+x)\,dt + (\mu + 1 - x)\,F(x)$$

$$= 1 - x - \int_x^1 F(u)du + (\mu + 1 - x)\, F(x)$$

3. a) Montrer que $\displaystyle\int_0^1 F(z)\, dz = 1 - \mu$ et $\displaystyle\int_0^1 z\, F(z)\, dz = \frac{1 - \mu^2 - \sigma^2}{2}$.

b) On suppose que le voyageur arrive au hasard entre midi et 13h00, c'est-à-dire que son instant d'arrivée est une variable aléatoire X qui suit la loi uniforme sur le segment $[0, 1]$.

On admet que son temps d'attente W vérifie $E(W) = \displaystyle\int_0^1 E(T_x)\, dx$.

Calculer $E(W)$. Quand cette espérance est-elle minimale ?

Solution :

1. a) On a $T_x(\Omega) = \mathbb{R}_+$.

• Si $0 \leqslant t < 1 - x$, *i.e.* $x \leqslant x + t < 1$, on a $[x, x + t[\ \subset [0, 1[$ et on est sûr d'avoir le bus de midi ; la probabilité d'attendre au maximum t est celle que le bus de midi arrive dans l'intervalle de temps $[x, x + t]$, soit

$$P(T_x \leqslant t) = F(x + t) - F(x)$$

• si $1 - x \leqslant t < 2 - x$, *i.e.* $1 \leqslant x + t < 2$, on a deux cas disjoints :

$\rightarrow$ soit le bus de midi n'est pas encore passé, ce qui a pour probabilité $1 - F(x)$;

$\rightarrow$ soit il est déjà passé (avec la probabilité $F(x)$), et la probabilité d'attendre au plus t est alors la probabilité que le bus de 13h00 arrive entre 0 et $x + t - 1$, soit $F(x + t - 1)$; alors

$$P(T_x \leqslant t) = 1 - F(x) + F(x)\, F(x + t - 1)$$

• si $2 - x \leqslant t$, c'est-à-dire $x + t \geqslant 2$, on est sûr d'avoir au pire le bus de 13h00, donc

$$P(T_x \leqslant t) = 1$$

b) Par dérivation, T_x a pour densité (sur $\mathbb{R}_+$)

$$\varphi(t) = \begin{cases} f(x + t) & \text{si } 0 \leqslant t < 1 - x \\ F(x)\, f(x + t - 1) & \text{si } 1 - x \leqslant t < 2 - x \\ 0 & \text{si } 2 - x \leqslant t \end{cases}$$

2. Les changements de variable $u = x + t$, $u = x + t - 1$ et une intégration par parties donnent :

$$E(T_x) = \int_0^{1-x} t\, f(x + t)dt + \int_{1-x}^{2-x} t\, F(x)\, f(x + t - 1)dt$$

$$= \int_x^1 (u - x)\, f(u)du + F(x) \int_0^1 (u - x + 1)\, f(u)du$$

$$= \int_x^1 u\,f(u)du - x \int_x^1 f(u)du + F(x) \int_0^1 u\,f(u)du + (1-x)\,F(x) \int_0^1 f(u)du$$
$$= \big[u\,F(u)\big]_x^1 - \int_x^1 F(u)du - x\,\big[F(1) - F(x)\big] + F(x)\,E(X) + (1-x)\,F(x)$$
$$= 1 - x\,F(x) - \int_x^1 F(u)du - x\,\big[1 - F(x)\big] + (\mu + 1 - x)\,F(x)$$
$$= 1 - x - \int_x^1 F(u)du + (\mu + 1 - x)\,F(x)$$

3. a) On a par intégration par parties

$\mu = E(Z) = \displaystyle\int_0^1 z\,f(z)dz = \big[z\,F(z)\big]_0^1 - \int_0^1 F(z)dz = 1 - \int_0^1 F(z)dz,$

d'où $\displaystyle\int_0^1 F(z)dz = 1 - \mu$

$\sigma^2 = \displaystyle\int_0^1 (z-\mu)^2 f(z)dz = \big[(z-\mu)^2\,F(z)\big]_0^1 - 2\int_0^1 (z-\mu)\,F(z)dz$

$\sigma^2 = (1-\mu)^2 - 2\displaystyle\int_0^1 z\,F(z)dz + 2\mu\int_0^1 F(z)dz = 1 - \mu^2 - 2\int_0^1 z\,F(z)dz$

soit : $\displaystyle\int_0^1 z\,F(z)dz = \frac{1-\mu^2-\sigma^2}{2}$

b) D'après la question 2, on a :

$$E(W) = \int_0^1 \left[1 - x - \int_x^1 F(u)du + (\mu + 1 - x)\,F(x)\right] dx$$

On calcule séparément :

$\displaystyle\int_0^1 (1-x)dx = 1 - \big[\tfrac{x^2}{2}\big]_0^1 = \frac{1}{2}$

$$\int_0^1 \left(\int_x^1 F(u)du\right) dx = \left[x\int_x^1 F(u)du\right]_0^1 - \textstyle\int_0^1 x\,\big[-F(x)\big]dx$$
$$= \int_0^1 x\,F(x)dx = \frac{1-\mu^2-\sigma^2}{2}$$

$$\int_0^1 \big[(\mu+1-x)\,F(x)\big]dx = (\mu+1)\int_0^1 F(x)dx - \int_0^1 x\,F(x)dx$$
$$= (1+\mu)\,(1-\mu) - \frac{1-\mu^2-\sigma^2}{2} = \frac{1-\mu^2+\sigma^2}{2}$$

d'où finalement $E(W) = \dfrac{1}{2} - \dfrac{1-\mu^2-\sigma^2}{2} + \dfrac{1-\mu^2+\sigma^2}{2} = \dfrac{1}{2} + \sigma^2$

On constate que l'espérance du temps d'attente du voyageur est minimum lorsque la loi des retards des bus est d'écart-type nul, c'est-à-dire lorsque le retard est constant.

Exercice 3.05.

Dans cet exercice X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ suivant la loi $\Gamma(1, n)$ (c'est-à-dire la loi $\gamma(n)$), avec $n \in \mathbb{N}^*$.

1. Rappeler l'expression d'une densité de X, ainsi que la valeur de son espérance $E(X)$ et de sa variance $V(X)$.

2. Pour tout réel $\lambda > 0$, on pose $\psi(\lambda) = \ln\big(E(\mathrm{e}^{-\lambda(X-E(X))})\big)$. Calculer $\psi(\lambda)$.

3. Montrer que pour tout $\lambda > 0$, on a $\psi(\lambda) \leqslant n\dfrac{\lambda^2}{2}$.

4. a) Montrer que pour tout $x > 0$ et tout $\lambda > 0$, on a

$$P(E(X) - X \geqslant x) \leqslant \mathrm{e}^{-\lambda x + \psi(\lambda)}$$

b) En déduire que pour tout $x > 0$ $P(E(X) - X \geqslant x) \leqslant \mathrm{e}^{-x^2/(2n)}$.

5. Comparer l'inégalité que l'on vient d'obtenir avec celle obtenue par l'inégalité de Bienaymé-Tchebicheff.

Solution :

1. Une densité de X est donnée par $f(x) = \begin{cases} \dfrac{x^{n-1}\mathrm{e}^{-x}}{(n-1)!} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$

On sait que $E(X) = n$ et $V(X) = n$.

2. On utilise le théorème de transfert :

$$\begin{aligned} E(\mathrm{e}^{-\lambda(X-E(X))}) &= \int_0^{+\infty} \mathrm{e}^{-\lambda(x-n)} \frac{x^{n-1}}{(n-1)!}\mathrm{e}^{-x}dx \\ &= \mathrm{e}^{\lambda n}\int_0^{+\infty} \frac{x^{n-1}}{(n-1)!}\mathrm{e}^{-(\lambda+1)x}dx \\ &= \frac{\mathrm{e}^{\lambda n}}{(\lambda+1)^n}\int_0^{+\infty} \frac{u^{n-1}}{(n-1)!}\mathrm{e}^{-u}du \text{ (avec } u = (\lambda+1)x) \\ &= \frac{\mathrm{e}^{\lambda n}}{(\lambda+1)^n} \end{aligned}$$

Ainsi :

$$\psi(\lambda) = \lambda n - n\ln(\lambda+1)$$

3. Une étude de la fonction $f : \lambda \mapsto \dfrac{n\lambda^2}{2} - \lambda n + n\ln(\lambda+1)$ montre que f est croissante sur $\mathbb{R}^+$ et que $f(0) = 0$.

4. a) Les événements $[E(X) - X \geqslant x]$ et $[\mathrm{e}^{\lambda(E(X)-X)} \geqslant \mathrm{e}^{\lambda x}]$ sont égaux. L'inégalité de Markov donne alors :

$$P([E(X) - X \geqslant x]) = P[\mathrm{e}^{\lambda(E(X)-X)} \geqslant \mathrm{e}^{\lambda x}]) \leqslant \frac{1}{\mathrm{e}^{\lambda x}}E(\mathrm{e}^{\lambda(E(X)-X)})$$

$$= \mathrm{e}^{\psi(\lambda)}\mathrm{e}^{-\lambda x}$$

b) Ainsi : $\psi(\lambda) - \lambda x \leqslant \dfrac{n\lambda^2}{2} - \lambda x$

Ceci étant vérifié pour tout $\lambda > 0$, il vient :

$$P([E(X) - X \geqslant x]) \leqslant \inf_{\lambda>0} \left(\frac{n\lambda^2}{2} - \lambda x\right) = -\frac{x^2}{2n}$$

5. L'inégalité de Bienaymé Tchebicheff permet d'écrire

$$P([E(X) - X \geqslant x]) \leqslant P([|E(X) - X| \geqslant x]) \leqslant \frac{V(X)}{x^2} = \frac{n}{x^2}$$

Or, on peut montrer que pour tout $x > 0$, $\mathrm{e}^{-x^2/(2n)} \leqslant \dfrac{n}{x^2}$ (faire par exemple une étude de fonction). Ainsi l'inégalité obtenue (grandes déviations) est-elle meilleure que celle obtenue par Bienaymé Tchebicheff.

Exercice 3.06.

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

1. Soit Z une variable aléatoire à valeurs dans $\mathbb{N}$. Montrer que la variable aléatoire 2^{-Z} admet une espérance. On la note $r(Z)$.

On suppose dans la suite de l'exercice que pour tout $n \in \mathbb{N}$, $P(Z = n) = \left(\frac{1}{2}\right)^{n+1}$.

2. a) Montrer que l'on définit ainsi une loi de probabilité et calculer $r(Z)$.

b) Montrer que pour tout $(n, q) \in \mathbb{N}^2$, $\sum\limits_{k=0}^{n} \dbinom{k+q}{q} = \dbinom{n+q+1}{q+1}$.

c) Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de même loi que Z et pour tout entier $q \geqslant 1$, on pose $S_q = \sum\limits_{i=1}^{q} X_i$.

Montrer que la loi de S_q est définie par :

$$\forall\, n \in \mathbb{N}, P(S_q = n) = \binom{n+q-1}{q-1}\left(\frac{1}{2}\right)^{n+q}$$

d) Calculer $r(S_q)$. En déduire que

$$\sum_{n=0}^{\infty} \binom{n+q-1}{q-1}\left(\frac{1}{4}\right)^{n} = \left(\frac{4}{3}\right)^{q}$$

3. On suppose dans cette question que Z représente le nombre de lionceaux devant naître en 2014 d'un couple de lions. Chaque lionceau a la probabilité 1/2 d'être mâle ou femelle, indépendamment des autres. On note F la variable aléatoire représentant le nombre de femelles devant naître en 2014.

Déterminer la loi de F.

Solution :

1. Comme $0 \leqslant P(Z=n)\left(\frac{1}{2}\right)^n \leqslant P(Z=n)$ et que $\sum P(Z=n)$ converge. Le théorème de transfert permet d'affirmer que $E(2^{-Z})$ existe (et est majorée par 1).

2. a) Ici $r(Z)=\sum\limits_{n=0}^{\infty}\left(\frac{1}{2}\right)^{2n+1}=\frac{1}{2}\sum\limits_{n=0}^{\infty}\left(\frac{1}{4}\right)^n=\frac{2}{3}$

b) La relation de Pascal :
$$\binom{n+q+1}{q+1}+\binom{n+q+1}{q}=\binom{n+q+2}{q+1}$$
donne le résultat à l'aide d'un raisonnement par récurrence banal.

c) Pour $q=1$, on a $S_1=Z$: la relation est donc vérifiée pour $q=1$.

Supposons que $\forall\, n\in\mathbb{N},\ P(S_q=n)=\binom{n+q-1}{q-1}\left(\frac{1}{2}\right)^{n+q}$.

Comme $S_{q+1}=S_q+X_{q+1}$, par indépendance :
$P(S_{q+1}=n)=\sum\limits_{k=0}^{n}P(S_q=k)P(X_{q+1}=n-k)=\left(\frac{1}{2}\right)^{n+q+1}\sum\limits_{k=0}^{n}\binom{k+q-1}{q-1}$

Soit : $P(S_{q+1}=n)=\left(\frac{1}{2}\right)^{n+q+1}\binom{n+q}{q}$. On conclut par le principe de récurrence.

d) Par indépendance des (X_n), on a indépendance des (2^{-X_n}) et $r(S_q)=r(X_1)^q$, ce qui s'écrit $\sum\limits_{n=0}^{\infty}\binom{n+q-1}{q-1}\left(\frac{1}{2}\right)^{n+q}\left(\frac{1}{2}\right)^n=\left(\frac{2}{3}\right)^q$, soit : $\sum\limits_{n=0}^{\infty}\binom{n+q-1}{q-1}\left(\frac{1}{4}\right)^n=\left(\frac{4}{3}\right)^q$.

3. La famille $([Z=n])_{n\geq 0}$ forme un système complet d'événements.
La loi conditionnelle de F conditionnée par $[Z=n]$ est la loi binomiale $\mathcal{B}(n,1/2)$. Donc, pour tout $k\in\mathbb{N}$:
$$\begin{aligned}P(F=k)&=\sum_{n=0}^{\infty}P_{[Z=n]}(F=k)P(Z=n)=\sum_{n=k}^{\infty}\binom{n}{k}\left(\frac{1}{2}\right)^n\\&=\left(\frac{1}{2}\right)\left(\frac{1}{4}\right)^k\sum_{n=0}^{\infty}\binom{n+k}{k}\left(\frac{1}{4}\right)^n=\frac{2}{3}\times\left(\frac{1}{3}\right)^k.\end{aligned}$$

Exercice 3.07.

Dans cet exercice m désigne un réel strictement positif et X une variable aléatoire suivant la loi exponentielle de paramètre $1/m$.
$(X_k)_{k\geq 1}$ désigne une suite de variables aléatoires *i.i.d.* de même loi que X.
On pose, pour tout entier $n\geqslant 1$, $S_n=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$.

1. Déterminer la loi de nS_n.
Soit f une fonction continue et bornée par une constante K sur $\mathbb{R}^+$.

2. Soit $\varepsilon > 0$. Montrer qu'il existe $\alpha > 0$ tel que si $|x-m| < \alpha$ alors $|f(x) - f(m)| < \varepsilon/2$.

3. Soit Z une variable aléatoire à densité à valeurs dans $\mathbb{R}^+$ admettant une espérance m et une variance σ^2.

a) Montrer que $f(Z)$ admet une espérance.

b) Justifier l'inégalité $|E(f(Z)) - f(m)| \leqslant \frac{\varepsilon}{2} + 2K\dfrac{\sigma^2}{\alpha^2}$

c) En déduire que la suite $(E(f(S_n)))_n$ converge vers $f(m)$.

4. Montrer que
$$\lim_{n\to+\infty}\left[\frac{1}{(n-1)!}\int_0^{+\infty} f(x)\Big(\frac{nx}{m}\Big)^{n-1}\exp\Big(-\frac{nx}{m}\Big)\frac{n\,dx}{m}\right] = f(m)$$

Solution :

1. Les variables aléatoires indépendantes $X_1,\dots,X_n$ suivant la même loi $\mathcal{E}(1/m)$, la variable aléatoire $\sum\limits_{i=1}^{n} X_i$ suit la loi $\Gamma(1/m, n)$ de densité :
$$g(x) = \frac{1}{m(n-1)!}\Big(\frac{x}{m}\Big)^{n-1}\mathrm{e}^{-x/m},\ \text{pour } x \geqslant 0$$

2. Il s'agit d'écrire la continuité de f en m, soit :
$$\forall\,\varepsilon > 0, \exists\,\alpha > 0, |x-m| < \alpha \Longrightarrow |f(x) - f(m)| < \varepsilon/2$$

3. a) La fonction f étant continue et majorée par K, on a $|f(t)f_Z(t)| \leqslant Kf_Z(t)$, ce qui assure l'existence de $\displaystyle\int_{\mathbb{R}^+}|f(t)f_Z(t)|dt$ et donc l'existence de $E(f(Z))$ par le théorème du transfert.

b) On écrit
$$\begin{aligned}|E(f(Z)) - f(m)| &\leqslant \int_0^{+\infty}|f(t) - f(m)|f_Z(t)dt\\ &\leqslant \int_0^{m-\alpha}\cdots + \int_{m-\alpha}^{m+\alpha}\cdots + \int_{m+\alpha}^{+\infty}\cdots = I_1 + I_2 + I_3\end{aligned}$$

- par la question 2, $0 \leqslant I_2 \leqslant \dfrac{\varepsilon}{2}\displaystyle\int_{m-\alpha}^{m+\alpha} f_Z(t)dt < \frac{\varepsilon}{2}$.
- la fonction $|f|$ étant majorée par K,

$$0 \leqslant I_1 + I_3 \leqslant 2K\Big(\int_0^{m-\alpha} f_Z(t)dt + \int_{m+\alpha}^{+\infty} f_Z(t)dt\Big) = 2KP(|Z-m| > \alpha)$$

$$\leqslant 2K\frac{\sigma^2}{\alpha^2}, \text{ par l'inégalité de Bienaymé-Tchebicheff.}$$

c) On applique l'inégalité précédente à $Z = S_n$, avec $E(S_n) = m$ et $\sigma^2 = \dfrac{m}{n}$, soit :

$$|E(f(Z)) - f(m)| \leqslant \frac{\varepsilon}{2} + 2K\frac{m}{n\alpha^2}$$

Il reste à faire tendre n vers $+\infty$ pour obtenir le résultat demandé.

4. Appliquons le résultat précédent à la loi de S_n.

D'abord : $P(S_n \leqslant x) = P\big(\sum\limits_{i=1}^{n} X_i \leqslant nx\big) \Longrightarrow f_{S_n}(x) = ng(nx)$

soit, pour $x \geqslant 0$ $f_{S_n}(x) = \dfrac{n}{m(n-1)!}\big(\dfrac{nx}{m}\big)^{n-1}\mathrm{e}^{-nx/m}$.

Par le théorème de transfert (l'intégrale converge puisque la fonction f est bornée)

$$E(f(S_n)) = \frac{1}{(n-1)!}\int_0^{+\infty} f(x)\big(\frac{nx}{m}\big)^{n-1}\exp\big(-\frac{nx}{m}\big)\frac{ndx}{m}$$

On termine par la question précédente.

Exercice 3.08.

Dans tout cet exercice, n désigne un entier naturel non nul. On note $\mathcal{E}$ l'ensemble des variables aléatoires X, à valeurs dans $[\![1, n]\!]$ et telles que pour tout entier $k \in [\![1, n]\!]$, $P(X = k) \neq 0$ et définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

Pour toute variable aléatoire X de $\mathcal{E}$, on note :

$$H(X) = -\sum_{k=1}^{n} P(X = k)\times \ln(P(X = k)).$$

On note U une variable aléatoire de loi uniforme sur $[\![1, n]\!]$.

1. Déterminer $H(U)$.

2. Montrer que pour tout $X \in \mathcal{E}$, $H(X) \leqslant H(U)$.

Soient $f : [\![1, n]\!] \to \mathbb{R}$ une fonction injective et $a \in \mathbb{R}$. On considère l'ensemble $\mathcal{E}_a$ des variables aléatoires de $\mathcal{E}$ telles que $E(f(X)) = a$. On suppose que $\mathcal{E}_a$ est non vide.

3. Montrer que la fonction $\varphi : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sum\limits_{k=1}^{n} (f(k) - a)\mathrm{e}^{(f(k)-a)x}$ est une bijection strictement croissante de $\mathbb{R}$ sur $\mathbb{R}$.

4. Pour tout $X \in \mathcal{E}_a$ et $k \in [\![1, n]\!]$, on note $p_k = P(X = k)$. On peut ainsi considérer H comme une fonction de $(p_1, \ldots, p_n)$. Déterminer la hessienne de H.

5. En déduire que H admet un minimum sur $\mathcal{E}_a$ atteint en un point unique.

Solution :

1. On rappelle que pour tout $k \in [\![1, n]\!]$, $P(U = k) = \dfrac{1}{n}$.
Ainsi, $H(U) = -\sum\limits_{k=1}^{n} \dfrac{1}{n} \ln \dfrac{1}{n} = \ln n$.

2. Soit $X \in \mathcal{E}$. Alors, en notant $p_k = P(X = k)$, la fonction logarithme néperien étant concave,

$$\begin{aligned} H(U) - H(X) &= \ln n + \sum_{k=1}^{n} p_k \ln p_k = \sum_{k=1}^{n} p_k \ln(np_k) = -\sum_{k=1}^{n} p_k \ln \frac{1}{np_k} \\ &\geqslant \ln\Big(\sum_{k=1}^{n} \frac{p_k}{np_k}\Big) \geqslant \ln 1 \end{aligned}$$

3. La fonction φ est continue et dérivable sur $\mathbb{R}$ et pour tout réel x,

$$\varphi'(x) = \sum_{k=1}^{n} (f(k) - a)^2 \mathrm{e}^{(f(k)-a)x}.$$

Ainsi, la fonction φ est une fonction continue et strictement croissante, donc elle réalise une bijection de $\mathbb{R}$ sur $\varphi(\mathbb{R})$.
Comme $\mathcal{E}_a$ est non vide, il existe $X \in \mathcal{E}_a$ telle que $E(f(X)) = a$, *i.e.* telle que $\sum\limits_{k=1}^{n} p_k(f(k) - a) = 0$.
Ainsi, comme les p_k sont strictement positifs, les $(f(k) - a)$ ne peuvent pas tous être de même signe ni tous nuls car f est injective. En factorisant donc par ceux qui sont maximaux ou par ceux qui sont minimaux, on obtient :

$$\lim_{x \to +\infty} \varphi(x) = +\infty, \quad \lim_{x \to -\infty} \varphi(x) = -\infty.$$

Finalement, φ réalise une bijection strictement croissante de $\mathbb{R}$ sur $\mathbb{R}$.

4. On remarque que pour tout $k \in [\![1, n]\!]$, $\dfrac{\partial H}{\partial p_k}(X) = -\ln p_k - 1$. Ainsi, la hessienne de H est diagonale et ses coefficients diagonaux sont les nombres $-\dfrac{1}{p_k}$.

5. En utilisant les notations précédentes, on cherche les extremums de la fonction

$$H : (p_1, \ldots, p_n) \mapsto -\sum_{k=1}^{n} p_k \ln p_k$$

sous les conditions $\sum\limits_{k=1}^{n} p_k = 1$ et $\sum\limits_{k=1}^{n} f(k)p_k = a$. Ainsi, le gradient de H doit appartenir à l'espace vectoriel engendré par les vecteurs $(1, \ldots, 1)$ et $(f(1), \ldots, f(n))$. Donc il existe $\lambda, \mu \in \mathbb{R}$ tels que pour tout $k \in [\![1, n]\!]$, $\ln p_k + 1 = \lambda + \mu f(k)$, ou encore $p_k = \mathrm{e}^{\lambda-1} e^{\mu f(k)}$.

De plus, on doit avoir : $\sum\limits_{k=1}^{n} p_k = 1$ et $\sum\limits_{k=1}^{n} (f(k)-a)p_k = 0$.

Or :

$$\begin{aligned}\sum_{k=1}^{n} (f(k)-a)p_k &= e^{\lambda-1}\sum_{k=1}^{n}(f(k)-a)e^{\mu f(k)} = e^{\lambda-1+a\mu}\sum_{k=1}^{n}(f(k)-a)e^{\mu(f(k)-a)}\\ &= e^{\lambda-1+a\mu}\varphi(\mu), \text{ donc } \varphi(\mu)=0.\end{aligned}$$

Alors, $\sum\limits_{k=1}^{n} p_k = e^{\lambda-1}\sum\limits_{k=1}^{n} e^{\mu f(k)} = 1$.

Comme la fonction φ est bijective, il existe un unique μ_0 tel que $\varphi(\mu_0) = 0$.

On pose alors $\lambda_0 = 1 - \ln\big(\sum\limits_{k=1}^{n} e^{\mu_0 f(k)}\big)$.

Ainsi, la fonction H possède un extremum atteint une fois sur $\mathcal{E}_a$. De plus, H est convexe, donc cet extremum est un minimum.

Exercice 3.09.

1. a) Montrer que la restriction de la fonction tan à l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$ réalise une bijection de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ dans $\mathbb{R}$. On note arctan sa bijection réciproque.

b) Déterminer les limites de arctan en $+\infty$ et en $-\infty$.

c) Montrer que arctan est impaire.

d) Montrer que arctan est dérivable sur $\mathbb{R}$ et calculer $(\arctan)'(x)$.

e) Montrer que pour $x > 0$, $\arctan x + \arctan\dfrac{1}{x} = \dfrac{\pi}{2}$ et que pour $x < 0$, $\arctan x + \arctan\dfrac{1}{x} = -\dfrac{\pi}{2}$.

f) Montrer que arctan admet un développement limité à l'ordre 3 en 0 et expliciter ce développement.

Soit $k \in \mathbb{R}$ et $f : x \longmapsto \dfrac{k}{1+x^2}$.

2. a) Déterminer k de telle sorte que f soit une densité de probabilité.

b) Soit X une variable aléatoire de densité f. X admet-elle un moment d'ordre 1 ?

3. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes et de même loi que X. Pour tout $n \geqslant 1$, on pose $Y_n = \dfrac{1}{n}\sup(X_1, \ldots, X_n)$.

a) Pour tout $n \geqslant 1$, déterminer la fonction de répartition de Y_n.

b) Montrer que (Y_n) converge en loi vers une variable aléatoire Y de fonction de répartition notée F que l'on déterminera.

c) On pose $Z = \dfrac{1}{Y}$. Déterminer la loi de Z.

Solution :

1. a) L'étude de la fonction tan montre qu'elle est strictement croissante de $]-\pi/2, \pi/2[$ sur $\mathbb{R}$ car $(\tan)'(x) = 1 + \tan^2(x)$.

C'est donc une bijection continue de $]-\pi/2, \pi/2[$ sur $\mathbb{R}$ de réciproque continue.

b) Comme $(\tan)'(x) \neq 0$, la fonction arctan est de classe C^1 sur $\mathbb{R}$ et

$$(\arctan)'(x) = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2}$$

c) Les limites sont $-\pi/2$ et $\pi/2$.

d) En dérivant, il vient $\dfrac{d}{dx}\left(\arctan x + \arctan \dfrac{1}{x}\right) = 0$.

Ainsi $x \mapsto \arctan x + \arctan \dfrac{1}{x}$ est constante sur chaque intervalle de continuité, c'est-à-dire sur $\mathbb{R}^{+*}$ et $\mathbb{R}^{-*}$.

Il reste à prendre la valeur, par exemple, en $x = \pi/4$ et $x = -\pi/4$.

e) Comme $\dfrac{d}{dx}(\arctan x) = \dfrac{1}{1+x^2} = 1 - x^2 + o(x^2)$

on intègre ce développement limité pour obtenir $\arctan x = x - \dfrac{x^3}{3} + o(x^3)$

2. a) On doit demander $k > 0$ et $\displaystyle\int_{\mathbb{R}} \frac{k}{1+x^2}dx = 1$ soit $k = \dfrac{1}{\pi}$.

b) X n'a pas d'espérance puisque l'intégrale $\displaystyle\int_{\mathbb{R}} \frac{|x|}{1+x^2}dx$ diverge.

3. a) Pour tout x réel

$P(Y_n \leqslant x) = P(\sup\limits_{1 \leq i \leq n} X_i \leqslant nx) = P\left(\bigcap\limits_{i=1}^{n} [X_i \leqslant nx]\right) = \prod\limits_{i=1}^{n} P([X_i \leqslant nx])$

$P(Y_n \leqslant x) = \left(\frac{1}{2} + \frac{1}{\pi}\arctan(nx)\right)^n$

b) • si $x < 0$, par la question 1 : $P(Y_n \leqslant x) = \dfrac{1}{\pi^n}(\arctan \dfrac{-1}{nx})^n \leqslant \dfrac{1}{2^n} \underset{n\to\infty}{\longrightarrow} 0$

• si $x = 0$, $P(Y_n \leqslant 0) = \dfrac{1}{2^n} \underset{n\to\infty}{\longrightarrow} 0$

• si $x > 0$, avec un développement limité (en fait un équivalent) :

$\ln[P(Y_n \leqslant x)] = n\ln\left(1 - \frac{1}{\pi}\arctan\frac{1}{nx}\right) = -\frac{1}{\pi x} + o(1) \underset{n\to\infty}{\longrightarrow} -\frac{1}{\pi x}$

Ainsi : $F(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \mathrm{e}^{-1/\pi x} & \text{si } x > 0 \end{cases}$

c) On sait que $Y(\Omega) = \mathbb{R}_+^*$ et, pour $x > 0$:

$F_Z(x) = P(Y \geqslant \frac{1}{x}) = 1 - P(Y \leqslant \frac{1}{x}) = 1 - \mathrm{e}^{\frac{-1}{\pi x}}$

Ainsi Z suit la loi exponentielle de paramètre $\dfrac{1}{\pi}$.

Exercice 3.10.

Soient n et p deux entiers supérieurs ou égaux à 2. L'entier p restera fixé dans tout l'exercice. On considère np variables aléatoires indépendantes $X_1, \dots, X_{np}$. Pour tout k appartenant à $[\![1, p]\!]$, on suppose que la variable aléatoire X_k suit une loi de Poisson de paramètre λ_1 et pour tout $k \in [\![p+1, np]\!]$ que X_k suit une loi de Poisson de paramètre λ_2. On pose

$$S_n = \sum_{i=1}^{np} X_i, \quad A = \sum_{i=1}^{p} X_i, \quad B_n = \sum_{i=p+1}^{np} X_i.$$

1. Déterminer les lois des variables aléatoires A et B_n.

2. Quelle est la loi de S_n ?

3. Soit $\ell \in \mathbb{N}$.

a) On considère une variable aléatoire Y_ℓ dont la loi est donnée par :

$P(Y_\ell = m) = P_{(S_n=\ell)}(A = m)$ pour tout $m \in \mathbb{N}$.

Déterminer explicitement cette loi.

b) De même, pour $\ell \in \mathbb{N}$, on considère une variable aléatoire Z_ℓ dont la loi est donnée par $P(Z_\ell = m) = P_{(S_n=\ell)}(B_n = m)$ pour tout $m \in \mathbb{N}$. Expliciter cette loi.

4. On définit les deux variables U_n et V_n par :

$$U_n = \frac{\lambda_1}{\lambda_1 + (n-1)\lambda_2} S_n \text{ et } V_n = \frac{(n-1)\lambda_2}{\lambda_1 + (n-1)\lambda_2} S_n$$

Montrer que les suites de variables aléatoires $(\frac{U_n}{n})_n$ et $(\frac{V_n}{n})_n$ convergent en probabilité vers deux variables aléatoires que l'on déterminera.

Solution :

1. Comme $X_1, \dots, X_p$ sont indépendantes et suivent la même loi de Poisson de paramètre λ_1, A suit une loi de Poisson de paramètre $p\lambda_1$.

De la même manière on montre que B_n suit une loi de Poisson de paramètre $(n-1)p\lambda_2$.

2. Les variables A et B_n sont construites à partir de deux blocs disjoints d'une famille de variables indépendantes, le cours nous dit que A et B_n sont indépendantes.

Prenant en compte les résultats de la question 1, on en déduit que $S_n = A + B_n$ suit une loi de Poisson de paramètre $p\lambda_1 + (n-1)p\lambda_2$.

3. a) Si $m > \ell$, on a $P(Y_\ell = m) = P(A = m|S_n = \ell) = 0$. Si $0 \leqslant m \leqslant \ell$, il vient :

$$\begin{aligned}
P(Y_\ell = m) = P(A = m|S_n = \ell) &= \frac{P((A = m) \cap (S_n = \ell))}{p(S_n = \ell)} \\
&= \frac{P((A = m) \cap (B_n = \ell - m))}{p(S_n = \ell)} \\
&= \frac{P(A = m)P(B_n = \ell - m)}{p(S_n = \ell)} \\
&= \frac{\ell!}{m!(\ell - m)!} \frac{(p\lambda_1)^m((n-1)p\lambda_2)^{\ell-m}}{(p\lambda_1 + (n-1)p\lambda_2)^\ell} \\
&= \binom{m}{\ell}\left(\frac{\lambda_1}{\lambda_1 + (n-1)\lambda_2}\right)^m\left(1 - \frac{\lambda_1}{\lambda_1 + (n-1)\lambda_2}\right)^{\ell-m}.
\end{aligned}$$

La variable Y suit donc une loi binomiale de paramètres $(\ell, \dfrac{\lambda_1}{\lambda_1 + (n-1)\lambda_2})$.

b) De façon analogue on trouve que Z suit une loi binomiale de paramètres $(\ell, \dfrac{(n-1)\lambda_2}{\lambda_1 + (n-1)\lambda_2})$.

4. Avec les formules de la question précédente, on trouve aisément

$$E\left(\frac{U_n}{n}\right) = \frac{1}{n}\frac{\lambda_1}{\lambda_1 + (n-1)\lambda_2}(p\lambda_1 + (n-1)p\lambda_2) = \frac{p}{n}\lambda_1$$

$$E\left(\frac{V_n}{n}\right) = (1 - \frac{1}{n})p\lambda_2$$

$$V\left(\frac{U_n}{n}\right) = \frac{1}{n^2}\left(\frac{\lambda_1}{\lambda_1 + (n-1)\lambda_2}\right)^2 V(S_n) = \frac{p\lambda_1^2}{n^2}\frac{1}{\lambda_1 + (n-1)\lambda_2}$$

$$V\left(\frac{V_n}{n}\right) = (1 - \frac{1}{n})^2 p\lambda_2^2 \frac{1}{\lambda_1 + (n-1)\lambda_2}.$$

Soit $\varepsilon > 0$, pour n assez grand on voit avec l'inégalité triangulaire que

$$\{|\frac{V_n}{n} - p\lambda_2| > \varepsilon\} \subseteq \{|\frac{V_n}{n} - (1 - 1/n)p\lambda_2| > \varepsilon/2\}.$$

En utilisant l'inégalité de Bienaymé-Tchebychev, on en déduit que la suite $(\dfrac{V_n}{n})$ converge en probabilité vers la variable certaine $p\lambda_2$ et la suite de variables $(\dfrac{U_n}{n})$ converge en probabilité vers 0.

Exercice 3.11.

Soit X et Y deux variables aléatoires à densité indépendantes définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ de densités respectives f et g nulles hors du segment $[0, 1]$.

1. a) On pose $U = \ln X$, déterminer une densité de U.

b) On pose $V = \ln(1 - Y)$, déterminer une densité de V.

c) Montrer qu'une densité de la variable aléatoire $Z = X(1-Y)$ est donnée par

$$x \mapsto \int_x^1 f\big(\frac{x}{t}\big)g(1-t)\frac{dt}{t}$$

pour $x \in]0,1[$ et $x \mapsto 0$ sinon.

Ce résultat pourra être utilisé dans la suite de cet exercice, même s'il n'a pas été démontré.

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, de même loi à valeurs dans $[0,1]$. On lui associe la suite $(S_n)_{n\in\mathbb{N}^*}$ de segments définie par $S_n = \big[\prod_{i=1}^{n+1} X_i, \prod_{i=1}^{n} X_i\big]$ de longueur L_n.

2. Montrer que pour tout $n \geqslant 2$, on a $L_n = X_1 Y_{n-1}$, où Y_{n-1} est une variable aléatoire indépendante de (X_1) de même loi que L_{n-1}.

3. On suppose que la loi commune des variables aléatoires X_n est la loi uniforme sur $[0,1]$. Déterminer, par récurrence, la loi de L_n.

4. Déterminer l'espérance et la variance de L_n.

Solution :

1. a) b) c) Pour déterminer la loi de $X(1-Y)$ passons au logarithme. Comme $X(\Omega) = Y(\Omega) =]0,1[$, on a $\ln X(\Omega) = \ln(1-Y)(\Omega) = \mathbb{R}^{-*}$.

Soit $x < 0$. On a, par croissance de la fonction exponentielle :

$$P(\ln(X) < x) = P(X < \mathrm{e}^x) = F_X(e^x) \Longrightarrow f_{\ln X}(x) = \mathrm{e}^x f(\mathrm{e}^x)$$

De même : $P(\ln(1-Y) < x) = P(Y > 1-\mathrm{e}^x) = 1 - F_Y(1-\mathrm{e}^x)$, donne :

$$f_{\ln(1-Y)}(x) = \mathrm{e}^x g(1-\mathrm{e}^x)$$

Par indépendance, comme $\ln Z = \ln(X) + \ln(1-Y)$, en demandant $u < 0$ et $x - u < 0$:

$$f_{\ln Z}(x) = \int_{\mathbb{R}} f_{\ln X}(x-u) f_{\ln(1-Y)}(u)du = \int_{-\infty}^0 f_{\ln X}(x-u)\mathrm{e}^u g(1-\mathrm{e}^u)du$$

$$f_{\ln Z}(x) = \int_x^0 \mathrm{e}^{x-u} f(\mathrm{e}^{x-u})\mathrm{e}^u g(1-\mathrm{e}^u)du$$

$$= \mathrm{e}^x \int_x^0 f(\mathrm{e}^{x-u})g(1-\mathrm{e}^u)du = \mathrm{e}^x \int_{\mathrm{e}^x}^0 f(\frac{\mathrm{e}^x}{v})g(1-v)\frac{dv}{v} \text{ (avec } v = e^u)$$

Or, pour $x \in]0,1[, P(X(1-Y) \leqslant x) = P(\ln Z \leqslant \ln x)$ et

$$f_{X(1-Y)}(x) = \frac{1}{x} f_{\ln Z}(\ln x), \text{ soit :}$$

$$f_{X(1-Y)}(x) = \frac{1}{x} \times x \int_x^1 f(\frac{x}{v})g(1-v)\frac{dv}{v} = \int_x^1 f(\frac{x}{v})g(1-v)\frac{dv}{v}$$

2. De manière évidente $L_n = X_1\big(\prod_{i=2}^{n} X_i - \prod_{i=2}^{n+1} X_i\big) = X_1 Y_{n-1}$, avec Y_{n-1} indépendante de X_1 et de même loi que L_{n-1}.

3. $\star$ Déterminons la loi de $L_1 = X_1(1 - X_2)$.

Par la première question, pour tout $x \in]0,1[$, $f_{L_1}(x) = \displaystyle\int_x^1 \frac{dt}{t} = -\ln x$.

$\star$ Déterminons la loi de $L_2 = X_1 L_1 = X_1(1-(1-L_1))$.

On a : $P(1 - L_1 \leqslant x) = P(L_1 \geqslant 1 - x) = 1 - F_{L_1}(1-x)$, donc :

$f_{1-L_1}(x) = f_{L_1}(1-x) = -\ln(1-x)$, pour $0 < x < 1$ et ainsi :

$$f_{L_2}(x) = -\int_x^1 1 \times \ln(v)\frac{dv}{v} = \frac{\ln^2 x}{2} = \frac{(-\ln x)^2}{2}$$

On montre alors par une récurrence facile que la loi de L_n est donnée pour $x \in]0,1[$ par $f_{L_n}(x) = \dfrac{(-\ln x)^n}{n!}$

4. Calculons les deux premiers moments de L_n. Avec le changement de variable $u = -\ln x$, il vient :

$$E(L_n) = \int_0^1 x f_{L_n}(x)dx = \int_0^1 x\frac{(-\ln(x))^n}{n!}dx = \int_0^{+\infty} \frac{u^n}{n!}\mathrm{e}^{-2u}du = \frac{1}{2^{n+1}}$$

(poser $u = \dfrac{t}{2}$ pour retrouver une intégrale du programme)

$$E(L_n^2) = \int_0^1 x^2\frac{(-\ln(x))^n}{n!}dx = \int_0^{+\infty} \frac{u^n}{n!}\mathrm{e}^{-3u}du = \frac{1}{3^{n+1}} \quad \text{(idem)}$$

et :

$$V(L_n) = \frac{1}{3^{n+1}} - \frac{1}{4^{n+1}}$$

Exercice 3.12.

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

Soit $(X_i)_{i\in[\![1,n]\!]}$ une suite de variables aléatoires indépendantes de même loi uniforme sur $[0,1]$. Pour tout entier naturel n non nul, on note $M_n = \sup\{X_i,\ i \in [\![1,n]\!]\}$.

1. Soit $n \in \mathbb{N}^*$. Déterminer la fonction de répartition F_n de la variable aléatoire $n(M_n - 1)$.

2. Montrer que pour tout $x \in \mathbb{R}$, $(F_n(x))_n$ converge vers une valeur $g(x)$.

3. Montrer que g', définie sur $\mathbb{R}^*$, est une densité de probabilité. On notera G une variable aléatoire de densité g'.

4. Déterminer $E(G)$.

Soient λ un réel strictement positif et $(X_i)_{1\leq i\leq n}$ une suite de variables aléatoires indépendantes de même loi exponentielle de paramètre λ. Pour tout entier naturel non nul n, on pose $m_n = \inf\{X_i, 1 \leq i \leq n\}$ et $M_n = \sup\{X_i,\, 1 \leqslant i \leqslant n\}$.

5. Soit $n \in \mathbb{N}^*$. Montrer que m_n suit une loi exponentielle de paramètre $n\lambda$. En déduire la loi de la variable aléatoire nm_n.

6. Soit $n \in \mathbb{N}^*$. Déterminer la fonction de répartition F_n de la variable aléatoire $Z_n = \lambda M_n - \ln n$.

7. Montrer que pour tout $t \in \mathbb{R}$, il existe $g(t) \in \mathbb{R}$ tel que $\lim\limits_{n\to+\infty} F_n(t) = g(t)$.

8. Montrer que g est une fonction de répartition.

Solution :

1. Soit $n \in \mathbb{N}^\star$, $n(M_n - 1)$ prend ses valeurs entre $-n$ et 0 et pour x dans ce domaine :

$$\begin{aligned}P(n(M_n - 1) \leqslant x) &= P(\forall\ i \in [\![1, n]\!],\, X_i \leqslant \frac{x}{n} + 1) = [P(X_1 \leqslant \frac{x}{n} + 1)]^n\\ &= \left(1 + \frac{x}{n}\right)^n.\end{aligned}$$

Si $x \in \mathbb{R}_+$, $P(n(M_n - 1) \leqslant x) = 1$ et si $x < -n, P(n(M_n - 1) \leqslant x) = 0$.

2. Pour tout $x \in \mathbb{R}_-$, $\lim\limits_{n\to+\infty} F_n(x) = \mathrm{e}^x$ et pour tout $x \in \mathbb{R}_+$, $\lim\limits_{n\to+\infty} F_n(x) = 1$. On pose ainsi $g(x) = \min\{\mathrm{e}^x, 1\}$.

3. La fonction g est croissante sur $\mathbb{R}$, de limite nulle en $-\infty$ et de limite 1 en $+\infty$. De plus g est dérivable sur $\mathbb{R}^*$, avec pour $x \in \mathbb{R}_-^*$, $g'(x) = \mathrm{e}^x$ et g' est nulle sur $\mathbb{R}_+^*$: g' est bien une densité de probabilité (on prend ce que l'on veut en 0).

4. En utilisant la formule de transfert, et en passant tout de suite à la limite :

$$E(G) = \int_{\mathbb{R}_-} x\mathrm{e}^x dx = \big[x\mathrm{e}^x\big]_{\to-\infty}^0 - \int_{\mathbb{R}_-} \mathrm{e}^x dx = -1$$

5. Soit $x \in \mathbb{R}_+$.

$$\begin{aligned}P(m_n > x) &= P(\forall\, i \in [\![1, n]\!],\, X_i > x) = P(X_1 > x)^n\\ &= \mathrm{e}^{-n\lambda x}\end{aligned}$$

Ainsi, $P(m_n \leqslant x) = 1 - \mathrm{e}^{-n\lambda x}$ et m_n suit une loi exponentielle de paramètre $n\lambda$.

En reprenant le calcul précédent, ou en utilisant les propriétés de la loi exponentielle, nm_n suit une loi exponentielle de paramètre λ.

6. Soit $x \in \mathbb{R}$.

$$\begin{aligned}F_n(x) &= P(M_n \leqslant \frac{x+\ln n}{\lambda}) = P(\forall\, i \in [\![1,n]\!], X_i \leqslant \frac{x+\ln n}{\lambda})\\ &= [P(X_1 \leqslant \frac{x+\ln n}{\lambda})]^n = \big(1-\mathrm{e}^{-\lambda\frac{x+\ln n}{\lambda}}\big)^n \mathbf{1}_{[x\geq -\ln n]}\\ &= \big(1-\frac{\mathrm{e}^{-x}}{n}\big)^n \mathbf{1}_{[x\geq -\ln n]}.\end{aligned}$$

7. Ainsi, pour tout $x \in \mathbb{R}$, $\lim\limits_{n\to+\infty} F_n(x) = \mathrm{e}^{-\mathrm{e}^{-x}}$.

8. La fonction g est continue, strictement croissante, de limite nulle en $-\infty$ et de limite 1 en $+\infty$. Tout est dit.

Exercice 3.13.

1. On définit sur $\mathbb{R}$ la fonction g par $g(x) = \displaystyle\int_{-1}^{1} |x-t|\,dt$.

Donner, suivant les valeurs de x, l'expression explicite de $g(x)$ en fonction de x et vérifier que g est continue sur $\mathbb{R}$.

Dans la suite de l'exercice, X est une variable à densité définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. On définit sur Ω l'application Y par :

$$\forall\, \omega \in \Omega, Y(\omega) = \int_{-1}^{1} \big|X(\omega)-t\big|\,dt$$

On admet que Y est une variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$.

2. On suppose dans cette question que X suit la loi uniforme sur $[-1,1]$.
   a) Exprimer Y en fonction de X.
   b) Donner la fonction de répartition de Y.
   c) Vérifier que Y est une variable à densité et donner une densité de Y.
   d) Calculer l'espérance de Y.

3. On considère dans cette question une suite de variables aléatoires $(X_n)_{n\geq 1}$, définies sur $(\Omega, \mathcal{A}, P)$, où, pour tout entier naturel $n \geqslant 1$, X_n suit la loi normale d'espérance nulle et d'écart-type $1/n$.

On définit, pour tout entier naturel n non nul, l'application Y_n par :

$$\forall \omega \in \Omega, Y_n(\omega) = \int_{-1}^{1} \big|X_n(\omega)-t\big|\,dt$$

On admet que, pour tout entier naturel non nul, Y_n est une variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$.

On note F_{Y_n} la fonction de répartition de Y_n et Φ celle de la loi normale centrée réduite.

   a) Exprimer, pour tout réel y, $F_{Y_n}(y)$ en fonction de $\Phi(y)$ et de n.
   b) Montrer que la suite $(Y_n)_{n\geq 1}$ converge en loi.

Solution :

1. Il faut distinguer trois cas :

$\star$ si $x < 1$, alors, pour tout t de $[-1, 1]$, on a $x < t$, d'où $|x - t| = t - x$. On a donc dans ce cas $g(x) = \displaystyle\int_{-1}^{1} (t - x)\, dt = -2x$.

$\star$ si $-1 \leq x \leq 1$, on procède de manière analogue et :

$$g(x) = \int_{-1}^{x} (x - t)dt + \int_{x}^{1} (t - x)dt = x^2 + 1.$$

$\star$ si $x > 1$, on a $|x - t| = x - t$ et $g(x) = 2x$.

En conclusion : $\forall\, x \in \mathbb{R}, g(x) = \begin{cases} -2x & \text{si } x < -1 \\ x^2 + 1 & \text{si } -1 \leqslant x \leqslant 1 \\ 2x & \text{si } x > 1 \end{cases}$

On vérifie alors sans problème que g est continue sur $\mathbb{R}$.

2. a) On utilise la question précédente ou pour tout ω de Ω, le nombre $x = X(\omega)$ est dans l'intervalle $[-1, 1]$.

On a donc, en utilisant la relation de Chasles, comme ci-dessus :

$Y(\omega) = X^2(\omega) + 1$, cette égalité étant valable pour tout élément ω de Ω, on a finalement, en simplifiant :

$$Y = X^2 + 1$$

b) On constate donc que $Y(\Omega) = [1, 2]$.

Si y est dans $[1, 2]$, on a donc : $P([Y \leq y]) = P([X^2 + 1 \leq y]) = P([X^2 \leq y - 1])$.

Soit encore : $P([Y \leq y]) = P([-\sqrt{y-1} \leq X \leq \sqrt{y-1}])$. On a donc finalement, pour $y \in [1, 2]$, $F_Y(y) = F_X(\sqrt{y-1}) - F_X(-\sqrt{y-1})$.

Pour tout réel x de $[-1, 1]$, on a $F_X(x) = \dfrac{x+1}{2}$.

On a donc, pour y dans $[1, 2]$, $F_Y(y) = \dfrac{\sqrt{y-1}+1}{2} - \dfrac{-\sqrt{y-1}+1}{2}$.

En conclusion : $\forall\, y \in \mathbb{R}, \quad F_Y(y) = \begin{cases} 0 & \text{si } y < 1 \\ \sqrt{y-1} & \text{si } 1 \leqslant y \leqslant 2 \\ 1 & \text{si } y > 1 \end{cases}$

c) F_Y est évidemment de classe $\mathcal{C}^1$ sur $]-\infty, 1[$, sur $]1, 2[$ et sur $]2, +\infty[$. On vérifie facilement qu'elle est continue en 1 et en 2. Y est donc bien une variable à densité. On obtient alors une densité en dérivant sur les trois intervalles ouverts précédents et en rajoutant arbitrairement des valeurs en 1 et en 2.

On obtient par exemple :

$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y-1}} & \text{si } y \in]1,2[\\ 0 & \text{sinon} \end{cases}$$

d) On a $E(Y) = \displaystyle\int_1^2 \frac{y}{\sqrt{y-1}}dy$, qui est une intégrale convergente (en fait Y est une variable aléatoire bornée, elle admet donc une espérance).

$$E(Y) = \int_0^1 \frac{y+1}{\sqrt{y}}\,dy = \left[\frac{2}{3}y^{\frac{3}{2}} + 2y^{\frac{1}{2}}\right]_0^1 = \frac{8}{3}$$

(On aurait pu remarquer que $E(Y) = E(X^2)+1$ et utliser la formule donnant le moment d'ordre 2 d'une loi uniforme)

3. a) On utilise le système complet d'événements

$$([X_n < -1], [-1 \leqslant X_n \leqslant 1], [X_n > 1])$$

Ainsi, pour tout y réel :

$$\begin{aligned} P(Y_n \leqslant y) &= P([-2X_n \leqslant y] \cap [X_n < -1]) \\ &\quad + P([X_n^2 + 1 \leqslant y] \cap [-1 \leqslant X_n \leqslant -1]) + P([2X_n \leqslant y] \cap [X_n > 1]) \end{aligned}$$

- si $y \leqslant 1$, on a $y/2 \leqslant 1/2$ et chacune des trois probabilités ci-dessus est nulle.
- si $1 \leqslant y \leqslant 2$, seule la seconde probabilité n'est pas nulle et
$$P(Y_n \leqslant y) = P(-\sqrt{y-1} \leqslant X_n \leqslant \sqrt{y-1})$$
- si $y \geqslant 2$, seule la troisième probabilité n'est pas nulle et
$$P(Y_n \leqslant y) = P([1 \leqslant X_n \leqslant y/2] \cup [-y/2 \leqslant X_n \leqslant -1])$$

On a donc :

$$\forall\, y \in \mathbb{R}, F_{Y_n}(y) = \begin{cases} 0 & \text{si } y \leqslant 1 \\ F_{X_n}\big(\sqrt{y-1}\big) - F_{X_n}\big(-\sqrt{y-1}\big) & \text{si } 1 \leqslant y \leqslant 2 \\ F_{X_n}\big(\frac{y}{2}\big) - F_{X_n}\big(-\frac{y}{2}\big) & \text{si } y \geqslant 2 \end{cases}$$

D'autre part, comme X_n suit la loi $\mathcal{N}(0, \frac{1}{n})$, nX_n suit la loi normale centrée réduite. On a donc, pour tout réel z, $F_{X_n}(z) = \Phi(nz)$.

On obtient finalement :

$$\forall\, y \in \mathbb{R}, F_{Y_n}(y) = \begin{cases} 0 & \text{si } y \leqslant 1 \\ \Phi\big(n\sqrt{y-1}\big) - \Phi\big(-n\sqrt{y-1}\big) & \text{si } 1 \leqslant y \leqslant 2 \\ \Phi\big(\frac{ny}{2}\big) - \Phi\big(-\frac{ny}{2}\big) & \text{si } y \geqslant 2 \end{cases}$$

b) $y \leq 1 \implies \lim\limits_{n\to+\infty} F_{Y_n}(y) = 0$, $y > 1 \implies \lim\limits_{n\to+\infty} F_{Y_n}(y) = 1$.

En tout point où la fonction de répartition F de la variable aléatoire constante égale à 1 est continue, on a donc $\lim\limits_{n\to\infty} F_{Y_n}(x) = F(x)$: $(Y_n)_n$ converge en loi vers la variable constante égale à 1.

Exercice 3.14.

1. Soit Y une variable aléatoire à valeurs positives. Montrer que pour tout réel $b > 0$

$$P(Y \geqslant b) \leqslant \frac{E(Y)}{b}$$

On considère dans la suite une variable aléatoire X et un réel t strictement positif. On suppose que la variable aléatoire e^{tX} possède une espérance.

2. Montrer que, pour tout réel a : $P(X \geqslant a) \leqslant \mathrm{e}^{-ta} \times E(\mathrm{e}^{tX})$.

Dans la suite, on suppose que X suit la loi normale centrée réduite.

3. a) Montrer que, pour tout réel t strictement positif, la variable aléatoire e^{tX} possède une espérance et que : $E(\mathrm{e}^{tX}) = \mathrm{e}^{t^2/2}$.

b) En déduire que pour tout réel $a > 0$, $P(X \geqslant a) \leqslant \mathrm{e}^{-a^2/2}$.

Solution :

1. Soit $\mathbf{1}_B$ la variable indicatrice de l'événement $B = [X \geqslant b]$.
Procédons par disjonction de cas :

- si $\omega \in B$, alors $X(\omega) \geqslant b \implies \dfrac{X(\omega)}{b} \geqslant 1 = \mathbf{1}_B(\omega)$
- si $\omega \notin B$, alors, comme $b > 0$, $\dfrac{X(\omega)}{b} \geqslant 0 = \mathbf{1}_B(\omega)$

Ainsi $\mathbf{1}_B \leqslant \dfrac{X}{b}$. Par croissance et linéarité de l'espérance
$E(\mathbf{1}_B) = P(B) = P(X \geqslant b) \leqslant \dfrac{E(X)}{b}$, ce qui permet de conclure.

2. On applique l'inégalité précédente à la variable positive e^{tX} en remplaçant b par $\mathrm{e}^{ta} > 0$ et on trouve : $P(\mathrm{e}^{tX} \geqslant \mathrm{e}^{ta}) \leqslant \dfrac{E(\mathrm{e}^{tX})}{\mathrm{e}^{ta}}$
La fonction ln est strictement croissante et t strictement positif, on en déduit :

$$P(X \geqslant a) = P(tX \geqslant ta) = P(\mathrm{e}^{tX} \geqslant \mathrm{e}^{ta}) \leqslant \frac{E(\mathrm{e}^{tX})}{\mathrm{e}^{ta}}$$

3. a) L'intégrale définissant $E(\mathrm{e}^{tX})$ est l'intégrale $\dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \mathrm{e}^{tx} \mathrm{e}^{-t^2/2} dx$.
Or : $tx - \dfrac{t^2}{2} = \dfrac{-(x-t)^2}{2} + \dfrac{t^2}{2}$ (forme canonique du trinôme).
On reconnaît l'intégrale d'une fonction proportionnelle à une densité normale et :

$$E(\mathrm{e}^{tX}) = \mathrm{e}^{t^2/2}$$

b) On sait que $P(X \geqslant a) \leqslant \mathrm{e}^{-ta} E(\mathrm{e}^{tX})$, ceci pour tout réel $t > 0$. En remplaçant on trouve :

$$P(X \geqslant a) \leqslant \mathrm{e}^{-ta} \mathrm{e}^{t^2/2}$$

Il reste à prendre $t = a$ pour obtenir le résultat demandé.

Exercice 3.15.

Dans cet exercice, f est une densité de probabilité nulle sur $]-\infty, 0[$ et positive et continue sur $[0, +\infty[$. On suppose de plus que l'intégrale $\displaystyle\int_0^{+\infty} xf(x)\,dx$ est convergente.

1. Soit $t > 0$ un réel fixé. Un client se présente à un automate pour y effectuer un retrait d'argent. On suppose que l'instant d'arrivée X de ce client à partir de l'instant initial noté 0 est une variable aléatoire suivant une loi uniforme sur $]0, t[$.
Le client utilise alors l'automate pendant une durée aléatoire Y, indépendante de X et admettant f comme densité.
On appelle p_t la probabilité que le client soit encore en train d'utiliser l'automate à l'instant t.

a) Montrer que $X + Y$ est une variable aléatoire à densité dont une densité est donnée par la fonction g :

$$g(z) = \frac{1}{t}\int_0^{\min(t,z)} f(z-u)du, \text{ si } z \geqslant 0,\ g(z) = 0 \text{ sinon}$$

b) On note F la fonction de répartition de Y. Montrer que :

$$\forall\, z \in [0, t], g(z) = \frac{F(z)}{t}$$

c) Montrer que $P(X + Y \leqslant t) = F(t) - \dfrac{1}{t}\displaystyle\int_0^t zf(z)\,dz$.

d) Déterminer $\displaystyle\lim_{t\to+\infty} [\,tP(Y > t)\,]$.

e) Justifier que Y admet une espérance et montrer que cette espérance vérifie $E(Y) = \displaystyle\lim_{t\to+\infty} (t\,p_t)$

2. Des clients se présentent dans un hall contenant un grand nombre d'automates. On fixe un instant initial 0 et pour tout réel $t > 0$, on note :

- C_t la variable aléatoire égale au nombre de clients se présentant à l'un des automates entre les instants 0 et t.
- D_t la variable aléatoire égale au nombre de clients encore en train d'utiliser un automate à l'instant t.

On suppose de plus que :

- C_t suit une loi de Poisson de paramètre λt, avec $\lambda \in \mathbb{R}_+^*$ fixé.
- Pour tout $n \in \mathbb{N}$, la loi conditionnelle de D_t sachant $[C_t = n]$ est binomiale de paramètre (n, p_t), où p_t est défini comme dans la première question.

a) Montrer que D_t suit une loi usuelle dont on précisera le ou les paramètre(s).

b) Montrer que D_t admet une espérance que l'on précisera.

c) Montrer la suite de variables aléatoires $(D_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire dont on précisera la loi.

Solution :

1. a) La variable aléatoire X suit la loi uniforme sur $]0,t[$. Une de ses densités est donc donnée par la fonction $h : x \to \mathbf{1}_{]0,t[}\dfrac{1}{t}$, où $\mathbf{1}_{]0,t[}$ désigne la fonction indicatrice de $]0,t[$. Comme X et Y sont supposées indépendantes, $X+Y$ est une variable aléatoire continue dont une densité est donnée par la fonction g définie par :

$$\forall\, z \in \mathbb{R}, g(z) = \int_{-\infty}^{+\infty} h(u)f(z-u)du = \frac{1}{t}\int_0^t f(z-u)du.$$

g est évidemment nulle sur $\mathbb{R}^-$ et on obtient pour $z \geqslant 0$:

$$g(z) = \frac{1}{t}\int_0^{\min(t,z)} f(z-u)du.$$

b) Si $z \in [0,t]$, $\min(t,z) = z$ et $g(z) = \dfrac{1}{t}\displaystyle\int_0^z f(z-u)du$. En effectuant dans l'intégrale le changement de variable $v = z-u$, on trouve

$$g(z) = \frac{1}{t}\int_0^z f(v)dv = \frac{F(z)}{t}.$$

c) Comme t est strictement positif,

$$P(X+Y \leqslant t) = \int_{-\infty}^t g(z)dz = \int_0^t g(z)dz = \int_0^t \frac{F(z)}{t}dz.$$

On effectue alors une intégration par parties en posant $u(z) = F(z)$ et $v'(z) = 1$. On obtient

$$P(X+Y \leqslant t) = F(t) - \frac{1}{t}\int_0^t zf(z)dz.$$

d) Pour tout $t \in \mathbb{R}_+^*$, $0 \leqslant tP[Y>t] = \displaystyle\int_t^{+\infty} tf(z)dz \leqslant \int_t^{+\infty} zf(z)dz.$

Or, par hypothèse, l'intégrale $\displaystyle\int_0^{+\infty} zf(z)dz$ converge.

Par conséquent : $\displaystyle\lim_{t\to+\infty}\int_t^{+\infty} zf(z)dz = 0$. Par encadrement, on en déduit que :

$$\lim_{t\to+\infty}\left[\,tP(Y>t)\,\right] = 0.$$

e) Comme, par hypothèse, l'intégrale $\int_0^{+\infty} zf(z)dz$ est convergente, la variable Y admet une espérance. De plus, pour tout $t \in \mathbb{R}_+^*$,

$$tp_t = tP[X+Y>t] = t(1-P[X+Y\leqslant t]) = t(1-F(t)) + \int_0^t zf(z)dz$$

$$tp_t = tP[Y>t] + \int_0^t zf(z)dz.$$

En passant à la limite, on trouve, d'après la question précédente, que

$$\lim_{t\to+\infty}(tp_t) = \int_0^{+\infty} zf(z)dz = E(Y).$$

2. a) On a déjà $D_t(\Omega) = \mathbb{N}$. En appliquant la formule des probabilités totales au système complet d'événements $([C_t = n])_{n\in\mathbb{N}}$ et en remarquant que l'on doit avoir $D_t \leqslant C_t$, on obtient :

$$\begin{aligned} P(D_t = n) &= \sum_{n=k}^{+\infty}\binom{n}{k}p_t^k(1-p_t)^{n-k}\mathrm{e}^{-\lambda t}\frac{(\lambda t)^n}{n!} = \frac{\mathrm{e}^{-\lambda t}p_t^k(\lambda t)^k}{k!}\sum_{n=k}^{+\infty}\frac{(q_t\lambda t)^{n-k}}{(n-k)!} \\ &= \frac{\mathrm{e}^{-\lambda t}(p_t\lambda t)^k}{k!}\mathrm{e}^{\lambda t q_t} \end{aligned}$$

D'où $\forall\, k\in\mathbb{N}, P(D_t = k) = \dfrac{(\lambda t p_t)^k}{k!}\mathrm{e}^{-\lambda t p_t}$, ce qui prouve que D_t suit une loi de Poisson de paramètre $\lambda t p_t$.

b) Par propriété de la loi de Poisson, $E(D_t) = \lambda t p_t$.

c) Pour tout $n \in \mathbb{N}^*$ et pour tout $k\in\mathbb{N}$, $P(D_n = k) = \dfrac{(\lambda n p_n)^k}{k!}\mathrm{e}^{-\lambda n p_n}$. Or, d'après la question 1.e., $\lim\limits_{n\to\infty}(np_n) = E(Y)$.

D'où, $\lim\limits_{n\to\infty} P(D_n = k) = \dfrac{(\lambda E(Y))^k}{k!}\mathrm{e}^{-\lambda E(Y)}$. La suite (D_n) converge donnc en loi vers une variable D qui suit une loi de Poisson de paramètre $\lambda E(Y)$.

Exercice 3.16.

Dans tout l'exercice, X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et suivant la loi de Poisson de paramètre $\lambda > 0$.

1. a) Montrer que $P(|X-\lambda| \geqslant \lambda) \leqslant \dfrac{1}{\lambda}$.

b) En déduire l'inégalité $P(X \geqslant 2\lambda) \leqslant \dfrac{1}{\lambda}$.

2. On considère dans toute cette question une variable aléatoire discrète Z définie sur $(\Omega, \mathcal{A}, P)$, d'espérance nulle et de variance σ^2.

a) Montrer que

$$\forall\, a>0, \forall\, x \geqslant 0, P(Z\geqslant a) \leqslant P[(Z+x)^2 \geqslant (a+x)^2].$$

b) Montrer que $\forall a>0, \forall x \geqslant 0, P(Z \geqslant a) \leqslant \dfrac{\sigma^2+x^2}{(a+x)^2}$.

c) Montrer que $\forall a>0, P(Z \geqslant a) \leqslant \dfrac{\sigma^2}{\sigma^2+a^2}$.

d) En déduire que $P(X \geqslant 2\lambda) \leqslant \dfrac{1}{\lambda+1}$.

3. Pour tout réel t, on pose $G_X(t) = \sum\limits_{k=0}^{+\infty} P(X=k)\, t^k$.

a) Pour tout réel t, justifier l'existence de $G_X(t)$ et calculer sa valeur.

b) Montrer que : $\forall t \geqslant 1, \forall a>0, P(X \geqslant a) \leqslant \dfrac{G_X(t)}{t^a}$.

c) En déduire que $P(X \geqslant 2\lambda) \leqslant (\frac{\mathrm{e}}{4})^\lambda$.

Solution :

1. a) On applique l'inégalité de Bienaymé-Tchebychev à X. Comme $E(X) = V(X) = \lambda$, on obtient
$$P(|X-\lambda| \geqslant \lambda) \leqslant \frac{V(X)}{\lambda^2} = \frac{1}{\lambda}.$$

b) On remarque que $[|X-\lambda| \geqslant \lambda] = [X-\lambda \geqslant \lambda] \cup [X-\lambda \leqslant -\lambda] \supset [X \geqslant 2\lambda]$. On a donc $P[X \geqslant 2\lambda] \leqslant P(|X-\lambda| \geqslant \lambda) \leqslant \frac{1}{\lambda}$.

2. a) Pour tout $a>0$ et tout $x \geqslant 0$, on a
$$[(Z+x)^2 \geqslant (a+x)^2] = [Z+x \geqslant a+x] \cup [Z+x \leqslant -a-x] \supset [Z \geqslant a].$$
D'où l'inégalité demandée par croissance de la probabilité.

b) On applique l'inégalité de Markov à la variable aléatoire positive $(Z+x)^2$. On obtient :
$$P[(Z+x)^2 \geqslant (a+x)^2] \leqslant \frac{E((Z+x)^2)}{(a+x)^2}.$$
Comme $E(Z)=0$, on a :
$$E((Z+x)^2) = E(Z^2) + 2xE(Z) + x^2 = V(Z) + x^2 = \sigma^2 + x^2.$$
En utilisant la question précédente, on trouve ainsi :
$$P(Z \geqslant a) \leqslant P[(Z+x)^2 \geqslant (a+x)^2] \leqslant \frac{\sigma^2+x^2}{(a+x)^2}.$$

c) Soit $a>0$ un réel fixé. On étudie la fonction $f : x \mapsto \dfrac{\sigma^2+x^2}{(a+x)^2}$ définie sur $\mathbb{R}^+$. Comme, pour tout $x \geqslant 0$, $f'(x) = \dfrac{2(ax-\sigma^2)}{(a+x)^3}$, il s'ensuit que f atteint son minimum en $x = \dfrac{\sigma^2}{a}$, avec $f(\dfrac{\sigma^2}{a}) = \dfrac{\sigma^2}{a^2+\sigma^2}$. D'où la formule demandée.

d) On applique le résultat précédent à la variable aléatoire $Z = X - \lambda$ vérifiant $E(Z) = 0$ et $V(Z) = \lambda = \sigma^2$. En prenant $a = \lambda > 0$, l'inégalité de la question précédente devient :

$$P(X \geqslant 2\lambda) = P(X - \lambda \geqslant \lambda) \leqslant \frac{\sigma^2}{\lambda^2 + \sigma^2} = \frac{1}{\lambda + 1}.$$

3. a) Pour tout $t \in \mathbb{R}$,

$$G_X(t) = \lim_{N \to +\infty} \Big(\sum_{k=0}^{N} P(X = t)t^k\Big) = \mathrm{e}^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} t^k = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda t} = \mathrm{e}^{\lambda(t-1)}.$$

b) Pour tout $t \geqslant 1$ et tout $a > 0$,

$$G_X(t) \geqslant \sum_{k \geq a} P(X = k)t^k \geqslant t^a \sum_{k \geq a} P(X = k) = t^a P(X \geqslant a).$$

D'où l'inégalité demandée.

c) Pour tout $a > 0$ et tout $t \geqslant 1$, on a donc $P(X \geqslant a) \leqslant \dfrac{G_X(t)}{t^a} = \dfrac{\mathrm{e}^{\lambda(t-1)}}{t^a}$.
En prenant $a = 2\lambda > 0$ et $t = 2$, on obtient $P(X \geqslant 2\lambda) \leqslant (\frac{\mathrm{e}}{4})^{\lambda}$.

Exercice 3.17.

Le génotype d'un individu est un ensemble de 2 gènes pris parmi a et A. Trois génotypes sont possibles : 1 (aa), 2 (aA) et 3 (AA) (L'ordre ne compte pas). On s'intéresse à l'évolution d'une population de grande taille (génération 0) dont la proportion du génotype i est notée u_i pour $i \in \{1, 2, 3\}$. On suppose les mariages aléatoires et on rappelle que le génotype d'un enfant est formé d'un gène issu de celui de chaque parent, les deux gènes d'un parent ayant la même probabilité d'être transmis.

Soit E le génotype d'un enfant de la première génération.

1. a) Exprimer les probabilités conditionnelles : $P_{(M=i,F=j)}(E = 1)$ pour $(i, j) \in \{1, 2, 3\}^2$ où F et M sont les génotypes du père et de la mère.

b) En déduire que : $P(E = 1) = (u_1 + \frac{u_2}{2})^2$, puis la valeur de $P(E = 3)$.

c) On pose $\theta = u_1 + \frac{u_2}{2}$. Déterminer en fonction de θ les proportions des divers génotypes à la première génération : q_1, q_2, q_3.

d) Calculer les proportions des divers génotypes à la deuxième génération et en déduire qu'elles sont inchangées au cours du temps.

2. On dispose d'un échantillon de n individus. On note X_i le génotype du $i^{\text{ème}}$ individu, on a alors : $P(X_i = j) = q_j$. Les variables X_i, $1 \leqslant i \leqslant n$ sont supposées indépendantes. Pour $j \in \{1, 2, 3\}$, on note N_j le nombre d'individus de l'échantillon possédant le génotype j.

a) Pour $j = 1, 2, 3$, déterminer la loi de N_j, son espérance et sa variance. En déduire un estimateur sans biais de q_j.

b) Soit n_1, n_2, n_3 3 entiers naturels tels que : $n = n_1 + n_2 + n_3$, montrer que :

$$P(N_1 = n_1, N_2 = n_2, N_3 = n_3) = \frac{n!}{n_1!n_2!n_3!} q_1^{n_1} q_2^{n_2} q_3^{n_3}$$

c) Calculer $\mathrm{Cov}(N_1, N_2)$.

3. Pour tout entier $N > 0$, on pose : $\theta_n = \dfrac{N_1}{n} + \dfrac{N_2}{2n}$

a) Montrer que θ_n est un estimateur sans biais de θ.

b) Est-il convergent ?

Solution :

1. a) Loi du couple (M, P) : $P(M = i, P = j) = u_i u_j$

- $P(E = aa | M = aa, P = aa) = 1$
- $P(E = aa | M = aA, P = aA) = \dfrac{1}{4}$
- $P(E = aa | M = aA, P = aa) = P(E = aa | M = aa, P = aA) = \dfrac{1}{2}$
- $P(E = aa | M = aA, P = AA) = 0$

b) On a :

$$[E = aa] = [E = aa, P = aa, M = aa] \cup [E = aa, P = aA, M = aa]$$
$$\cup [E = aa, P = aa, M = aA] \cup [E = aa, P = aA, M = aA]$$

(dès que l'un des parents est AA l'enfant ne peut être aa !)

Donc par conditionnement :

$P(E = 1) = P(E = aa) = 1u_1^2 + u_2^2 \dfrac{1}{4} + 2u_1 u_2 \dfrac{1}{2} = (u_1 + \dfrac{u_2}{2})^2$.

Par symétrie : $P(E = 3) = (u_3 + \dfrac{u_2}{2})^2$

c) On pose $\theta = u_1 + \dfrac{u_2}{2}$,

- $q_1 = P(E = 1) = (u_1 + \dfrac{u_2}{2})^2 = \theta^2$
- $q_3 = P(E = 3) = (u_3 + \dfrac{u_2}{2})^2 = (1 - u_1 - u_2 + \dfrac{u_2}{2})^2 = (1 - \theta)^2$
- $q_2 = P(E = 2) = 1 - P(E = 1) - P(E = 3) = 2\theta(1 - \theta)$

d) Le paramètre de la deuxième génération est $\theta_2 = q_1 + \dfrac{q_2}{2} = \theta^2 + \theta(1 - \theta) = \theta$ ce paramètre est donc stationnaire au cours du temps.

2. a) La loi de N_j est la loi $\mathcal{B}(n, q_j)$, $E(N_j) = nq_j$ et $V(N_j) = nq_j(1 - q_j)$ et $\dfrac{N_j}{n}$ est un estimateur sans biais de q_j

b) On note A l'ensemble des n-uplets $(x_1, \dots, x_n)$ de $\{1, 2, 3\}^n$ dont n_1 termes valent 1, n_2 termes valent 2 et n_3 termes valent 3.

On a, par indépendance des X_i :

$P(N_1 = n_1, N_2 = n_2, N_3 = n_3) = \sum_A P(X_1 = x_1, \cdots, X_n = x_n)$

$$= \sum_A P(X_1 = x_1) \cdots P(X_n = x_n)$$

$P(N_1 = n_1, N_2 = n_2, N_3 = n_3) = \sum_A q_1^{n_1} q_2^{n_2} q_3^{n_3} = \dfrac{n!}{n_1!n_2!n_3!} q_1^{n_1} q_2^{n_2} q_3^{n_3}$

car $\dfrac{(n_1+n_2+n_3)!}{n_1!n_2!n_3!}$ est le nombre de partitions d'un ensemble à $n = n_1 + n_2 + n_3$ éléments en 3 sous-ensembles à n_1, n_2 et n_3 éléments.

3. Il vient, par indépendance des X_i :

$\mathrm{Cov}(N_1, N_2) = \mathrm{Cov}(\sum_{i=1}^n \mathbf{1}_{X_i=1}, \sum_{j=1}^n \mathbf{1}_{X_j=2}) = \sum_{i=1}^n \sum_{j=1}^n \mathrm{Cov}(\mathbf{1}_{X_i=1}, \mathbf{1}_{X_j=2})$

$\qquad = \sum_{i=1}^n \mathrm{Cov}(\mathbf{1}_{X_i=1}, \mathbf{1}_{X_i=2}) = n(-q_1 q_2)$

et

$E(\theta_n) = E(\frac{N_1}{n} + \frac{N_2}{2n}) = \frac{1}{n}E(N_1) + \frac{1}{2n}E(N_2) = q_1 + \frac{q_2}{2} = \theta^2 + \theta(1-\theta) = \theta.$

Donc θ_n est un estimateur sans biais de θ.

$V(\theta_n) = \dfrac{1}{n^2}(V(N_1) + \frac{1}{4}V(N_2) + \mathrm{Cov}(N_1, N_2)) = \dfrac{\theta(1-\theta)}{2n}$

Donc θ_n est convergent.

Exercice 3.18.

Soit n un entier de $\mathbb{N}^*$.

1. On note F_n l'ensemble des **permutations** de $[\![1, n]\!]$. On choisit au hasard une de ces permutations, et on définit la variable aléatoire Y_n égale au nombre de points fixes de cette permutation.

a) Montrer que le nombre d'éléments de F_n n'admettant aucun point fixe vaut :

$$d_n = n! \sum_{k=0}^n \frac{(-1)^k}{k!}$$

b) Déterminer la loi de Y_n.

c) Montrer que la suite $(Y_n)_n$ converge en loi vers une variable aléatoire dont on déterminera la loi.

2. On note E_n l'ensemble des **applications** de $[\![1, n]\!]$ dans $[\![1, n]\!]$. On choisit au hasard une de ces applications, et on définit la variable X_n égale au nombre de points fixes de cette application.

a) Déterminer la loi de X_n.

b) Montrer que la suite $(X_n)_n$ converge en loi vers une variable aléatoire dont on déterminera la loi.

Solution :

1. a) Soit d_n le nombre de permutations sans point fixes et $q_n = n! - p_n$ le nombre de permutations admettant au moins un point fixe.
Notons A_i les permutation σ de $[\![1, n]\!]$ telles que $\sigma(i) = i$. Par la formule de Poincaré :

$$q_n = \text{card}(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} (-1)^{i+1} \sum_{1 \leq k_1 < \cdots < k_i \leq n} \text{card}(A_{k_1} \cap \cdots \cap A_{k_i})$$
$$= \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} (n-i)! = n! \sum_{i=1}^{n} \frac{(-1)^{i+1}}{i!}$$

Ainsi $d_n = n! \sum\limits_{i=0}^{n} \dfrac{(-1)^i}{i!}$.

b) La variable aléatoire Y_n prend ses valeurs dans $[\![1, n]\!]$. On a, pour tout $k \in [\![1, n]\!]$,

$$P(Y_n = k) = \binom{n}{k} \frac{d_{n-k}}{n!} = \frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!}$$

c) Pour tout $k \in [\![1, n]\!]$, $\lim\limits_{n \to +\infty} \dfrac{1}{k!} \sum\limits_{i=0}^{n-k} \dfrac{(-1)^i}{i!} = \dfrac{\mathrm{e}^{-1}}{k!}$.
Ainsi $(Y_n)_n$ converge en loi vers une loi de Poisson de paramètre 1.

2. a) La variable aléatoire X_n prend ses valeurs dans $[\![0, n]\!]$. Pour tout $k \in [\![0, n]\!]$, pour obtenir l'événement $[X_n = k]$, il faut et il suffit de

- choisir k éléments parmi n, soit $\binom{n}{k}$ possibilités,
- ces k éléments sont fixés,
- donner une image aux $n - k$ éléments restants, autre que lui-même, soit $(n-1)^{n-k}$ choix possibles.

Ainsi : $P(X_n = k) = \dbinom{n}{k} \dfrac{(n-1)^{n-k}}{n^n} = \dbinom{n}{k} \dfrac{1}{n^k} \Big(1 - \dfrac{1}{n}\Big)^{n-k}$

b) On a $P(X_n = k) = \dfrac{n(n-1)\cdots(n-k+1)}{n^k k!} \big(1 - \dfrac{1}{n}\big)^{n-k}$.
Or $\ln\big(1 - \dfrac{1}{n}\big)^{n-k} = (n-k)\ln\big(1 - \dfrac{1}{n}\big) \underset{n \to +\infty}{\sim} -\dfrac{n-k}{n} \underset{n \to +\infty}{\sim} -1$
et $\lim\limits_{n \to +\infty} \dfrac{n(n-1)\cdots(n-k+1)}{n^k k!} = \dfrac{1}{k!}$
Ainsi $\lim\limits_{n \to +\infty} P(X_n = k) = \dfrac{\mathrm{e}^{-1}}{k!}$
Donc $(X_n)_n$ converge en loi vers une loi de Poisson de paramètre 1.

Exercice 3.19.

1. Soit a un réel strictement positif. Montrer qu'il existe une variable aléatoire réelle de variance égale à a.

2. On suppose dans cette question que $M = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ est la matrice de variance-covariance d'un couple de variables aléatoires discrètes (X, Y). Montrer que $a \geqslant 0, b \geqslant 0, c \in \mathbb{R}$ et $ab - c^2 \geqslant 0$.

3. Réciproquement, on suppose que la matrice $M = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ vérifie $a \geqslant 0$, $b \geqslant 0$, $c \in \mathbb{R}$ et $ab - c^2 \geqslant 0$.

On suppose $ab = 0$. Montrer que M est la matrice de variance-covariance d'un vecteur (X, Y).

4. On suppose $ab \neq 0$. On pose

$$\sigma_1 = \sqrt{a}, \sigma_2 = \sqrt{b} \text{ et } \rho = \frac{c}{\sqrt{ab}}$$

Soit (U, V) un couple de variables aléatoires discrètes, avec U et V indépendantes et possédant chacune une variance égale à 1.
Montrer que $(\sigma_1 U, \sigma_2 \rho U + \sigma_2 \sqrt{1 - \rho^2} V)$ a pour matrice de variance-covariance M.

Solution :

1. Par exemple, si X suit la loi normale $\mathcal{N}(0, 1)$, alors $\sqrt{a}\, X$ suit la loi normale $\mathcal{N}(0, a)$.

2. La variance étant positive, on a $a \geqslant 0, c \geqslant 0$ et $ab - c^2 \geqslant 0$ n'est autre que l'inégalité de Cauchy-Schwarz.

3. On suppose $ab = 0$, alors $c^2 \leqslant 0 \Rightarrow c = 0$ et $M = \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix}$ par exemple.

Auquel cas, A est la matrice de variance-covariance de $(C, \sqrt{b}X)$, où C est une variable aléatoire constante.

4. $V(X) = \sigma_1^2, V(Y) = \sigma_2^2 \rho^2 + \sigma_2^2(1 - \rho^2) = \sigma_2^2$

$$\operatorname{Cov}(X, Y) = E(\sigma_1 U \times (\sigma_2 \rho U + \sigma_2 \sqrt{1 - \rho^2} V) \\ -\sigma_1 E(U) \times E(\sigma_2 \rho U + \sigma_2 \sqrt{1 - \rho^2} V)$$

Après calculs

$$\operatorname{Cov}(X, Y) = \sqrt{ab} \frac{c}{\sqrt{ab}} = c$$

On vient de montrer dans le cas 2×2 que toute matrice de variance covariance est une matrice symétrique réelle positive et réciproquement que toute matrice symétrique réelle positive est une matrice de variance covariance.