3

PROBABILITÉS

Exercice 3.1.

Dans tout cet exercice, les variables aléatoires sont supposées à valeurs dans $\mathbb{R}^+$ et à densité continue sur $\mathbb{R}_+^*$.

1. Soit X une variable aléatoire telle que pour tout $x \in \mathbb{R}_+^*$, $P(X > x) > 0$.

Justifier, pour tout $x > 0$, l'existence de :

$$\varphi(x) = \lim_{h \to 0^+} \left[\frac{1}{h} \times \frac{P\big(x < X \leqslant x + h\big)}{P(X > x)} \right]$$

On appelle *taux de panne* de X la fonction φ ainsi définie sur $\mathbb{R}_+^*$.

2. Soit $\alpha > 0$ et $\beta > 0$ deux réels. Déterminer la constante K pour que la fonction

$$g(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{K}{(1 + \beta\, x)^{\alpha+1}} & \text{si } x > 0 \end{cases}$$

soit une densité de probabilité.

Calculer le taux de panne d'une variable aléatoire Y de densité g.

3. Déterminer les lois des variables aléatoires ayant un taux de panne constant.

4. Déterminer la (les) loi(s) (densité et fonction de répartition) d'une variable aléatoire W admettant pour taux de panne la fonction :

$$h(t) = a\, \lambda^a\, t^{a-1}, \text{ pour } t > 0,$$

où a et λ sont deux paramètres strictement positifs.

Que retrouve-t-on pour $a = 1$?

Solution :

1. On a $P(X > x) \neq 0$ et en notant F la fonction de répartition de X :

$$\frac{1}{h}\times\frac{P\big(x<X\leqslant x+h\big)}{P(X>x)}=\frac{1}{1-F(x)}\times\frac{F(x+h)-F(x)}{h}$$

Comme F est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$, φ est bien définie sur $\mathbb{R}_+^*$, avec :

$$\varphi(x)=\frac{1}{1-F(x)}\times f(x)$$

2. $\star$ Pour $K\geqslant 0$, la fonction g est définie sur $\mathbb{R}$, continue sur $\mathbb{R}^*$ et admet une limite à gauche et une limite à droite en 0.

$\star$ On a : $\displaystyle\int_0^{+\infty}\frac{dx}{(1+\beta x)^{\alpha+1}}=\Big[-\frac{1}{\alpha\beta}(1+\beta x)^{-\alpha}\Big]_0^{\to+\infty}=\frac{1}{\alpha\beta}$

Ainsi g est une densité de probabilité si et seulement si $K=\alpha\beta$.

En notant alors G la fonction de répartition associée, le calcul précédent donne :

$$\forall\, x\geqslant 0, G(x)=1-\frac{1}{(1+\beta x)^{\alpha}}$$

et

$$\varphi(x)=\frac{g(x)}{1-G(x)}=\frac{\alpha\beta}{1+\beta x}$$

3. Si φ est constante, cette constante ne peut être nulle (sinon f serait la fonction nulle sur $\mathbb{R}_+^*$, ce qui n'est pas raisonnable pour une densité de variable aléatoire à valeurs dans $\mathbb{R}^+$). En notant $\lambda>0$ cette constante, on a donc :

$$\forall\, x>0, \frac{F'(x)}{1-F(x)}=\lambda \text{ , soit } -\ln\big(1-F(x)\big)+\ln\big(1-F(0)\big)=\lambda x$$

et comme $F(0)=0$:

$$\forall\, x>0, F(x)=1-\mathrm{e}^{-\lambda x} \text{ et } X\hookrightarrow\mathcal{E}(\lambda)$$

4. On a, pour $t>0$: $\dfrac{F_W'(t)}{1-F_W(t)}=h(t)=a\,\lambda^a\,t^{a-1}$, d'où :

$$-\ln\big(1-F_W(t)\big)+\ln\big(1-F_W(0)\big)=\lambda^a t^a$$

et en intégrant sur $[0,x]$:

$$\forall\, x\geqslant 0, F_W(x)=1-\mathrm{e}^{-(\lambda x)^a}$$

Par dérivation,une densité de W est donc :

$$\forall\, x>0, f_W(x)=a\lambda^a x^{a-1}\mathrm{e}^{-(\lambda x)^a}\text{, sinon } f_W(x)=0$$

Si $a=1$, on retrouve $h(t)=\lambda$ et pour $x\geqslant 0$, $F_W(x)=1-\mathrm{e}^{-\lambda x}$.

La variable W suit donc la loi $\mathcal{E}(\lambda)$.

Exercice 3.2.

Soit N un entier naturel non nul. On dispose d'un sac contenant N jetons numérotés de 1 à N dans lequel on peut effectuer une succession de tirages **avec remise** d'un jeton en notant, à chaque fois, le numéro obtenu.

Pour tout entier naturel n non nul, on note T_n le nombre (aléatoire) de numéros distincts obtenus au cours des n premiers tirages.

1. Soit n un entier naturel non nul.
   a) Quelles sont les valeurs prises par T_n ?
   b) Calculer $P([T_n=1])$ et $P([T_n=n])$.
   c) Déterminer $P([T_n=2])$.

2. Soit (k,n) un couple d'entiers naturels non nuls avec $1 \leqslant k \leqslant N$.
Déterminer une relation entre $P([T_{n+1}=k])$, $P([T_n=k])$ et $P([T_n=k-1])$.

3. Pour tout entier naturel n non nul, on considère le polynôme :

$$G_n = \sum_{k=1}^{N} P([T_n=k])X^k$$

a) Prouver l'égalité :

$$G_{n+1} = \frac{1}{N}(X-X^2)G'_n + XG_n$$

b) Pour tout entier naturel n non nul, en reliant l'espérance $E(T_n)$ à G_n, exprimer $E(T_{n+1})$ à l'aide de $E(T_n)$, N et n, puis déterminer $E(T_n)$ en fonction de N et n.

c) Déterminer $\lim\limits_{N\to\infty} \dfrac{E(T_N)}{N}$.

Solution :

1. a) Si $n \leqslant N$, alors $T_n(\Omega) = [\![1,n]\!]$ et si $n \geqslant N$, $T_n(\Omega) = [\![1,N]\!]$, ce que l'on peut écrire :

$$T_n(\Omega) = [\![1, \min(n,N)]\!]$$

b) $\star$ Il y a N^n listes de tirages possibles, toutes équiprobables et N listes pour lesquelles $[T_n=1]$ est réalisé (obtenir toujours le même numéro), soit :

$$P(T_n=1) = \frac{1}{N^{n-1}}$$

$\star$ Si $n > N$, l'événement $[T_n=n]$ est impossible, et si $n \leqslant N$ les listes réalisant $[T_n=n]$ sont les arrangements de n éléments pris parmi les N éléments présents, soit :

$$P([T_n=n]) = \frac{\mathrm{A}_N^n}{N^n}$$

c) Les listes réalisant $(T_n=2)$ sont constituées ainsi :
$\to$ on choisit les deux éléments a et b obtenus parmi les N de $\dbinom{N}{2}$ façons ;
$\to$ il existe 2^n mots de longueur n ne contenant pas d'autres lettres que les lettres a et b, et il faut exclure les deux mots $a\ldots a$ et $b\ldots b$ qui réalisent $[T_n=1]$ et pas $[T_n=2]$. Ainsi :

$$P([T_n=2]) = \frac{\binom{N}{2}(2^n-2)}{N^n}$$

2. Si on réalise $(T_n=k)$, alors au rang précédent T_{n-1} n'a pu prendre que les valeurs $k-1$ et/ou k. Dans le premier cas on obtient au n-i'eme tirage un nouveau numéro, parmi les $N-(k-1)$ numéros non encore obtenus, et dans le second cas, on obtient au $n^{\text{ème}}$ tirage, un des k numéros déjà obtenus. De manière plus formelle :

$$P(T_{n+1}=k) = P(T_n=k)P_{(T_n=k)}(T_{n+1}=k) + P(T_n=k-1)P_{(T_n=k-1)}(T_{n+1}=k)$$

Soit, compte tenu de ce que nous venons de dire :

$$P(T_{n+1}=k) = \frac{k}{N}P(T_n=k) + \frac{N-(k-1)}{N}P(T_n=k-1)$$

3. a) On a $G_n = \sum\limits_{k=1}^{N} P(T_n=k)X^k$, donc $G'_n = \sum\limits_{k=1}^{N} kP(T_n=k)X^{k-1}$.

D'où :

$$\begin{aligned} G_{n+1} &= \sum_{k=1}^{N} P(T_{n+1}=k)X^k \\ &= \sum_{k=1}^{N} \frac{k}{N} P(T_n=k)X^k + \sum_{k=1}^{N} \frac{N-(k-1)}{N} P(T_n=k-1)X^k \\ &= \frac{X}{N}\sum_{k=1}^{N} kP(T_n=k)X^{k-1} + \frac{1}{N}\sum_{j=0}^{N-1}(N-j)P(T_n=j)X^{j+1} \\ &= \frac{X}{N}G'_n + \sum_{j=1}^{N} P(T_n=j)X^{j+1} - \frac{1}{N}\sum_{j=1}^{N} jP(T_n=j)X^{j+1} \end{aligned}$$

soit finalement :

$$G_{n+1} = \frac{X}{N}G'_n + XG_n - \frac{X^2}{N}G'_n$$

b) On remarque que $E(T_n) = G'_n(1)$, donc en dérivant la relation précédente, il vient :

$$G'_{n+1} = \frac{1}{N}(1-2X)G'_n + \frac{1}{N}(X-X^2)G''_n + G_n + XG'_n$$

ce qui donne en 1, en sachant que $G_n(1) = 1$:

$$E(T_{n+1}) = -\frac{1}{N}E(T_n) + 1 + E(T_n)$$

c'est-à-dire :

$$E(T_{n+1}) = \big(1 - \frac{1}{N}\big)E(T_n) + 1$$

Le traitement d'une telle suite arithmético-géométrique est standard : son point fixe est N, d'où l'on déduit : $E(T_n) - N = \big(1 - \frac{1}{N}\big)^n (E(T_0) - N)$, soit :

$$\forall\, n \in \mathbb{N}, E(T_n) = N\big[1 - \big(1 - \frac{1}{N}\big)^n\big]$$

c) $\dfrac{E(T_N)}{N} = 1 - \big(1 - \frac{1}{N}\big)^N = 1 - \mathrm{e}^{N\ln(1-\frac{1}{N})}$.

Comme on sait que $\ln(1-u) \underset{(0)}{\sim} -u$, il vient :

$$\lim_{N\to\infty} \frac{E(T_N)}{N} = 1 - \mathrm{e}^{-1}$$

Exercice 3.3.

Soit $n > 0$ un entier naturel fixé.

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

$$P_1(x) = x^{n-1} \text{ et pour } j \geqslant 1,\ P_{j+1}(x) = P_j(x) + \frac{1-x}{n}P'_j(x)$$

Montrer par récurrence que : $P_j(x) = \displaystyle\sum_{i=0}^{n-1} \binom{n-1}{i} (1 - \frac{i}{n})^{j-1}(x-1)^i$.

2. On dispose de n boîtes dans lesquelles on lance au hasard l'une après l'autre des billes. Les résultats des lancers sont indépendants les uns des autres. Pour $j \geqslant 1$, on note X_j le nombre de boîtes non vides après les j premiers lancers.

a) i) Quelles valeurs peut prendre la variable aléatoire X_j ?

ii) Déterminer les lois de X_j pour $j \in \{1, 2\}$.

iii) Déterminer les probabilités conditionnelles $P_{(X_j=i)}(X_{j+1} = k)$ pour $1 \leqslant i \leqslant n$.

iv) En déduire l'expression de $P(X_{j+1}=k)$, en fonction des $P(X_j=i)$, pour $1 \leqslant i \leqslant n$.

b) Soit la fonction $G_j(x) = \sum\limits_{k=1}^{n} P(X_j=k)x^{n-k}$.

Vérifier que les suites $(G_j(x))_{j\geqslant 1}$ et $(P_j(x))_{j\geqslant 1}$ sont égales.

c) En déduire : $P(X_j=n) = \sum\limits_{i=0}^{n-1}(-1)^i\binom{n-1}{i}(1-\frac{i}{n})^{j-1}$.

d) Montrer que pour tout $j \in [\![1, n-1]\!]$, on a :

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

Solution :

1. $\star$ Pour $j=1$, on écrit :

$$\sum_{i=0}^{n-1}(x-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i} = (1+(x-1))^{n-1} = x^{n-1} = P_1(x)$$

$\star$ Supposons le résultat acquis pour un certain rang j, alors en dérivant :

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

et en remplaçant :

$$\begin{aligned}
P_j(x) + \frac{1-x}{n}P_j'(x) &= 1 + \sum_{i=1}^{n-1}(x-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i} \\
&\qquad + i(x-1)^{i-1}(1-\frac{i}{n})^{j-1}\binom{n-1}{i} \\
&= 1 + \sum_{i=1}^{n-1}(x-1)^i(1-\frac{i}{n})^{j-1}\binom{n-1}{i}(1-\frac{i}{n}) \\
&= \sum_{i=0}^{n-1}(x-1)^i(1-\frac{i}{n})^{j}\binom{n-1}{i} = P_{j+1}(x)
\end{aligned}$$

ce qui montre que le résultat est valide au rang $j+1$. On conclut par le principe de récurrence.

2. a) i) X_j prend ses valeurs entre 1 et $\min(j, n)$.

ii) Sans problème :

$$P(X_1=1) = 1 \text{ et } P(X_2=1) = \frac{1}{n},\ P(X_2=2) = \frac{n-1}{n}.$$

iii) En suivant l'évolution du nombre de boîtes non vides entre la fin du $j^{\text{ème}}$ et la fin du $(j+1)^{\text{ème}}$ tirage selon le résultat obtenu au $(j+1)^{\text{ème}}$ tirage :

$P_{(X_j=k)}(X_{j+1}=k) = \frac{k}{n}$

$P_{(X_j=k-1)}(X_{j+1}=k) = \frac{n-k+1}{n}$

si $i \notin \{k-1, k\}$, $P_{(X_j=i)}(X_{j+1}=k) = 0$.

iv) La formule des probabilités totales se réduit alors à :

$$P(X_{j+1}=k) = \frac{k}{n}P(X_j=k) + \frac{n-k+1}{n}P(X_j=k-1)$$

b) On a $G_{j+1}(x) = \sum\limits_{k=1}^{n} P(X_{j+1}=k)x^{n-k}$, soit en remplaçant :

$$\begin{aligned}G_{j+1}(x) &= \sum_{k=1}^{n} \frac{k}{n} P(X_j = k)x^{n-k} + \sum_{k=1}^{n} \frac{n-k+1}{n} P(X_j = k-1)x^{n-k} \\ &= \sum_{k=1}^{n} \frac{k-n+n}{n} P(X_j = k)x^{n-k} + \sum_{k=1}^{n} \frac{n-k+1}{n} P(X_j = k-1)x^{n-k} \\ &= \sum_{k=1}^{n} P(X_j = k)x^{n-k} + \frac{1-x}{n} \sum_{k=1}^{n} (n-k)P(X_j = k)x^{n-k-1}\end{aligned}$$

$$G_{j+1}(x) = G_j(x) + \frac{1-x}{n} G'_j(x)$$

Comme $G_1(x) = \sum_{k=1}^{n} P(X_1 = k)x^{n-k} = P(X_1 = 1)x^{n-1} = x^{n-1}$, les suites (P_j) et (G_j) ont les mêmes termes initiaux et vérifient les mêmes relations de récurrence, donc sont égales :

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

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

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

d) Pour j tel que $1 \leqslant j < n$, l'événement $P(X_j = n)$ est impossible, donc de probabilité nulle, ce qui s'écrit :

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

Exercice 3.4.

Une urne contient des boules noires et blanches, la proportion de boules noires étant p, avec $0 < p < 1$. On effectue une suite de tirages d'une boule avec, à chaque fois, remise de la boule obtenue avant le tirage suivant.

I. On note N le rang aléatoire où l'on obtient pour la première fois une boule noire et B le rang aléatoire où l'on obtient pour la première fois une boule blanche.

1. Donner les lois, espérances et variances des variables aléatoires N et B.

2. N et B sont-elles indépendantes ?

II. On note X la longueur de la première suite de boules de même couleur et Y la longueur de la deuxième suite de boules de même couleur.

Ainsi, l'événement $(X = 2, Y = 3)$ est réalisé si et seulement si on a tiré $(n, n, b, b, b, n, \ldots)$ ou $(b, b, n, n, n, b, \ldots)$, où n désigne le tirage d'une boule noire et b celui d'une boule blanche.

1. Quelle est la loi du couple (X, Y) ?

2. Déterminer la loi de X et son espérance $E(X)$. Montrer que $E(X) \geqslant 2$.

3. Déterminer la loi de Y, son espérance et sa variance.

4. Calculer la probabilité de l'événement $(X = Y)$.

5. Montrer que X et Y sont indépendantes si et seulement si $p = 0{,}5$.

6. Déterminer la loi de $X+Y$. On distinguera les cas $p \neq 0{,}5$ et $p = 0{,}5$.

Solution :

I) 1. Les tirages ayant lieu avec remise, on sait que :

$$N \hookrightarrow \mathcal{G}(p) \text{ et } B \hookrightarrow \mathcal{G}(q)$$

et donc :

$$E(N) = \frac{1}{p},\ E(B) = \frac{1}{q} \text{ et } V(N) = \frac{q}{p^2},\ V(B) = \frac{p}{q^2}$$

2. $P([N=1] \cap [B=1]) = 0 \neq P(N=1)P(B=1)$, donc N et B ne sont pas indépendantes.

II) 1. Réaliser $(X = i) \cap (Y = j)$, c'est obtenir d'abord i boules noires, puis j boules blanches et enfin une boule noire, ou bien obtenir d'abord i boules blanches, puis j boules noires et enfin une boule blanche. Donc par indépendance des résultats des différents tirages :

$$P([X=i] \cap [Y=j]) = p^i q^j p + q^i p^j q = p^{i+1}q^j + q^{i+1}p^j$$

2. La loi marginale de X est donc donnée par :

$$\forall\, i \in \mathbb{N}^*, P(X=i) = \sum_{j=1}^{\infty}(p^{i+1}q^j + q^{i+1}p^j) = p^{i+1}\sum_{j=1}^{\infty} q^j + q^{i+1}\sum_{j=1}^{\infty} p^j$$
$$= p^{i+1}\times\frac{q}{1-q} + q^{i+1}\times\frac{p}{1-p}$$
$$\forall\, i \in \mathbb{N}^*, P(X=i) = qp^i + pq^i$$

Notons que $\sum\limits_{i=1}^{\infty} P(X=i) = q\times\frac{p}{1-p} + p\times\frac{q}{1-q} = p+q = 1$ et X est bien une variable aléatoire à valeurs dans $\mathbb{N}^*$.

La convergence des série rencontrées étant claire, on a :

$$E(X) = \sum_{i=1}^{\infty} iP(X=i) = qp\sum_{i=1}^{\infty} ip^{i-1} + pq\sum_{i=1}^{\infty} iq^{i-1}$$
$$E(X) = qp\frac{1}{(1-p)^2} + pq\frac{1}{(1-q)^2}$$

soit :

$$E(X) = \frac{p}{q} + \frac{q}{p} = \frac{p^2+q^2}{pq}$$

Comme $p^2 + q^2 - 2pq = (p-q)^2 \geqslant 0$, on a $p^2+q^2 \geqslant 2pq$ et $E(X) \geqslant 2$ (avec égalité seulement pour $p = q = \frac{1}{2}$).

3. De la même façon, on obtient la loi marginale de Y, par sommation :

$$\forall\, j \in \mathbb{N}^*, P(Y=j) = \sum_{i=1}^{\infty}(p^{i+1}q^j + q^{i+1}p^j) = q^j p^2\sum_{i=1}^{\infty} p^{i-1} + p^j q^2\sum_{i=1}^{\infty} q^{i-1}$$
$$= q^j p^2\times\frac{1}{1-p} + p^j q^2\times\frac{1}{1-q}$$
$$P(Y=j) = p^2 q^{j-1} + q^2 p^{j-1}$$

(on peut encore vérifier que $\sum\limits_{j=1}^{\infty} P(Y=j) = p^2\times\frac{1}{1-q} + q^2\times\frac{1}{1-p} = 1$)

La convergence ne pose toujours pas de problème, et :

$$E(Y) = p^2\sum_{j=1}^{\infty} jq^{j-1} + q^2\sum_{j=1}^{\infty} jp^{j-1}$$
$$= p^2\times\frac{1}{(1-q)^2} + q^2\times\frac{1}{(1-p)^2} = 1+1=2$$

Enfin :

$$E(Y(Y-1)) = p^2q\sum_{j=2}^{\infty} j(j-1)q^{j-2} + q^2p\sum_{j=2}^{\infty} j(j-1)p^{j-2}$$
$$= p^2q\times\frac{2}{(1-q)^3} + q^2p\times\frac{2}{(1-p)^3} = 2\big(\frac{q}{p}+\frac{p}{q}\big)$$

D'où par la formule de Koenig :

$$V(Y) = E(Y(Y-1)) + E(Y) - E(Y)^2 = 2\big(\frac{q}{p}+\frac{p}{q}-1\big)$$

4. $P(X=Y) = \sum_{i=1}^{\infty} P((X=i)\cap(Y=i)) = \sum_{i=1}^{\infty}(p+q)(pq)^i = \sum_{i=1}^{\infty}(pq)^i$

$$P(X=Y) = \frac{pq}{1-pq}$$

5. $\star$ Si $p=\frac{1}{2}$, on a :

$$P((X=i)(Y=j)) = \big(\frac{1}{2}\big)^{i+j} = \big(\frac{1}{2}\big)^i\big(\frac{1}{2}\big)^j = P(X=i)P(Y=j)$$

Donc X et Y sont indépendantes.

$\star$ Si $p\neq\frac{1}{2}$, on a :

$P(X=1)=2pq$, $P(Y=1)=p^2+q^2$, $P((X=1)\cap(Y=1)) = p^2q+q^2p = pq$

et :

$pq = 2pq(p^2+q^2) \iff p^2+q^2=\frac{1}{2} \iff 2p^2-2p+\frac{1}{2}=0 \iff (p-\frac{1}{2})^2=0$

Donc, pour $p\neq\frac{1}{2}$, les événements $(X=1)$ et $(Y=1)$ ne sont pas indépendants et *a fortiori* les variables aléatoires X et Y ne sont pas indépendantes.

6. $X+Y$ prend ses valeurs dans $[\![2,+\infty[\![$, et pour $k\geqslant 2$:

$$P(X+Y=k) = \sum_{i=1}^{k-1} P((X=i)\cap(Y=k-i)) = \sum_{i=1}^{k-1}(p^{i+1}q^{k-i}+q^{i+1}p^{k-i})$$
$$= pq^k\sum_{i=1}^{k-1}\big(\frac{p}{q}\big)^i + qp^k\sum_{i=1}^{k-1}\big(\frac{q}{p}\big)^i$$

$\star$ Si $p=q=\frac{1}{2}$, alors $\frac{p}{q}=\frac{q}{p}=1$ et $P(X+Y=k)=2(k-1)\big(\frac{1}{2}\big)^{k+1}$

$\star$ Si $p\neq\frac{1}{2}$, alors $\frac{p}{q}\neq 1$ et $\frac{q}{p}\neq 1$, l'identité géométrique donnant :

$$P(X+Y=k) = pq^k\frac{\frac{p}{q}-\big(\frac{p}{q}\big)^k}{1-\frac{p}{q}} + qp^k\frac{\frac{q}{p}-\big(\frac{q}{p}\big)^k}{1-\frac{q}{p}}$$

Ce que l'on peut aussi écrire :

$$P(X+Y=k) = pq\frac{(p^k-q^k)+pq(p^{k-2}-q^{k-2})}{p-q} = pq\frac{p^{k-1}-q^{k-1}}{p-q}$$

Exercice 3.5.

Soit $p\in\mathbb{N}^*$ et $s=(s_1,\ldots,s_p)\in\{0,1\}^p$.

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre $\alpha \in]0,1[$, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On pose pour $n \in \mathbb{N}$:

$$B_n = \{\omega \in \Omega / X_{np+1}(\omega) = s_1, X_{np+2}(\omega) = s_2, \ldots, X_{(n+1)p}(\omega) = s_p\}$$

Si E est un événement, on désignera par E^c son complémentaire dans l'univers Ω.

1. Justifier l'assertion :

$\omega \in \bigcap\limits_{k=0}^{\infty} (\bigcup\limits_{i=k}^{\infty} B_i) \implies s$ apparaît une infinité de fois dans $(X_1(\omega), \ldots, X_n(\omega), \ldots)$

2. Montrer que la suite $(B_n)_{n\geqslant 0}$ est formée d'événements mutuellement indépendants.

3. Déterminer l'événement $\Big(\bigcap\limits_{k=0}^{\infty} (\bigcup\limits_{i=k}^{\infty} B_i)\Big)^c$.

4. Montrer que pour tout entier naturel k, on a : $P\Big(\bigcap\limits_{i=k}^{\infty} B_i^c \Big) = 0$.

5. En déduire que $P\Big(\bigcap\limits_{k=0}^{\infty} (\bigcup\limits_{i=k}^{\infty} B_i)\Big) = 1$ et interpréter ce résultat.

Solution :

1. $\omega \in \bigcap\limits_{k=0}^{\infty} (\bigcup\limits_{i=k}^{\infty} B_i) \implies \forall\, k \in \mathbb{N}, \omega \in \bigcup\limits_{i=k}^{\infty} B_i \implies \forall\, k \in \mathbb{N}, \exists\, i \geqslant k, \omega \in B_i$,
et donc ω appartient à une infinité de B_i (sinon à partir d'un certain rang ω n'appartiendrait plus à aucun B_i).

2. B_n dépend des variables aléatoires $X_{np+1}, X_{np+2}, \ldots, X_{(n+1)p}$. Donc des événements $B_{n_1}, B_{n_2}, \ldots$ (avec les indices n_i deux à deux distincts) utilisent chacun des variables autres que les variables utilisées par les autres, on sait alors que l'indépendance des variables aléatoires $X_i, i \in \mathbb{N}^*$ donne l'indépendance des événements $B_n, n \in \mathbb{N}^*$.

3. $$\begin{aligned}\omega \in \Big(\bigcap_{k=0}^{\infty} (\bigcup_{i=k}^{\infty} B_i)\Big)^c &\iff \omega \notin \bigcap_{k=0}^{\infty} (\bigcup_{i=k}^{\infty} B_i) \iff \exists\, k \in \mathbb{N}, \omega \notin \bigcup_{i=k}^{\infty} B_i \\ &\iff \exists\, k \in \mathbb{N}, \forall\, i \geqslant k, \omega \notin B_i \iff \exists\, k \in \mathbb{N}, \omega \in \bigcap_{i=k}^{\infty} B_i^c \\ &\iff \omega \in \bigcup_{k=0}^{\infty} (\bigcap_{i=k}^{\infty} B_i^c)\end{aligned}$$

(on peut aussi appliquer directement les lois de Augustus de Morgan).

4. $\star$ Par indépendance des variables $X_1, \ldots, X_p$, on a :

$$P(B_0) = P(X_1 = s_1, X_2 = s_2, \ldots, X_p = s_p) = \alpha^r (1-\alpha)^{p-r},$$

où r est le nombre de 1 de la séquence s. Le résultat est évidemment le même pour n'importe quel événement B_i.

$\star$ Donc $P(B_i^c) = 1 - \alpha^r(1-\alpha)^{p-r}$ et par indépendance :

$$P\big(\bigcap_{i=k}^{k+j} B_i^c\big) = [1-\alpha^r(1-\alpha)^{p-r}]^{j+1}$$

Comme $0 \leqslant 1-\alpha^r(1-\alpha)^{p-r} < 1$, on a $\lim\limits_{j\to+\infty} P\big(\bigcap\limits_{i=k}^{k+j} B_i^c\big) = 0$. Par le théorème de limite monotone , on a donc :

$$P\big(\bigcap_{i=k}^{+\infty} B_i^c\big) = 0$$

5. La probabilité d'une réunion étant majorée par la somme des probabilités, on a :

$$P\big(\bigcup_{k=0}^{\infty}\bigcap_{i=k}^{\infty} B_i^c\big) \leqslant \sum_{k=0}^{\infty} P\big(\bigcap_{i=k}^{\infty} B_i^c\big) = 0$$

Donc :

$$P\big(\bigcap_{k=0}^{\infty}\bigcup_{i=k}^{\infty} B_i\big) = 1$$

Ainsi, dans la succession des expériences de Bernoulli de ce problème, la liste s de résultats consécutifs, à partir d'un rang de la forme $kp+1$, apparaît presque sûrement une infinité de fois.

Exercice 3.6.

Un sac contient n billes numérotées de 1 à n. On tire une bille au hasard, on note son numéro et on la remet dans le sac.

On appelle X la variable aléatoire qui prend pour valeur ce numéro. Lorsque ce numéro est k, on tire sans remise k billes que l'on distribue au hasard dans p boîtes $B_1, \dots, B_p$. On désigne par Y_i la variable aléatoire égale au nombre de billes reçues par la boîte B_i ($i \in \{1, \dots, p\}$).

1. Déterminer la loi du couple (X, Y_i) pour $i \in \{1, \dots, p\}$.

2. En déduire, pour tout $i \in \{1, \dots, p\}$, la loi de Y_i et calculer son espérance.

3. Déterminer l'espérance de la variable aléatoire $\dfrac{Y_i}{X}$.

4. Quelle est la loi du vecteur aléatoire $(Y_1, \dots, Y_p)$?

Solution :

1. La variable X suit la loi uniforme sur $[\![1, n]\!]$ et Y_i est à valeurs dans $[\![0, n]\!]$. Soit donc $(k, \ell) \in [\![1, n]\!] \times [\![0, n]\!]$, on a

$\star$ Si $\ell > k$, il est clair que $P((X = k) \cap (Y_i = \ell)) = 0$

$\star$ Si $\ell \leqslant k$, $P((X = k) \cap (Y_i = \ell)) = P(X = k)P_{(X=k)}(Y_i = \ell)$

Or $P(X = k) = \dfrac{1}{n}$ et $P_{(X=k)}(Y_i = \ell) = \dbinom{k}{\ell}\big(\dfrac{1}{p}\big)^{\ell}\big(1 - \dfrac{1}{p}\big)^{k-\ell}$ (succession de k expériences indépendantes à deux issues : tomber dans B_i ou tomber dans une autre boîte). Ainsi :

$$P((X = k) \cap (Y_i = \ell)) = \frac{1}{n}\binom{k}{\ell}\big(\frac{1}{p}\big)^{\ell}\big(1 - \frac{1}{p}\big)^{k-\ell}$$

2. Pour $\ell \in [\![0, n]\!]$, il vient par sommation :

$$P(Y_i=\ell)=\sum_{k=1}^{n}P((X=k)\cap(Y_i=\ell))=\sum_{k=\ell}^{n}P((X=k)\cap(Y_i=\ell))$$

$$P(Y_i=\ell)=\frac{1}{n}\sum_{k=\ell}^{n}\binom{k}{\ell}\Big(\frac{1}{p}\Big)^{\ell}\Big(1-\frac{1}{p}\Big)^{k-\ell}$$

Par conditionnement : $E(Y_i)=\sum_{k=1}^{n}E(Y_i/X=k)P(X=k)$

Or la loi conditionnelle de Y_i, conditionnée par la réalisation de l'événement $(X=k)$ est la loi $\mathcal{B}(k,\frac{1}{p})$ d'espérance $\frac{k}{p}$. Ainsi :

$$E(Y_i)=\sum_{k=1}^{n}\frac{k}{np}=\frac{n(n+1)}{2np}=\frac{n+1}{2p}$$

3. On a :

$$E\Big(\frac{Y_i}{X}\Big)=\sum_{k=1}^{n}E\Big(\frac{Y_i}{X}/X=k\Big)P(X=k)=\sum_{k=1}^{n}E\Big(\frac{Y_i}{k}/X=k\Big)P(X=k)$$
$$=\sum_{k=1}^{n}\frac{1}{k}E\big(Y_i/X=k\big)P(X=k)=\sum_{k=1}^{n}\frac{1}{k}\times\frac{k}{np}=\frac{1}{p}$$
$$E\Big(\frac{Y_i}{X}\Big)=\frac{1}{p}$$

4. On remarque que $Y_1+\cdots+Y_p=X$, donc :

$\star$ Si $\ell_1+\cdots+\ell_p\notin[\![1,n]\!]$, $P((Y_1=\ell_1)\cap\ldots\cap(Y_p=\ell_p))=0$

$\star$ Si $\ell_1+\cdots+\ell_p\in[\![1,n]\!]$,

$\alpha=P((Y_1=\ell_1)\cap\ldots\cap(Y_{p-1}=\ell_{p-1})\cap(Y_p=\ell_p))$

$\quad=P((Y_1=\ell_1)\cap\ldots\cap(Y_{p-1}=\ell_{p-1})\cap(X=\ell_1+\cdots+\ell_p))$

$\quad=P((Y_1=\ell_1)\cap\ldots\cap(Y_{p-1}=\ell_{p-1})/X=\ell_1+\cdots+\ell_p)P(X=\ell_1+\cdots+\ell_p)$

Sachant que $(X=\ell_1+\cdots+\ell_p)$ est réalisé, nous sommes en présence d'une distribution multinomiale (urne à p catégories) ; les proportions valant $\frac{1}{p},\frac{1}{p},\ldots,\frac{1}{p}$ et donc, facilement :

$$\alpha=\frac{(\ell_1+\cdots+\ell_p)!}{\ell_1!\ell_2!\ldots\ell_p!}\times\Big(\frac{1}{p}\Big)^{\ell_1}\ldots\Big(\frac{1}{p}\Big)^{\ell_p}\times\frac{1}{n}$$

$$P((Y_1=\ell_1)\cap\ldots\cap(Y_p=\ell_p))=\frac{(\ell_1+\cdots+\ell_p)!}{\ell_1!\ell_2!\ldots\ell_p!}\times\Big(\frac{1}{p}\Big)^{\ell_1+\cdots+\ell_p}\times\frac{1}{n}$$

Exercice 3.7.

Un point se déplace dans un plan muni d'un repère orthonormé. Il part de l'origine O des coordonnées à l'instant 0.

Si à l'instant $t=k, k\in\mathbb{N}$ il se situe au point de coordonnées (X_k,Y_k), alors à l'instant $t=k+1$ il se trouve au point de coordonnées (X_{k+1},Y_{k+1}) de sorte que $A_{k+1}=X_{k+1}-X_k$ et $B_{k+1}=Y_{k+1}-Y_k$ suivent la loi normale $\mathcal{N}(0,1)$.

On suppose les A_i et les B_j mutuellement indépendantes.

Soit $\Phi_{m,\sigma}$ la fonction de répartition de la loi normale $\mathcal{N}(m,\sigma^2)$ et $\varphi_{m,\sigma}$ une densité de cette loi.

On rappelle que $\Gamma(1/2)=\sqrt{\pi}$.

1. a) Quelle est la loi suivie par X_n ?

   b) Soit M_n le point de coordonnées (X_n,Y_n).

Exprimer, à l'aide de la fonction $\Phi_{0,1}$, la probabilité qu'à l'instant n le point M_n se trouve dans le carré $C = [-1, 1]^2$.

2. Soit D_n la distance de M_n à l'origine : $D_n = \sqrt{X_n^2 + Y_n^2}$.

a) Reconnaître la loi de X_n^2, puis celle de D_n^2. En déduire la loi de D_n et calculer son espérance.

b) Quelle est la probabilité qu'à l'instant n le point se trouve dans le disque de centre O et de rayon 1 ?

Solution :

1. a) Par télescopage $X_n = \sum\limits_{k=1}^{n} A_k$. Les variables $A_i, i \in \mathbb{N}$ étant indépendantes et de même loi $\mathcal{N}(0,1)$, on sait que $X_n \hookrightarrow \mathcal{N}(0,n)$.

b) M est dans le carré voulu si et seulement si $-1 \leqslant X_n \leqslant 1$ et $-1 \leqslant Y_n \leqslant 1$, soit par indépendance des variables X_n et Y_n (les variables A_i sont aussi indépendantes des variables B_j) :

$$\begin{aligned} P(M_n \in C) &= P(-1 \leqslant X \leqslant 1)P(-1 \leqslant Y \leqslant 1) = \big(\Phi_{0,\sqrt{n}}(1) - \Phi_{0,\sqrt{n}}(-1)\big)^2 \\ &= \big(2\Phi_{0,\sqrt{n}}(1) - 1\big)^2 \end{aligned}$$

Or $\Phi_{0,\sqrt{n}}(1) = P(X_n \leqslant 1) = P\big(\frac{X_n}{\sqrt{n}} \leqslant \frac{1}{\sqrt{n}}\big) = \Phi_{0,\sqrt{n}}\big(\frac{1}{\sqrt{n}}\big)$, d'où :

$$P(M_n \in C) = \big(2\Phi\big(\frac{1}{\sqrt{n}}\big) - 1\big)^2$$

2. a) X_n^2 prend ses valeurs dans $\mathbb{R}^+$ et pour $x \geqslant 0$:

$$P(X_n^2 \leqslant x) = P(-\sqrt{x} \leqslant X_n \leqslant \sqrt{x}) = 2\Phi_{0,\sqrt{n}}(\sqrt{x})$$

Par dérivation, une densité de X_n^2 sur $\mathbb{R}_+^*$ est la fonction :

$$x \mapsto \frac{1}{\sqrt{x}}\varphi_{0,\sqrt{n}}(\sqrt{x}) = \frac{1}{\sqrt{x}\sqrt{2n\pi}}\mathrm{e}^{-\frac{x}{2n}}$$

Ainsi :

$$\text{pour } x > 0,\ f_{X^2}(x) = \frac{1}{\Gamma(1/2)}\big(\frac{x}{2n}\big)^{\frac{1}{2}-1}\frac{1}{2n}\mathrm{e}^{-\frac{x}{2n}}$$

On reconnaît une densité d'une variable aléatoire qui suit la loi $\Gamma(2n, \frac{1}{2})$.

Comme $D_n^2 = X_n^2 + Y_n^2$, avec X_n^2 et Y_n^2 indépendantes, on sait alors que :

$$D_n^2 \hookrightarrow \Gamma(2n, 1) = \mathcal{E}\big(\frac{1}{2n}\big)$$

de densité sur $\mathbb{R}_+^*$, $f_{D_n^2}(x) = \frac{1}{2n}\mathrm{e}^{-\dfrac{x}{2n}}$

La variable D_n prend ses valeurs dans $\mathbb{R}^+$ et pour $x \geqslant 0$,

$$P(D_n \leqslant x) = P(D_n^2 \leqslant x^2)$$

ce qui donne par dérivation :

$$f_{D_n}(x) = 2xf_{D_n^2}(x^2) = \frac{x}{n}\mathrm{e}^{-\frac{x^2}{2n}}$$

Par le théorème de transfert et sous réserve de convergence (absolue) de l'intégrale :

$$E(D_n) = \int_0^{+\infty} \sqrt{t} \times \frac{1}{2n}\,\mathrm{e}^{-\frac{t}{2n}}\,dt$$

La convergence est claire et le changement de variable $u = \dfrac{t}{2n}$ donne :

$$E(D_n) = \sqrt{2n}\int_0^{+\infty} \sqrt{u}\,\mathrm{e}^{-u}\,du = \sqrt{2n}\Gamma(3/2)$$

Comme $\Gamma(3/2) = \frac{1}{2}\Gamma(1/2)$, il vient finalement :

$$E(D_n) = \sqrt{\frac{n\pi}{2}}$$

b) M_n est dans le disque unité si et seulement si $(D_n^2 \leqslant 1)$ est réalisé. Ainsi, la connaissance de la fonction de répartition de D_n^2 donne :

$$P(D_n \leqslant 1) = 1 - \mathrm{e}^{-\frac{1}{2n}}$$

Exercice 3.8.

Soit a un réel strictement positif.
On considère la fonction f définie sur $\mathbb{R}$ par :

$$f(x) = \begin{cases} \dfrac{1}{x^2}\ \exp\left(a\times\dfrac{x-1}{x}\right) & \text{si } x \in]0,1] \\ 0 & \text{sinon} \end{cases}$$

où exp désigne la fonction exponentielle.

1. Étudier la continuité de f.

2. Déterminer la constante C telle que $C\times f$ soit une fonction densité (on pourra utiliser le changement de variable $u = 1 - \frac{1}{x}$).

3. Soit X une variable aléatoire de densité $C\times f$. Soit Y la variable aléatoire définie par $Y = \dfrac{1}{X} - \left\lfloor \dfrac{1}{X} \right\rfloor$, où $\lfloor\ \rfloor$ désigne la fonction partie entière.

a) Déterminer la loi de probabilité de $\left\lfloor \dfrac{1}{X} \right\rfloor$.

b) Déterminer la fonction de répartition de Y, puis une densité de Y.

Solution :

1. La fonction f est continue sur $\mathbb{R}$ sauf peut-être en 0 et en 1, et :

$\star$ $\lim\limits_{x\to 0^+} f(x) = \lim\limits_{t\to+\infty} t^2\mathrm{e}^{-at}\mathrm{e}^{a} = 0 = f(0) = \lim\limits_{t\to 0^-} f(t)$, donc f est continue au point 0.

$\star$ En revanche $\lim\limits_{t\to 1^+} f(t) = 0 \neq 1 = f(1)$ et f n'est pas continue en 1.

2. L'existence de $\displaystyle\int_0^1 f(x)\,dx$ ne pose aucun problème et le changement de variable $x \mapsto u = 1 - \dfrac{1}{x}$ est de classe $\mathcal{C}^1$ strictement monotone, donc légitime et puisque $du = \dfrac{dx}{x^2}$, il donne :

$$\int_0^1 f(x)\,dx = \int_{-\infty}^0 \mathrm{e}^{au}\,du = \left[\frac{1}{a}\,\mathrm{e}^{au}\right]_{\to-\infty}^0 = \frac{1}{a}$$

Soit :

$$C = \frac{1}{a}$$

3. a) La variable aléatoire X prend ses valeurs dans $]0,1]$, donc $\dfrac{1}{X}$ dans $[1,+\infty[$ et sa partie entière dans $\mathbb{N}^*$.
Pour $k \in \mathbb{N}^*$, en notant F la fonction de répartition de X :

$$P\big(\lfloor \tfrac{1}{X} \rfloor = k\big) = P\big(k \leqslant \tfrac{1}{X} < k+1\big) = P\big(\tfrac{1}{k+1} < X \leqslant \tfrac{1}{k}\big)$$
$$= F\big(\tfrac{1}{k}\big) - F\big(\tfrac{1}{k+1}\big)$$

Or, pour $\alpha \geqslant 0$:

$$F(\alpha) = \frac{1}{a}\int_0^{\alpha} f(x)\,dx = \frac{1}{a}\int_{-\infty}^{1-\frac{1}{\alpha}} \mathrm{e}^{au} du = \mathrm{e}^{a(1-\frac{1}{\alpha})}$$

Donc :

$$P\big(\lfloor \tfrac{1}{X} \rfloor = k\big) = \mathrm{e}^{a(1-k)} - \mathrm{e}^{a(1-(k+1))} = \mathrm{e}^{-ak}(\mathrm{e}^{a} - 1)$$

b) Y prend ses valeurs entre 0 et 1 et pour $y \in [0,1[$, on a par disjonction des cas :

$$P(Y \leqslant y) = \sum_{k=1}^{\infty} P\big((\lfloor \tfrac{1}{X} \rfloor = k) \cap (Y \leqslant y)\big) = \sum_{k=1}^{\infty} P\big(\tfrac{1}{k+y} \leqslant X < \tfrac{1}{k}\big)$$
$$= \sum_{k=1}^{\infty} \big[F\big(\tfrac{1}{k}\big) - F\big(\tfrac{1}{k+y}\big)\big] = \sum_{k=1}^{\infty} \big[\mathrm{e}^{a(1-k)} - \mathrm{e}^{a(1-k-y)}\big]$$
$$= (1-\mathrm{e}^{-ay}) \sum_{k=1}^{\infty} (\mathrm{e}^{-a})^{k-1}$$

soit, par l'identité géométrique :

$$\forall\, y \in [0,1[, F_Y(y) = P(Y \leqslant y) = \frac{1-\mathrm{e}^{-ay}}{1-\mathrm{e}^{-a}}$$

Par dérivation une densité f_Y de Y est, par exemple :

$$\forall\, y \in [0,1], f_Y(y) = \frac{a}{1-\mathrm{e}^{-a}}\,\mathrm{e}^{-ay}\ ;\ f_Y(y) = 0 \text{ sinon}$$

Exercice 3.9.

On effectue une suite de lancers d'une pièce de monnaie. On suppose que les résultats des lancers sont indépendants et qu'à chaque lancer, la pièce donne Face avec la probabilité p $(0 < p < 1)$ et Pile avec la probabilité $q = 1-p$. On s'intéresse au nombre de lancers nécessaires pour obtenir deux Face de suite (c'est-à-dire lors de deux lancers consécutifs). On suppose donné un espace probabilisé, muni d'une probabilité P, modélisant cette expérience.

Pour tout entier $n \geqslant 1$, on note U_n l'événement « on obtient deux Face de suite, pour la première fois, aux lancers numéro n et $n+1$ », et on pose $u_n = P(U_n)$.

Pour tout entier $n \geqslant 2$, on note A_n l'événement « les n premiers lancers ne donnent pas deux Face de suite et le $n^{\text{ème}}$ lancer donne Face », et B_n l'événement « les n premiers lancers ne donnent pas deux Face de suite et le $n^{\text{ème}}$ lancer donne Pile ».
Enfin, on pose $x_n = P(A_n), y_n = P(B_n)$.

1. a) Déterminer $u_1, x_2, y_2, u_2, x_3, y_3, u_3$.

b) Trouver pour $n \geqslant 2$, une relation simple entre x_n et u_n.

c) Pour tout $n \geqslant 2$, déterminer les probabilités conditionnelles :

$$P(A_{n+1}/A_n), P(A_{n+1}/B_n), P(B_{n+1}/A_n), P(B_{n+1}/B_n)$$

d) En déduire, pour tout $n \geqslant 2$, les relations suivantes :

$$\begin{cases} x_{n+1} = py_n \\ y_{n+1} = q(x_n + y_n) \end{cases}$$

2. On suppose que $p = 1/2$.

a) On pose $v_n = 2^n y_n$. Déterminer une relation de récurrence entre v_{n+1}, v_n et v_{n-1}.

b) En déduire, pour tout $n \geqslant 2$, une expression de x_n puis de u_n, en fonction de n.

c) Vérifier que $\sum\limits_{n=1}^{\infty} u_n = 1$, et en donner une interprétation.

Solution :

1. a) Avec des notations évidentes :

$$u_1 = P(F_1F_2) = p^2,$$
$$x_2 = P(A_2) = P(P_1F_2) = pq,\ y_2 = P(B_2) = P(P_1P_2 \cup F_1P_2) = P(P_2) = q,$$
$$u_2 = P(P_1F_2F_3) = qp^2,$$
$$x_3 = P(P_2F_3) = pq,\ y_3 = P(P_3) - P(F_1F_2P_3) = q - p^2q$$
$$u_3 = P(P_2F_3F_4) = p^2q.$$

b) Clairement $u_n = P(A_nF_{n+1})$ et comme F_{n+1} est indépendant des résultats des lancers précédents :

$$u_n = P(A_n)P(F_{n+1}) = px_n$$

c) $\star$ Si A_n est réalisé, la série de lancers se termine par un résultat Face et A_{n+1} ne peut plus se réaliser (on aurait eu deux fois de suite Face)

$$P(A_{n+1}/A_n) = 0$$

$\star$ En suivant le résultat du rang n, on a :

$$P(A_{n+1}/B_n) = P(F_{n+1}) = p\,;\ P(B_{n+1}/A_n) = P(P_{n+1}) = q$$
$$P(B_{n+1}/B_n) = P(P_{n+1}) = q$$

d) Pour $n \geqslant 2$, (A_n, B_n) est un système complet d'événements et par la formule des probabilités totales :

$x_{n+1} = P(A_{n+1}) = P(B_n)P(A_{n+1}/B_n) = y_np$

$y_{n+1} = P(B_{n+1}) = P(A_n)P(B_{n+1}/A_n) + P(B_n)P(B_{n+1}/B_n) = x_nq + y_nq$

$$\begin{cases} x_{n+1} = py_n \\ y_{n+1} = qx_n + qy_n \end{cases}$$

2. On a ici : $\begin{cases} x_{n+1} = \dfrac{1}{2}\,y_n \\ y_{n+1} = \dfrac{1}{2}(x_n + y_n) \end{cases}$

a) Donc, pour $n \geqslant 2$, $y_{n+2} = \dfrac{1}{2}(x_{n+1} + y_{n+1}) = \dfrac{1}{2}\,y_{n+1} + \dfrac{1}{4}\,y_n$

ce que l'on peut écrire :

$$v_{n+2} = v_{n+1} + v_n$$

b) Le traitement des suites de Fibonacci est classique et il existe des scalaires λ et μ tels que :

$$\forall\, n \geqslant 2, v_n = \lambda\Big(\frac{1+\sqrt{5}}{2}\Big)^n + \mu\Big(\frac{1-\sqrt{5}}{2}\Big)^n$$

En posant $\alpha = \dfrac{1+\sqrt{5}}{2}$ et $\beta = \dfrac{1-\sqrt{5}}{2}$, les conditions initiales donnent alors

$$v_n = \frac{\beta^{n+1} - \alpha^{n+1}}{\beta - \alpha}$$

d'où l'on déduit :

$$y_n = \frac{1}{2^n}\Big(\frac{\beta^{n+1} - \alpha^{n+1}}{\beta - \alpha}\Big) \;;\; x_n = \frac{1}{2^n}\Big(\frac{\beta^{n} - \alpha^{n}}{\beta - \alpha}\Big)$$

$$u_n = \frac{1}{2^{n+1}}\Big(\frac{\beta^{n} - \alpha^{n}}{\beta - \alpha}\Big)$$

c) La formule donnant u_n est encore valable pour $n = 1$ et donc, la convergence des séries rencontrées étant évidente :

$$\sum_{n=1}^{\infty} u_n = \frac{1}{2(\beta-\alpha)}\Big(\sum_{n=1}^{\infty}\big(\frac{\beta}{2}\big)^n - \sum_{n=1}^{\infty}\big(\frac{\alpha}{2}\big)^n\Big)$$

$$= \frac{1}{2(\beta-\alpha)}\Big(\frac{1}{1-\frac{\beta}{2}} - 1 - \frac{1}{1-\frac{\alpha}{2}} + 1\Big) = \frac{1}{(2-\beta)(2-\alpha)}$$

Comme $\alpha + \beta = 1$ et $\alpha\beta = -1$ (racines de l'équation caractéristique ... ou calcul direct), on a finalement :

$$\sum_{n=1}^{\infty} u_n = 1$$

Ce qui prouve que l'on est quasi-certain d'obtenir au moins une fois deux Face de suite dans une succession indéfinie de lancers d'une pièce honnête.

Exercice 3.10.

Soit n un entier supérieur ou égal à 2. Une urne contient une boule noire et $(n-1)$ boules blanches.

On vide l'urne en effectuant des tirages d'une boule de la manière suivante : le premier tirage s'effectue sans remise, le deuxième s'effectue avec remise, le troisième s'effectue sans remise, le quatrième s'effectue avec remise ... D'une manière générale, les tirages d'ordre impair s'effectuent sans remise et les tirages d'ordre pair s'effectuent avec remise de la boule tirée.

1. a) Quel est le nombre total N de tirages effectués lors de cette épreuve ?

b) Pour tout $j \in [\![1, n-1]\!]$, combien reste-t-il de boules avant le $(2j)^{\text{ème}}$ tirage ? Combien en reste-t-il avant le $(2j+1)^{\text{ème}}$ tirage ?

2. On désigne par X_k la variable aléatoire qui vaut 1 si la boule noire est obtenue au $k^{\text{ème}}$ tirage (que ce soit la première fois ou non) et 0 sinon.
On désigne par X la variable aléatoire égale au nombre d'apparitions de la boule noire lors de cette épreuve.

a) Calculer $P(X_1 = 1), P(X_2 = 1)$.

b) Pour tout entier naturel $j \in [\![1, n-1]\!]$, calculer $P(X_{2j+1} = 1)$ et $P(X_{2j} = 1)$.

3. Pour tout $j \in [\![1, n-1]\!]$, on note U_j l'événement « On obtient la boule noire pour la première fois au $(2j-1)^{\text{ème}}$ tirage ».

a) En considérant l'état de l'urne avant le $(2n-2)^{\text{ème}}$ tirage, montrer que $P(U_n) = 0$.

b) Montrer que pour tout $j \in [\![1, n-1]\!]$, on a : $P(U_j) = \dfrac{n-j}{n(n-1)}$.

c) Exprimer l'événement $(X = 1)$ en fonction des événements (U_j), et en déduire la valeur de $P(X = 1)$.

Calculer $P(X = n)$.

Solution :

1. a) Suivons le nombre de boules restant dans l'urne après chaque pas du tirage :

$$n-1, n-1, n-2, n-2, \cdots, 1, 1, 0.$$

Il y a donc eu exactement $N = 2(n-1)+1 = 2n-1$ tirages pour vider l'urne.

b) Comme il y a remise après le $(2j)$-ième tirage, il y a autant de boules restantes avant le $(2j)$-ième tirage qu'après le $(2j)$-ième tirage (voir la question précédente). Il reste donc $n-j$ boules.

2. a) L'événement $(X_1 = 1)$ est l'événement «la boule noire est sortie au premier tirage» ; donc $P(X_1 = 1) = \dfrac{1}{n}$.
L'événement $(X_2 = 1)$ est l'événement «la boule noire est sortie au second tirage et pas au premier» ; donc

$$P(X_2 = 1) = P(X_1 = 0)P(X_2 = 1/X_1 = 0) = (1 - \frac{1}{n})\frac{1}{n-1} = \frac{1}{n}$$

b) Comme il n'y a pas remise lors des tirages impairs, il vient

$$(X_{2j} = 1) = (X_1 = 0) \cap (X_3 = 0) \cap \cdots \cap (X_{2j-1} = 0) \cap (X_{2j} = 1)$$

Par la formule des probabilités composées (c'est-à-dire en suivant la composition de l'urne au cours du temps) :

$$P(X_{2j} = 1) = \frac{n-1}{n} \times \frac{n-2}{n-1} \times \cdots \times \frac{n-j}{n-j+1} \times \frac{1}{n-j} = \frac{1}{n}$$

Et comme après le $(2j)$-ième tirage, l'urne se retrouve dans le même état qu'auparavant, on a

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

Ainsi, pour tout $k \in [\![1, 2n-1]\!]$, X_k suit la loi de Bernoulli de paramètre $1/n$.

3. a) L'événement U_n correspond à « la boule noire est sortie au dernier tirage». Le tirage précédent étant pair, il y a eu remise, et c'est la même boule qui est sortie au $(2n-2)$-ième tirage. Ainsi la boule noire ne peut être sortie au dernier tirage, et $P(U_n) = 0$.

b) Comme

$$U_j = (X_1 = 0) \cap (X_2 = 0) \cap \cdots \cap (X_{2j-2} = 0) \cap (X_{2j-1} = 1),$$

par la formule des probabilités composées, il vient :

$$P(U_j) = \frac{n-1}{n} \times \frac{n-2}{n-1} \times \frac{n-2}{n-1} \times \cdots \times \frac{n-j}{n-j+1} \times \frac{1}{n-j+1} = \tfrac{n-j}{n(n-1)}$$

c) Seuls les tirages d'ordre impair évacuent la boule noire dès son apparition. Donc $(X = 1) = \bigcup_{j=1}^{n} U_j$. Ces derniers événements étant deux à deux incompatibles, il vient :

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

L'événement $(X=n)$ correspond à « la boule noire est sortie à tous les tirages pairs et n'a pas été choisie lors des tirages impairs, sauf lors du dernier ». Soit :

$$(X=n)=(X_1=0)\cap(X_2=1)\cap(X_3=0)\cap\cdots\cap(X_{2n-2}=1)\cap(X_{2n-1}=1)$$

En utilisant de nouveau la formule des probabilités composées, il vient :

$$P(X=n)=\frac{n-1}{n}\times\frac{1}{n-1}\times\frac{n-2}{n-1}\times\frac{1}{n-2}\times\cdots\times 1\times 1=\frac{1}{n!}$$

Exercice 3.11.

Deux trains sont prévus au départ d'une ville A : un premier train doit partir à 12 heures et le suivant à 13 heures, mais ces départs peuvent subir un retard. Les retards sont des variables aléatoires Z_1 et Z_2 indépendantes, de même loi à valeurs dans $[0,1]$ (l'unité de temps est donc l'heure), de densité f, de fonction de répartition F, d'espérance μ et de variance σ^2.

1. Un voyageur arrive à l'instant x après 12 heures, avec $x\in[0,1]$. On note T_x le temps qu'il faudra à ce voyageur pour commencer son voyage.

a) Montrer que T_x prend ses valeurs entre 0 et $2-x$.

b) Montrer que :

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

En déduire une densité de T_x.

2. Montrer que l'espérance $m(x)$ de T_x vérifie :

$$\begin{aligned}m(x)&=\int_0^{1-x}tf(t+x)\,dt+(\mu+1-x)F(x)\\&=1-x-\int_x^1F(u)\,du+(\mu+1-x)F(x)\end{aligned}$$

3. a) Calculer $\int_0^1F(z)\,dz$ et $\int_0^1zF(z)\,dz$ en fonction de μ et σ^2.

b) On suppose que le voyageur arrive au hasard entre midi et 13 heures.

On admet que l'espérance M de son temps d'attente avant de commencer son voyage est alors donnée par :

$$M=\int_0^1m(x)\,dx$$

Calculer M.

Solution :

1. a) La variable aléatoire T_x prend ses valeurs entre 0 et $2-x$, puisque soit il prend le premier train (qui avait du retard et part juste à ce moment), soit le second (qui peut avoir jusqu'à une heure de retard).

b) • Si $x+t < 1$. Dans ce cas $P(T_x \leqslant t)$ représente la probabilité que le train de 12 h. parte dans l'intervalle $[x, x+t]$. Donc $P(T_x \leqslant t) = F(t+x) - F(x)$.

• Si $1 \leqslant x+t < 2$. La probabilité de partir avec le train de midi est $1 - F(x)$, et celle de partir avec le train de 13 h. est la probabilité d'avoir raté celui de midi multiplié par la probabilité d'avoir un retard dans l'intervalle $[0, x+t-1]$. Ainsi $P(T_x \leqslant t) = 1 - F(x) + F(x)F(t+x-1)$.

• Si $x+t \geqslant 2$. On est alors certain d'avoir le train de 13 h. Donc $P(T_x \leqslant t) = 1$.

Par dérivation (par rapport à t) une densité de T_x est :

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

2. Un calcul évident donne :

$$\begin{aligned} E(T_x) &= \int_0^{1-x} tf(t+x)\,dt + \int_{1-x}^{2-x} tF(x)f(x+t-1)\,dt \\ &= \int_x^1 (u-x)f(u)du + F(x)\int_0^1 (u-x+1)f(u)\,du \\ &= \int_x^1 uf(u)\,du - x\int_x^1 f(u)\,du + \mu F(x) + (1-x)F(x) \\ &= 1 - x - \int_x^1 F(u)\,du + (\mu+1-x)F(x) \end{aligned}$$

la dernière ligne s'obtenant par une intégration par parties de la première intégrale.

3. a) Il vient :

$$\mu = \int_0^1 zf(z)\,dz = \big[zF(z)\big]_0^1 - \int_0^1 F(z)\,dz, \text{ d'où : } \int_0^1 F(z)\,dz = 1 - \mu$$

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

$$\int_0^1 zF(z)\,dz = \frac{1-\mu^2-\sigma^2}{2}$$

b) Il vient :

$$\begin{aligned} M &= \int_0^1 E(T_x)\,dx \\ &= \int_0^1 (1-x)\,dx - \int_0^1 \left(\int_x^1 F(z)dz\right) dx + \int_0^1 (\mu+1-x)F(x)\,dx \end{aligned}$$

À l'aide d'une intégration par parties :

$$\int_0^1 \left(\int_x^1 F(z)\,dz\right) dx = \left[x\int_x^1 F(z)\,dz\right]_0^1 + \int_0^1 xF(x)\,dx = \frac{1-\mu^2-\sigma^2}{2}$$

et,

$$M = \frac{1}{2} + \frac{1-\mu^2-\sigma^2}{2} + 1 - \mu^2 - \frac{1-\mu^2-\sigma^2}{2} = \frac{1}{2} + \sigma^2$$

Ainsi $E(W)$ est minimal lorsque $\sigma^2 = 0$, par exemple lorsqu'on est sûr que les trains sont à l'heure …

Exercice 3.12.

Soient X_1, X_2, X_3 trois variables aléatoires réelles discrètes et centrées.
On note M leur matrice de variance-covariance, soit : $M = (m_{i,j})_{1\leqslant i,j\leqslant 3}$, avec

$$m_{i,j} = E(X_iX_j)$$

1. Montrer que la matrice M est diagonalisable et que ses valeurs propres sont positives ou nulles.

2. On suppose dans cette question que :

$$M = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Déterminer les valeurs propres de M.
Les variables aléatoires X_1, X_2, X_3 sont-elles alors mutuellement indépendantes ?

3. Soit Y une variable aléatoire centrée. On définit une fonction F sur $\mathbb{R}^3$ par :

$$F(x_1, x_2, x_3) = E\big[\big(Y - \sum_{i=1}^{3} x_iX_i\big)^2\big]$$

Montrer qu'une condition nécessaire et suffisante pour que F admette un *minimum* en (a, b, c) est :

$$M \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} E(YX_1) \\ E(YX_2) \\ E(YX_3) \end{pmatrix}$$

4. On suppose dans cette question que la matrice M est celle de la question 2.

a) Donner une condition nécessaire et suffisante sur (α, β, γ) pour que le système linéaire $M \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ admette une solution.

b) Sous cette condition, donner l'expression de cette solution.

Solution :

1. La matrice M est symétrique réelle donc diagonalisable. Ses valeurs propres sont positives ou nulles car :

$$X^TMX = \sum_{i=1}^{3}\sum_{j=1}^{3} x_ix_jE(X_iX_j) = \sum_{i=1}^{3} x_i^2E(X_i^2) + 2\sum_{i\neq j} \mathrm{Cov}(x_iX_i, x_jX_j)$$
$$= V\big(\sum_{i=1}^{3} x_iX_i\big) \geqslant 0$$

2. On remarque que

$$M - 3I = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} = -J$$

On sait que les valeurs propres de J sont 0, de sous-espace propre associé le plan engendré par $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ et 3, la droite propre associée étant dirigée par la colonne $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Les valeurs propres de M sont donc 3 et 0, les sous-espaces propres étant les mêmes.

Comme $E(X_iX_j) = -1 \neq E(X_i)E(X_j) = 0$, les variables aléatoires (X_i) ne sont pas indépendantes.

3. On a :

$F(x_1, x_2, x_3) = E(Y^2) - 2\sum_{i=1}^{3} x_i E(YX_i) + \sum_{i=1}^{3} x_i^2 E(X_i^2) + 2\sum_{i<j} x_i x_j E(X_iX_j)$

Les points critiques de F, qui est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$, sont donnés par :

$$\begin{cases} \dfrac{\partial F}{\partial x_1} = -2E(YX_1) + 2x_1E(X_1^2) + 2x_2E(X_1X_2) + 2x_3E(X_1X_3) = 0 \\ \dfrac{\partial F}{\partial x_2} = -2E(YX_2) + 2x_2E(X_2^2) + 2x_1E(X_1X_2) + 2x_3E(X_2X_3) = 0 \\ \dfrac{\partial F}{\partial x_3} = -2E(YX_3) + 2x_3E(X_3^2) + 2x_3E(X_1X_3) + 2x_2E(X_2X_3) = 0 \end{cases}$$

Ce système d'équations est équivalent à l'équation matricielle

$$M\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} E(X_1Y) \\ E(X_2Y) \\ E(X_3Y) \end{pmatrix}$$

Ainsi :

• si F admet un minimum en (a, b, c), c'est un point critique donné par la résolution de l'équation matricielle $M\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} E(X_1Y) \\ E(X_2Y) \\ E(X_3Y) \end{pmatrix}$.

• Si $M\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} E(X_1Y) \\ E(X_2Y) \\ E(X_3Y) \end{pmatrix}$, alors (a, b, c) est un extremum de F et $F(x_1, x_2, x_3) \geqslant 0$ entraîne que cet extremum est un minimum.

4. a) L'équation $M\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ admet une solution si et seulement si $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ appartient à l'image de M, donc si et seulement s'il existe (λ, μ) réels tels que :

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \lambda\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} + \mu\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

soit :

$$\begin{cases} \alpha &= 2\lambda - \mu \\ \beta &= -\lambda + 2\mu \\ \gamma &= -\lambda - \mu \end{cases}$$

ce qui est équivalent à $\alpha + \beta + \gamma = 0$ (en fait $x + y + z = 0$ est l'équation de l'image de M).

b) Sous la condition $\alpha + \beta + \gamma = 0$, la solution est :

$$x_1 = \frac{\alpha - \gamma}{3}, x_2 = \frac{\beta - \gamma}{3}, x_3 = \frac{2\gamma - \alpha - \beta}{3}$$

Exercice 3.13.

Soient $n \in \mathbb{N}^*$ quelconque, $p \in]0, 1[$, $(X_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires, indépendantes, suivant la même loi uniforme à densité sur $[0, n]$, et N_n une variable aléatoire indépendante des X_i suivant la loi binomiale $\mathcal{B}(n, p)$.

On pose :

$U_n = \max(X_0, \ldots, X_n)$, $V_n = \min(X_0, \ldots, X_n)$, $W_n = \min(X_0, \ldots, X_{N_n})$

1. Déterminer la loi de U_n , son espérance et sa variance.

2. Montrer que la loi de V_n est celle de $n - U_n$. En déduire $E(V_n)$ et $V(V_n)$.

3. On note G_n la fonction de répartition de V_n. Déterminer $\lim\limits_{n\to+\infty} G_n(t)$ suivant les valeurs de t. Que peut-on en déduire ?

4. Déterminer la fonction de répartition H_n de W_n, puis déterminer $\lim\limits_{n\to+\infty} H_n(t)$ suivant les valeurs de t. Conclure.

Solution :

1. Il vient immédiatement par indépendance :

$$F_{U_n}(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ P\big(\bigcap\limits_{i=0}^{n}(X_i \leqslant x)\big) = \big(\frac{x}{n}\big)^{n+1} & \text{si } 0 \leqslant x \leqslant n \\ 1 & \text{si } x > n \end{cases}$$

Une densité de U_n est alors, par dérivation :

$$f_{U_n}(x) = \begin{cases} \dfrac{n+1}{n}\big(\frac{x}{n}\big)^n & \text{si } 0 \leqslant x \leqslant n \\ 0 & \text{sinon} \end{cases}$$

Un calcul immédiat donne :

$$E(U_n) = \frac{n(n+1)}{n+2}, V(U_n) = \frac{n^2(n+1)}{(n+3)(n+2)^2}$$

2. Comme $n - U_n = n - \max(X_0, \ldots, X_n) = \min(n - X_0, \ldots, n - X_n)$, et comme $n - X_i$ suit la loi uniforme sur $[0, n]$, la variable aléatoire $n - U_n$ suit la même loi que la variable V_n, et

$$E(V_n) = n - E(U_n) = \frac{n}{n+2}, V(V_n) = V(U_n)$$

3. Pour tout x réel :

$$G_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - \big(1 - \frac{x}{n}\big)^n & \text{si } 0 \leqslant x \leqslant n \\ 1 & \text{si } x \geqslant n \end{cases}$$

et de manière évidente, pour x fixé :

$$\lim_{n\to+\infty} G_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1-\mathrm{e}^{-x} & \text{si } x > 0 \end{cases}$$

Ainsi (V_n) converge en loi vers la loi exponentielle de paramètre 1.

4. On a bien évidemment $H_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 & \text{si } x \geqslant n \end{cases}$, et $H_n(x) = 1 - P(W_n \geqslant x)$.

Lorsque $x \in [0, n]$, en appliquant la formule des probabilités totales au système complet d'événements $((N = k))_{0\leqslant k\leqslant n}$, il vient :

$$\begin{aligned} H_n(x) &= 1 - \sum_{k=0}^{n} P(W_n > x \mid N = k)P(N = k) \\ &= 1 - \sum_{k=0}^{n} \big(1 - \frac{x}{n}\big)^{k+1}\binom{n}{k} p^k q^{n-k} = 1 - \big(1-\frac{x}{n}\big)\big(p(1-\frac{x}{n}) + q\big)^n \\ &= 1 - \big(1-\frac{x}{n}\big)\big(1-\frac{px}{n}\big)^n \end{aligned}$$

Donc, à x fixé, $\lim\limits_{n\to+\infty} H_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1-\mathrm{e}^{-px} & \text{si } x > 0 \end{cases}$, et :

(W_n) converge en loi vers la loi exponentielle de paramètre p.

Exercice 3.14.

Une urne contient $n_1 \geqslant 1$ boules blanches, $n_2 \geqslant 1$ boules noires et $r \geqslant 0$ boules rouges. On note $n = n_1 + n_2$. Le jeu consiste à effectuer une succession de tirages au hasard une boule de l'urne, sans remise, jusqu'à ce que la boule tirée soit :

— ou blanche, auquel cas la partie est gagnée,

— ou noire, et la partie est perdue.

La longueur de la partie est le nombre de boules sorties de l'urne à la fin de la partie.

1. Compléter le programme Pascal qui suit pour qu'il simule le déroulement d'une partie :

```
program simul-escp
var n1,n2,r :integer ; L :integer ; x : char ;
begin
  readln(n1) ; readln(n2) ; readln(r) ;
  randomize ;
  ...
  repeat
    ...
    alea :=1+ random(n1+n2+r) ;
    if ((1 <= alea) and (alea <= n1))
    then ...
    ...
  until (x \= 'r') ;
  if (x='n') then write('perdu') else write('gagné') ;
  writeln (' et en ',L,' coups.') ;
end.
```

2. L'urne contenant au départ r boules rouges, on note G_r la variable aléatoire égale à 1 si le joueur gagne et à 0 sinon et $E(G_r)$ l'espérance de G_r.

a) Déterminer $E(G_r)$ pour $r = 0$ puis $r = 1$.

b) Montrer que l'on a, pour tout entier naturel non nul r :

$$E(G_r) = \frac{n_1}{n+r} + \frac{r}{n+r}E(G_{r-1}).$$

En déduire la valeur de $E(G_r)$.

3. On note L_r la variable aléatoire égale à la longueur d'une partie lorsque l'urne contient r boules rouges.

a) Quelles sont les valeurs possibles de L_r ?

b) Calculer $E(L_0), E(L_1)$.

c) Déterminer une relation entre $P(L_{r-1} = k-1)$ et $P(L_r = k)$, pour $r \geqslant 1$ et $k \geqslant 2$.

d) Etablir une relation de récurrence entre : $E(L_r)$ et $E(L_{r-1})$, pour $r \geqslant 1$.

e) En déduire l'expression de $E(L_r)$.

Solution :

1. Une proposition de programme :

```
Program simul ;
Var n1,n2,r,L : integer ; x : char ;
Begin
   readln(n1) ; readln(n2) ; readln(r) ;
   randomize ;
   L := 0 ;
   Repeat
      L := L+1 ;
      alea := 1 + random(n1+n2+r) ;
      if ((1<= alea) and (alea <= n1)) then
         Begin
         x : = 'b' ; n1 := n1-1 ;
         End
      else if ((n1+1 <= alea) and ( alea <= n1+n2)) then
         Begin
         x := 'n' ; n2 := n2-1
         end
      else
         Begin
         x := 'r' ; r := r-1
         end
   until (x <> 'r') ;
   if x='n' then write ('perdu') else write('gagné') ;
   writeln(' en',L,' coups')
End.
```

2. a) La variable aléatoire G suit la loi de Bernoulli de paramètre $P(G = 1) = E(G)$.

• Si $r = 0$, l'événement $(G = 1)$ est l'événement «tirer une boule blanche au premier tirage». Donc $E(G) = \frac{n_1}{n}$.

• Si $r = 1$, l'événement $(G = 1)$ est l'événement «tirer une boule blanche au premier tirage ou tirer une rouge puis une blanche au second tirage».

$$P(G = 1) = \frac{n_1}{n+1} + \frac{n_1}{n} \times \frac{1}{n+1} = \frac{n_1}{n}$$

b) On note B_1, N_1, R_1 les événements «tirer une boule blanche, (resp. noire, rouge) au premier tirage». C'est un système complet d'événements.

Notons $f(r) = E(G)$. Alors :

$P(G=1) = P(G_1 = 1/B_1)P(B_1) + P(G_1 = 1/N_1)P(N_1) + P(G_1 = 1/R_1)P(R_1)$

d'où :

$$f(r) = \frac{n_1}{n+r} + 0 + f(r-1)\times\frac{r}{n+r} = \frac{n_1}{n+r} + f(r-1)\times\frac{r}{n+r}$$

On sait que $f(0) = f(1) = \dfrac{n_1}{n}$.

Supposons que $f(r-1) = \dfrac{n_1}{n}$. La relation précédente montre que $f(r) = \dfrac{n_1}{n}$.

Finalement, par le principe de récurrence :

$$E(G) = \frac{n_1}{n}$$

3. a) On a $L(\Omega) = [\![1, r+1]\!]$.

b) Comme $L_0 = 1$, on a $E(L_0) = 1$.

On a $P(L_1 = 1) = \dfrac{n}{n+r}$ et $P(L_1 = 2) = 1 - P(L_1 = 1)$. D'où $E(L_1) = \dfrac{n+2}{n+1}$.

c) Soir $r \geqslant 1$. Alors, si $k \geqslant 2$,

$$P(L_r = k) = P(R_1)P(L_r = k/R_1) = \frac{r}{n+r}P(L_{r-1} = k-1).$$

d) Il vient :

$$\begin{aligned} E(L_r) &= P(L_r = 1) + \sum_{k=2}^{r+1} kP(L_r = k) \\ &= \frac{n}{n+r} + \frac{r}{n+r}\sum_{k=2}^{r+1} kP(L_{r-1} = k-1) = \frac{n}{n+r} + \frac{r}{n+r}\big(E(L_{r-1})+1\big) \end{aligned}$$

e) On montre par récurrence (immédiate) que $E(L_r) = \dfrac{n+r+1}{n+r}$.

Exercice 3.15.

Soit n un entier naturel non nul et λ un nombre réel fixé. On pose

$$P_n(\lambda) = \mathrm{e}^{-\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!}$$

Soit $X_1, \ldots, X_n$ n variables aléatoires définies sur le même espace probabilisé, indépendantes et suivant toutes la loi de Poisson de paramètre 1.

1. Montrer que P_n est strictement concave sur l'intervalle $[n-1, n]$.

2. Montrer que pour tout entier $n > 0$, on a : $P_n(n-1) + P'_n(n-1) = P_{n-1}(n-1)$.

3. A l'aide d'une formule de Taylor, montrer que la suite $\big(P_n(n)\big)_{n\geqslant 1}$ est strictement décroissante.

4. Quelle est la loi de $S_n = X_1 + \cdots + X_n$? A l'aide du théorème de la limite centrée, démontrer que $\lim\limits_{n\to\infty} P_n(n) = \dfrac{1}{2}$.

5. Montrer que pour tout n de $\mathbb{N}^*$, on a $P_n(n) > \dfrac{1}{2}$.

Solution :

1. En dérivant, il vient :

$$P_n'(\lambda) = -e^{-\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!} + e^{-\lambda}\sum_{k=1}^{n}\frac{\lambda^{k-1}}{(k-1)!} = -e^{-\lambda}\frac{\lambda^n}{n!}$$

et

$$P_n''(\lambda) = e^{-\lambda}\frac{\lambda^{n-1}}{(n-1)!}\big(\frac{\lambda}{n} - 1\big) \leqslant 0, \qquad \text{si } \lambda \in]n-1, n[$$

Ainsi $\lambda \mapsto P_n(\lambda)$ est strictement concave sur $[n-1, n]$.

2. Par calcul

$$\begin{aligned} P_n(n-1) + P_n'(n-1) &= e^{-n+1}\sum_{k=0}^{n}\frac{(n-1)^k}{k!} - e^{-(n-1)}\sum_{k=0}^{n}\frac{(n-1)^k}{k!} \\ &\qquad + e^{-n+1}\sum_{k=1}^{n}\frac{(n-1)^{k-1}}{(k-1)!} \\ &= e^{-n+1}\sum_{k=1}^{n}\frac{(n-1)^{k-1}}{(k-1)!} = e^{-(n-1)}\sum_{k=0}^{n-1}\frac{(n-1)^k}{k!} \\ &= P_{n-1}(n-1) \end{aligned}$$

3. Utilisons la formule de Taylor avec reste intégral à l'ordre 2 appliquée à P_n sur $[n-1, n]$. Il vient :

$$P_n(n) = P_n(n-1) + P_n'(n-1) + \int_{n-1}^{n}\frac{(n-t)^2}{2}P_n''(t)\,dt$$

Donc, par la question précédente, et par concavité de $t \mapsto P_n(t)$

$$\begin{aligned} P_n(n) - P_{n-1}(n-1) &= P_n(n) - P_n(n-1) + P_n(n-1) - P_{n-1}(n-1) \\ &= P_n(n) - P_n(n-1) - P_n'(n-1) \\ &= \int_{n-1}^{n}\frac{(n-t)^2}{2}P_n''(t)dt < 0 \end{aligned}$$

4. On remarque immédiatement que $P_n(n) = P(S_n \leqslant n)$. Ainsi :

$$P_n(n) = P(S_n - n \leqslant 0) = P(S_n - E(S_n) \leqslant 0) = P\big(\frac{S_n - n}{\sqrt{n}} \leqslant 0)$$

On termine en invoquant le théorème de la limite centrée, qui nous assure de la convergence en loi de $\dfrac{S_n - n}{\sqrt{n}}$ vers une variable aléatoire suivant la loi normale centrée réduite.

5. La suite $(P_n(n))$ est strictement décroissante et tend vers 1/2, par valeurs supérieures, d'où le résultat.

Exercice 3.16.

Soit $X_1, X_2, \dots, X_n$ n variables aléatoires indépendantes et suivant la même loi de Poisson de paramètre θ. On désire estimer, de la meilleure façon possible, $e^{-\theta}$.

1. Pour tout entier i compris entre 1 et n, on pose $Y_i = 1$ si $X_i = 0$ et $Y_i = 0$ sinon.

On pose $N = \sum\limits_{i=1}^{n} Y_i$ et $S_n = \sum\limits_{i=1}^{n} X_i$.

a) Montrer que $\dfrac{1}{n}N$ est un estimateur sans biais de $e^{-\theta}$.

b) Calculer sa variance, et montrer que cet estimateur est convergent.

c) Interpréter ce résultat en expliquant pourquoi cet estimateur est un estimateur « naturel » de $\mathrm{e}^{-\theta}$.

2. Pour tout entier naturel j, on définit la probabilité conditionnelle :

$$\varphi(j) = P(X_1 = 0/S_n = j)$$

Calculer $\varphi(j)$ (on vérifiera que le résultat est indépendant de la valeur de θ).

3. a) Montrer que $\varphi(S_n)$ est un estimateur sans biais de $\mathrm{e}^{-\theta}$.

b) Calculer la variance de $\varphi(S_n)$ et montrer que $\varphi(S_n)$ est un estimateur convergent.

c) Montrer que, quelle que soit la valeur de θ, $\varphi(S_n)$ a une variance inférieure à celle de $\frac{1}{n}N$.

Solution :

1. a) Pour tout i, la variable aléatoire Y_i suit la loi de Bernoulli de paramètre p_i, avec

$$p_i = P(Y_i = 1) = P(X_i = 0) = e^{-\theta}$$

Donc

$$E\big(\frac{N}{n}\big) = \frac{1}{n}\times n\,\mathrm{e}^{-\theta} = \mathrm{e}^{-\theta}$$

Ainsi $\frac{1}{n}N$ est un estimateur sans biais de $\mathrm{e}^{-\theta}$.

b) Les variables aléatoires Y_i sont indépendantes, puisque les variables X_i le sont. Au vu de la loi de Y_i, $V(Y_i) = \mathrm{e}^{-\theta}(1-\mathrm{e}^{-\theta})$ et :

$$V\big(\frac{N}{n}\big) = \frac{1}{n^2}\times nV(Y) = \frac{\mathrm{e}^{-\theta}(1-e^{-\theta})}{n}$$

Ainsi $\frac{1}{n}N$ est un estimateur sans biais convergent de $\mathrm{e}^{-\theta}$.

c) $\frac{1}{n}N$ représente la proportion, parmi les variables $X_1, \ldots, X_n$, des variables aléatoires égales à 0. C'est donc un estimateur naturel de $P(X_i = 0)$.

2. On a :

$$\begin{aligned}\varphi(j) = P(X_1 = 0/S_n = j) &= \frac{P(X_1 = 0)\cap(S_n = j))}{P(S_n = j)}\\ &= \frac{P(X_1 = 0)\cap(X_2 + \cdots + X_n = j))}{P(S_n = j)}\end{aligned}$$

$$\varphi(j) = \frac{P(X_2 + \cdots + X_n = j)P(X_1 = 0)}{P(S_n = j)}$$

Or, par le cours, $\sum\limits_{i=1}^{n} X_i$ suit la loi de Poisson de paramètre $n\theta$ et $\sum\limits_{i=2}^{n} X_i$ suit la loi de Poisson de paramètre $(n-1)\theta$. Aussi :

$$\varphi(j) = \frac{\frac{\mathrm{e}^{-(n-1)\theta}(n-1)^j\theta^j}{j!}}{\mathrm{e}^{-n\theta}\frac{n^j\theta^j}{j!}}\times\mathrm{e}^{-\theta} = \big(\frac{n-1}{n}\big)^j$$

3. a) Par le théorème de transfert :

$$E(\varphi(S_n)) = E\big(\frac{n-1}{n}\big)^{S_n} = \sum_{k=0}^{\infty}\big(\frac{n-1}{n}\big)^k\mathrm{e}^{-n\theta}\frac{(n\theta)^k}{k!} = \mathrm{e}^{-\theta}$$

Ainsi $\varphi(S_n)$ est un estimateur sans biais de $\mathrm{e}^{-\theta}$.

b) On peut écrire :

$$E(\varphi(S_n)^2) = E\big(\big(\frac{n-1}{n}\big)^{2S_n}\big) = \sum_{k=0}^{\infty}\big(\frac{n-1}{n}\big)^{2k}\mathrm{e}^{-n\theta}\frac{(n\theta)^k}{k!}$$
$$= \mathrm{e}^{-n\theta}\sum_{k=0}^{\infty}\frac{1}{k!}\big[(\frac{n-1}{n})^2 n\theta\big]^k = \mathrm{e}^{-n\theta}\mathrm{e}^{(\frac{n-1}{n})^2 n\theta}$$

D'où :

$$V(\varphi(S_n)) = \mathrm{e}^{(n-2+\frac{1}{n})\theta - n\theta} - \mathrm{e}^{-2\theta} = \mathrm{e}^{-2\theta}(\mathrm{e}^{\theta/n}-1)$$

Donc $\lim\limits_{n\to+\infty} V(\varphi(S_n)) = 0$ et $\varphi(S_n)$ qui est un estimateur sans biais est un estimateur convergent de $\mathrm{e}^{-\theta}$.

c) Il s'agit de montrer ici l'inégalité $\mathrm{e}^{-\theta}(\mathrm{e}^{\theta/n}-1) \leqslant \dfrac{1-\mathrm{e}^{-\theta}}{n}$, ce qui revient à montrer que $\mathrm{e}^{\theta/n} - 1 \leqslant \dfrac{\mathrm{e}^{\theta}-1}{n}$

Pour cela, on étudie sur $\mathbb{R}^+$ les variations de la fonction $h : x \mapsto \mathrm{e}^x - 1 - n.\mathrm{e}^{x/n} + n$, et on montre sa croissance, puisque sa dérivée est $h'(x) = \mathrm{e}^x - \mathrm{e}^{x/n}$ qui est positive. Comme $h(0) = 0$, h est strictement positive sur $\mathbb{R}_+^*$.

Exercice 3.17.

Soit $(U_k)_{k\geqslant 1}$ une suite de variables aléatoires indépendantes suivant la même loi uniforme sur l'intervalle $[0,1]$.

Pour $k \in \mathbb{N}^*$, on définit la variable aléatoire S_k par : $S_k = \max(U_1, U_2, \ldots, U_k)$.

Soit $n \in \mathbb{N}^*$. On considère une variable aléatoire X_n qui suit la loi uniforme sur l'ensemble $\{1, 2, \ldots, n\}$, et on pose alors $T_n = S_{X_n}$.

1. Donner la fonction de répartition de la variable aléatoire T_n.

2. En déduire que T_n est une variable aléatoire à densité. Donner une densité de T_n, puis calculer l'espérance $E(T_n)$ et la variance $V(T_n)$.

3. On rappelle que $\sum\limits_{k=1}^{n}\dfrac{1}{k} \sim \ln(n)$.

Déterminer $\lim\limits_{n\to\infty} E(T_n)$ et $\lim\limits_{n\to\infty} V(T_n)$.

4. Étudier la convergence en loi de la suite $(T_n)_n$.

5. Écrire une fonction PASCAL dont l'intitulé est :

```
function maxi(S :array[0..9] of real) :real ;
```

qui rend le plus grand nombre du tableau `S`.

6. Écrire un programme en PASCAL permettant de simuler la variable aléatoire T_{10}.

Solution :

1. la famille $((X_n = 1), \ldots, (X_n = n))$ forme un système complet d'événements. Donc, pour tout x réel,

$$P(T_n \leqslant x) = \sum_{k=1}^{n} P(T_n \leqslant x / X_n = k)P(X_n = k) = \frac{1}{n}\sum_{k=1}^{n} P(S_k \leqslant x)$$

Par indépendance des variables aléatoires (U_i), il vient :

$$P(S_k \leqslant x) = \prod_{i=1}^{k} P(U_i \leqslant x) = \begin{cases} 0 & \text{si } x < 0 \\ x^k & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}$$

Soit :

$$F_{T_n}(x) = P(T_n \leqslant x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{1}{n}\sum\limits_{k=1}^{n} x^k & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}$$

2. La fonction F_{T_n} est continue, croissante, de classe C^1 par morceaux, de limite nulle en $-\infty$ et de limite 1 en $+\infty$. Par dérivation, la variable aléatoire T_n admet une fonction densité définie par :

$$f_{T_n}(x) = \begin{cases} 0 & \text{si } x \notin [0,1] \\ \dfrac{1}{n}\sum\limits_{k=1}^{n} kx^{k-1} & \text{si } x \in [0,1] \end{cases}$$

Un calcul immédiat d'espérance et de variance donne :

$$E(T_n) = \frac{1}{n}\sum_{k=1}^{n} \frac{k}{k+1}, V(T_n) = \frac{1}{n}\sum_{k=1}^{n} \frac{k}{k+2} - \big(\frac{1}{n}\sum_{k=1}^{n} \frac{k}{k+1}\big)^2$$

3. En écrivant $k = (k+1) - 1$, il vient :

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

et par l'équivalent proposé :

$$\lim_{n\to+\infty} E(T_n) = 1 - \lim_{n\to+\infty} \frac{\ln n}{n} = 1$$

Un argument similaire permet de dire que : $\lim\limits_{n\to+\infty} V(T_n) = 1 - 1 = 0$.

4. Si $x \leqslant 0$, alors, $\lim\limits_{n\to+\infty} P(T_n \leqslant x) = 0$.

Si $x \geqslant 1$, alors, $\lim\limits_{n\to+\infty} P(T_n \leqslant x) = 1$.

Si $x \in]0,1[$, alors, $\lim\limits_{n\to+\infty} P(T_n \leqslant x) = \lim\limits_{n\to+\infty} \dfrac{x}{n} \times \dfrac{1-x^n}{1-x} = 0$.

Donc $(T_n)_n$ converge en loi vers la variable aléatoire constante égale à 1.

5. Une proposition de programme :

```
Function maxi(S :array[0..9] of real) : real ;
Var k : integer ; m : real ;
begin
m := S[0] ;
For k := 1 to 9 do if S[k]>m then m := S[k] ;
maxi := m
end ;
```

6. Une proposition de programme :

```
Program exo ;
Var k,x : integer ; S : array[0..9] of real
function maxi(...)
Begin
For k := 0 to 9 do S[k]=0 ;
randomize ;
x := random(10)+1 ;
For k := 1 to x-1 do S[k] := random
```

```
writeln(maxi(S)) ;
End.
```

Exercice 3.18.

Le but de l'exercice est l'estimation d'un paramètre θ d'une variable aléatoire X de densité f. On désigne par n un entier naturel non nul. Soit $(X_1, X_2, \dots, X_n)$ un n-échantillon de la loi de X. On note $(x_1, x_2, \dots, x_n)$ un échantillon observé. Pour un tel échantillon observé donné, on appelle vraisemblance de θ la fonction L définie par :

$$L(\theta) = \prod_{i=1}^{n} f(x_i) = f(x_1)f(x_2)\cdots f(x_n)$$

La méthode du maximum de vraisemblance consiste à utiliser comme estimation de θ pour un échantillon observé donné $(x_1, x_2, \dots, x_n)$ une valeur notée $\widehat{\theta}$ pour laquelle la vraisemblance $L(\theta)$ ou, de façon équivalente, $\ln(L(\theta))$ atteint un maximum. Cette valeur $\widehat{\theta}$ est fonction de $(x_1, x_2, \dots, x_n)$ et on peut alors construire un estimateur Y_n du paramètre θ comme fonction correspondante des variables aléatoires $(X_1, X_2, \dots, X_n)$.
Par exemple, si $\widehat{\theta} = x_1 + x_2 + \cdots + x_n$, on prendra : $Y_n = X_1 + X_2 + \cdots + X_n$.

1. *Estimation de la moyenne d'une loi normale*
On suppose que X suit une loi normale de paramètres m et σ^2, où $\theta = m$ est un réel inconnu et σ un réel strictement positif connu.

a) Pour un échantillon observé donné $(x_1, x_2, \dots, x_n)$, déterminer l'expression de $L(\theta)$.

b) Étudier les variations sur $\mathbb{R}$ de la fonction qui à θ associe $\ln(L(\theta))$ et montrer que la valeur $\widehat{\theta}$ pour laquelle cette fonction atteint son maximum est :

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

c) En déduire un estimateur Y_n de m. Est-il sans biais ? La suite d'estimateurs $(Y_n)_{n\in\mathbb{N}^*}$ de m est-elle convergente ?

2. *Estimation de la variance d'une loi normale*
On suppose que X suit une loi normale de paramètres m et σ^2, où $\theta = \sigma^2$ est un réel inconnu et m un réel connu.

a) Pour un échantillon observé donné $(x_1, x_2, \dots, x_n)$, déterminer l'expression de $L(\theta)$.

b) Étudier les variations sur $]0, +\infty[$ de la fonction qui à θ associe $\ln(L(\theta))$ et montrer que la valeur $\widehat{\theta}$ pour laquelle cette fonction atteint son maximum est :

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} (x_i - m)^2$$

c) En déduire un estimateur Z_n de σ^2. Est-il sans biais ? La suite d'estimateurs $(Z_n)_{n\in\mathbb{N}^*}$ de σ^2 est-elle convergente ?

Solution :

Rappelons que si X suit la loi normale de paramètres m et σ^2, une densité f de X est : $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}$

1. a) On peut écrire :

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x_i - m}{\sigma}\right)^2} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - m}{\sigma}\right)^2\right)$$

b) Posons $\varphi(\theta) = \ln L(\theta)$. Il vient :

$$\varphi(\theta) = -n\ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\theta - x_i)^2 = -n\ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}g(\theta)$$

Ainsi, maximiser φ revient à minimiser g. Or :

$$g'(\theta) = 2\sum_{i=1}^{n}(\theta - x_i) \geqslant 0 \Longleftrightarrow \theta \geqslant \frac{1}{n}\sum_{i=1}^{n} x_i = \widehat{\theta}$$

Ainsi φ est maximale pour la moyenne empirique $\widehat{\theta}$.

c) D'après l'énoncé, $Y_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$.

Immédiatement, $E(Y_n) = m$ et $V(Y_n) = \dfrac{\sigma^2}{n}$. La suite d'estimateurs (Y_n) est sans biais et convergente.

2. a) Cette fois, pour $\theta > 0$:

$$L(\theta) = \left(\frac{1}{\theta\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\theta}\sum_{i=1}^{n}(x_i - m)^2\right)$$

b) Posons $\psi(\theta) = \ln(L(\theta))$. Il vient :

$$\psi(\theta) = -\frac{1}{2\theta}\sum_{i=1}^{n}(x_i - m)^2 - \frac{n}{2}\ln(2\pi\theta) = -\frac{1}{2}h(\theta)$$

avec :

$$h(\theta) = \frac{1}{\theta}\sum_{i=1}^{n}(x_i - m)^2 + n\ln(2\pi\theta)$$

Maximiser ψ revient à minimiser h. Comme

$$h'(\theta) = \frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n}(x_i - m)^2$$

$$h'(\theta) \geqslant 0 \Longleftrightarrow \theta \geqslant \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 = \widehat{\theta}$$

c) Un estimateur de σ^2 est donc :

$$Z_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 = \widehat{\theta}$$

Un calcul immédiat donne $E(Z_n) = \sigma^2$. C'est un estimateur sans biais.

Par indépendance des (X_i), il vient $V(Z_n) = \dfrac{1}{n}V((X-m)^2)$, ce qui entraîne que $\displaystyle\lim_{n\to+\infty} V(Z_n) = 0$. La suite d'estimateurs (Z_n) est une suite d'estimateurs sans biais et convergente de σ^2.

Exercice 3.19.

n joueurs ($n \geqslant 2$) $J_1, J_2, \ldots, J_n$ jouent l'un après l'autre dans l'ordre de leurs indices. À chaque joueur J_k est imparti un événement A_k dont la probabilité de réalisation est un réel $p_k \in]0,1[$. On notera $q_k = 1 - p_k$ et on pose $q_0 = 1$.

Le jeu se termine lorsque l'un des joueurs réalise l'événement qui lui est imparti.
Lorsqu'aucun joueur n'a gagné, on recommence un tour, puis si de nouveau aucun joueur n'a gagné, on recommence un tour, *etc.*

On note G_k l'événement «le joueur J_k gagne» et on suppose l'indépendance mutuelle de toutes les suites de résultats des coups joués.

1. Montrer que pour tout $p \in \mathbb{N}$ et tout $k \in [\![1, n]\!]$, le joueur J_k ne peut jouer qu'aux coups $pn + k$.

2. En déduire la probabilité de G_k en fonction des réels $(q_0, q_1, \ldots, q_n)$ et p_k.

3. Montrer que le jeu se termine de façon presque certaine après un nombre fini de coups.

4. On dit que le jeu est équitable si chaque joueur à la même probabilité de gagner soit $\frac{1}{n}$. Donner une condition nécessaire et suffisante sur la suite $(p_1, \ldots, p_n)$ pour que le jeu soit équitable.

5. On suppose que le jeu est équitable.
 a) Montrer que $p_1 \leqslant \frac{1}{n}$.
 b) On suppose que l'on a $p_1 = \frac{1}{n}$. Déterminer alors $p_2, \ldots, p_n$.

6. On suppose encore le jeu équitable. Soit X la variable aléatoire représentant le nombre de coups joués jusqu'à la fin du jeu. Déterminer l'espérance de X.

Solution :

1. Le joueur J_k joue (éventuellement) au coup k. Il rejouera (éventuellement) après un tour complet soit au coup $n + k$, puis (éventuellement) au coup $2n + k$, etc.

2. Notons $A_{k,p}$ l'événement «le joueur J_k gagne au coup $(np + k)$». Au vu de la définition du jeu :

$$P(A_{k,p}) = (q_0 q_1 \ldots q_n)^p (q_0 q_1 \ldots q_{k-1}) p_k$$

Comme $G_k = \bigcup_{p \geqslant 0} A_{k,p}$, il vient par indépendance :

$$P(G_k) = \sum_{p=0}^{\infty} P(A_{k,p}) = \sum_{p=0}^{\infty} (q_0 q_1 \ldots q_n)^p (q_0 q_1 \ldots q_{k-1}) p_k$$

$$P(G_k) = (q_0 q_1 \ldots q_{k-1}) p_k \times \frac{1}{1 - q_0 q_1 \ldots q_n}$$

ou

$$P(G_k) = \frac{1}{1 - q_0 q_1 \ldots q_n} (q_0 q_1 \ldots q_{k-1} - q_0 q_1 \ldots q_k)$$

3. On calcule :

$$\sum_{k=1}^{n} P(G_k) = \frac{1}{1 - q_0 q_1 \ldots q_n} \sum_{k=1}^{n} (q_0 q_1 \ldots q_{k-1} - q_0 q_1 \ldots q_k)$$

$$= \frac{1}{1 - q_0 q_1 \ldots q_n} (q_0 - q_0 q_1 \ldots q_n) = 1$$

(ne pas oublier que $q_0 = 1$)

4. On veut que pour tout $k \in [\![1, n-1]\!]$, $P(G_k) = P(G_{k+1})$, soit pour tout $k \in [\![1, n-1]\!]$:

$$q_0 \dots q_{k-1} - q_0 \dots q_k = q_0 \dots q_k - q_0 \dots q_{k+1}$$

Si l'on note $a_k = q_0 \dots q_k$, l'équation précédente s'écrit, pour tout $k \in [\![1, n-1]\!]$:

$$a_{k-1} - 2a_k + a_{k+1} = 0$$

Son équation caractéristique est $x^2 - 2x + 1 = (x-1)^2 = 0$. On sait alors qu'il existe $(\lambda, \mu) \in \mathbb{R}^2$ tels que $a_k = \lambda + \mu k$, avec :

$$\begin{cases} a_0 = q_0 = 1 = \lambda \\ a_1 = q_1 = 1 + \mu \end{cases}$$

(on peut aussi remarquer que la relation précédente caractérise les suites arithmétiques)

Ainsi, le jeu est équitable si et seulement si :

$$\text{pour tout } k \in [\![1, n-1]\!],\ a_k = 1 - k(1-q_1) = 1 - kp_1.$$

5. a) Supposons que le jeu est équitable. Alors comme $a_n = 1 - np_1 \leqslant 1$, il vient $p_1 \leqslant \dfrac{1}{n}$.

b) Si $p_1 = \dfrac{1}{n}$, alors pour tout $k \in [\![1, n-1]\!]$, $a_k = q_0 \dots q_k = 1 - \dfrac{k}{n}$.

• $a_1 = q_1 = 1 - p_1 = 1 - \dfrac{1}{n}$ et $p_1 = \dfrac{1}{n}$.

• $a_2 = q_1 q_2 = \left(1 - \dfrac{1}{n}\right) q_2 = 1 - \dfrac{2}{n}$ et $p_2 = \dfrac{1}{n-1}$.

• Supposons que pour un certain rang $i < k$, on ait $p_{i-1} = \dfrac{1}{n-i+2}$. alors :
$a_i = \left(1 - \dfrac{1}{n}\right)\left(1 - \dfrac{1}{n-1}\right) \dots \left(1 - \dfrac{1}{n-i+2}\right) q_i = 1 - \dfrac{i}{n} \implies q_i = \dfrac{n-i}{n-i-1}$
et $p_i = \dfrac{1}{n-i+1}$; on conclut par le principe de récurrence limité.

6. On a $X(\Omega) = [1, +\infty[$ et $(X = pn + k) = A_{k,p}$. On sait que

$$P(A_{k,p}) = (1-np_1)^p (1-(k-1)p_1) \frac{p_1}{1-(k-1)p_1} = (1-np_1)^p p_1$$

Donc :

$$E(X) = \sum_{p=0}^{+\infty} \sum_{k=1}^{n} (pn+k) P(A_{k,p}) = \sum_{p=0}^{+\infty} pn^2 (1-np_1)^p p_1 \times \frac{n(n+1)}{2}$$

$$E(X) = p_1 \times \frac{n^3(n+1)}{2} \sum_{p=0}^{+\infty} p(1-np_1)^p$$

Après calculs :

$$E(X) = \frac{n(n+1)}{2} \times \frac{1-np_1}{p_1}$$

Exercice 3.20.

Dans tout cet exercice, les variables aléatoires sont supposées à valeurs dans $\mathbb{R}_+$ et à densité continue sur $\mathbb{R}_+^*$.

1. Soit X une variable aléatoire telle que pour tout $x \in \mathbb{R}_+^*$, $P(X > x) > 0$.

Justifier, pour tout $x > 0$, l'existence de :

$$\varphi(x) = \lim_{h \to 0^+} \left[\frac{1}{h} \times \frac{P\big(x < X \leqslant x+h\big)}{P(X > x)}\right]$$

On appelle *taux de panne* de X la fonction φ ainsi définie sur $\mathbb{R}_+^*$.

2. Calculer le taux de panne d'une variable aléatoire Y suivant une loi exponentielle de paramètre λ.

3. Soit X une variable aléatoire à valeurs dans $\mathbb{R}_+$, de densité $f > 0$, continue sur $\mathbb{R}_+^*$, de fonction de répartition F et de taux de panne φ.
On note $G = 1 - F$.

a) Étudier, pour $a > 0$, la convergence des intégrales

$$\int_0^a \varphi(t)\,dt \text{ et } \int_a^{+\infty} \varphi(t)\,dt$$

et justifier que, pour $x > 0$, on peut définir

$$\Phi(x) = \int_0^x \varphi(t)\,dt$$

b) La fonction Φ est-elle bijective de $\mathbb{R}_+^*$ sur $\mathbb{R}_+^*$?

4. a) Déterminer la loi de la variable aléatoire $Z = \Phi(X)$.

b) En déduire la loi de la variable aléatoire $Y = F(X)$.

Solution :

1. On a : $\dfrac{1}{h}\times\dfrac{P(x < X \leqslant x+h)}{P(X > x)} = \dfrac{1}{1-F(x)}\times\dfrac{1}{h}\displaystyle\int_x^{x+h} f(t)\,dt$
donc :

$$\lim_{h\to 0}\frac{1}{h}\times\frac{P(x < X \leqslant x+h)}{P(X > x)} = \frac{f(x)}{1-F(x)}$$

puisque la fonction f est continue sur $\mathbb{R}_+^*$.

2. Si Y suit la loi $\mathcal{E}(\lambda)$, il vient, pour tout $x > 0$:

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

3. a) Les hypothèses impliquent que $\varphi \geqslant 0$, continue, donc intégrable sur tout segment inclus dans $\mathbb{R}_+^*$. Comme $\varphi = \dfrac{f}{1-F}$, une primitive de φ est $-\ln(1-F)$, et :

$$\lim_{x\to 0^+}\int_x^a \varphi(t)\,dt = -\ln(1-F(a))$$

Ainsi $\Phi(a) = \displaystyle\int_0^a \varphi(t)\,dt$ existe bien, pour tout $a > 0$.
En revanche, on a :

$$\lim_{x\to+\infty}\int_a^x \varphi(t)\,dt = \ln(1-F(a)) - \lim_{x\to+\infty}\ln(1-F(x)) = +\infty$$

et donc l'intégrale $\displaystyle\int_a^{+\infty}\varphi(t)\,dt$ diverge

b) Comme $f > 0$, la fonction φ est strictement positive et Φ est une bijection de $\mathbb{R}_+^*$ sur lui-même car elle est continue et strictement croissante et $\displaystyle\lim_{+\infty}\Phi = +\infty$.

4. a) Pour $y > 0$, il vient :

$$P(\Phi(X) \leqslant y) = P(X \leqslant \Phi^{-1}(y)) = F(\Phi^{-1}(y)) = 1 - \exp -(\Phi(\Phi^{-1}(y))) \\ = 1 - \mathrm{e}^{-y}$$

Donc X suit la loi exponentielle $\mathcal{E}(1)$.

b) On a $F = 1 - \mathrm{e}^{-\Phi}$, d'où :

$$P(Y \leqslant y) = P(1 - y \leqslant e^{-\Phi(X)}) = \begin{cases} 0 & \text{si } y < 0 \\ 1 & \text{si } y \geqslant 1 \\ 1 - \mathrm{e}^{\ln(1-y)} = y & \text{si } 0 < y < 1 \end{cases}$$

Ainsi $F(X)$ suit la loi uniforme sur $[0, 1]$.

Exercice 3.21.

Une fabrique d'appareils de précision utilise dans une de ses deux usines des machines de modèle 1 dont la caractéristique de production suit une loi normale d'espérance m_1 et d'écart-type σ_1 et dans l'autre des machines de modèle 2 dont la caractéristique de production suit une loi normale d'espérance m_2 et d'écart-type σ_2.

Pour contrôler l'homogénéité de la production, on considère d'une part n_1 variables aléatoires X_i associées à la production des machines de type 1 et d'autre part n_2 variables aléatoires Y_j associées à la production des machines de type 2, toutes ces variables aléatoires étant indépendantes.

1. Montrer que $\overline{X} = \dfrac{1}{n_1} \sum\limits_{i=1}^{n_1} X_i$ est un estimateur sans biais et convergent de m_1 et que $\overline{Y} = \dfrac{1}{n_2} \sum\limits_{j=1}^{n_2} Y_j$ est un estimateur sans biais et convergent de m_2.

En déduire un estimateur sans biais T de $\theta = m_1 - m_2$. Donner la loi de T, son espérance et sa variance.

2. On suppose $\sigma_1^2 = 5.10^{-4}$, $\sigma_2^2 = 6.10^{-4}$, $n_1 = 50$ et $n_2 = 40$.

a) Déterminer la probabilité d'avoir une estimation qui s'écarte de la valeur nulle de plus de 0,01 lorsque $m_1 = m_2$. (*)

b) même question lorsque $m_1 - m_2 = 0{,}005$. (*)

3. On règle les machines de façon que $m_1 = m_2$.

a) On suppose $m_1 = m_2$, $\sigma_1^2 = 5.10^{-4}$, $\sigma_2^2 = 6.10^{-4}$, $n_1 = 50$ et $n_2 = 40$. Déterminer a tel que $P(|T| \leqslant a) \geqslant 0{,}95$. (*)

b) Le réglage est considéré comme convenable si l'estimation de θ donné par les réalisations des deux échantillons (X_i) et (Y_j) a une valeur qui s'écarte d'au plus 0,01 de la valeur nulle. Le réglage peut-il être considéré comme convenable pour les observations suivantes : $\sum\limits_{i=1}^{50} X_i = 90{,}2$ et $\sum\limits_{j=1}^{40} Y_j = 72{,}2$?

(*) Φ étant la fonction de répartition de la loi normale réduite, on donne :

$\Phi(1) \simeq 0{,}841$, $\Phi(1,96) \simeq 0{,}975$, $\Phi(2) \simeq 0{,}977$ et $\Phi(3) \simeq 0{,}999$.

Solution :

1. Un calcul immédiat donne :

$$E(\overline{X}) = m_1, V(\overline{X}) = \frac{\sigma_1^2}{n_1} \text{ et } E(\overline{Y}) = m_2, V(\overline{Y}) = \frac{\sigma_2^2}{n_2}$$

Donc $\overline{X}$ (resp. $\overline{Y}$) est un estimateur sans biais convergent de m_1 (resp. m_2). Ainsi $T = \overline{X} - \overline{Y}$ est un estimateur sans biais de $m_1 - m_2$.
La variable T suit la loi normale de paramètres :

$$E(T) = m_1 - m_2 \text{ et } V(T) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

2. Un calcul donne $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = 0.25 \times 10^{-4}$ et $\sigma(T) = 0.5 \times 10^{-2}$.

a) Lorsque $m_1 = m_2$; T suit la loi normale $\mathcal{N}(0; 0.25 \times 10^{-4})$, et $T^\star = 200T$ suit la loi normale centrée réduite. Donc :

$$P(|T| \geqslant 0.01) = 1 - P(|T^\star| < 2) = 1 - 0.954 = 0.046$$

b) Lorsque $m_1 - m_2 = 0.005$; T suit une loi normale $\mathcal{N}(0.005; 0.25 \times 10^{-4})$, et $T^\star = 200(T - 0.005)$ suit la loi normale centrée réduite. Donc :

$$P(|T| \geqslant 0.01) = 1 - P(-3 < T^\star < 1) = 0.16$$

3. a) On sait que dans cette configuration :

$$P(|T| \leqslant a) = P(|T^\star| \leqslant 200a) = 2\Phi(200a) - 1$$

et

$$P(|T| \leqslant a) \geqslant 0.95 \iff \Phi(200a) \geqslant 0.975 \iff 200a \geqslant 1.96$$
$$\iff a \geqslant 0.0098$$

b) On a :

$$\theta = \frac{90.2}{50} - \frac{72.2}{40} = -0.001$$

Le réglage peut être considéré comme correct.

Exercice 3.22.

Un concessionnaire vend un type de véhicule selon trois modes différents, l'un de ces modes étant la vente par le biais d'un site Internet. On suppose que chacun des achats est effectué indépendamment des autres et ne porte que sur un véhicule. Pour une suite aléatoire d'achats, on s'intéresse à la probabilité que le nombre d'achats par Internet soit exactement $\frac{1}{3}$ des ventes effectuées, c'est à dire que la fréquence des choix d'achat par Internet soit $\frac{1}{3}$. On suppose que la probabilité qu'un acheteur choisisse Internet est $p \in]0,1[$.

1. On observe une suite de $3n$ achats ($n \in \mathbb{N}^*$).

a) quelle est la probabilité p_n que la fréquence d'achats sur Internet soit $\frac{1}{3}$?

b) Montrer que $\frac{p_{n+1}}{p_n} \leqslant \frac{27}{4}(1-p)^2 p$.

En déduire que $p_n \leqslant \left(\frac{27}{4}p(1-p)^2\right)^{n-1} p_1$.

c) En étudiant les variations sur $[0,1]$ de la fonction $\varphi : x \mapsto \frac{27}{4}x^2(1-x)$, montrer que si $p \neq \frac{1}{3}$, la série de terme général p_n converge.

On suppose dans la suite $p \neq \frac{1}{3}$.

2. Lors de l'observation d'une série d'achats, on note q_n la probabilité que la fréquence d'achat par Internet soit égale, pour la première fois, à $\frac{1}{3}$ lors du $(3n)^{\text{ème}}$ achat.

a) Montrer que $p_1 = q_1$ et que $\forall\, n \geqslant 2, p_n = \sum\limits_{i=1}^{n-1} q_i p_{n-i} + q_n$.

b) En déduire que pour $n \geqslant 2$, $\sum\limits_{k=1}^{n} (p_k - q_k) = \sum\limits_{2 \leqslant i+j \leqslant n} p_i q_j$, puis que :

$$\sum_{k=1}^{n} (p_k - q_k) \leqslant \Big(\sum_{i=1}^{n} p_i\Big)\Big(\sum_{j=1}^{n} q_j\Big) \leqslant \sum_{k=1}^{2n} (p_k - q_k).$$

3. On note Q_n la probabilité d'obtenir au moins une fois la fréquence $\frac{1}{3}$ d'achats par Internet au cours d'une série de $3n$ achats.

a) Quelle est la valeur de Q_n en fonction des probabilités q_k ?

b) Montrer que la série de terme général q_n est convergente.

c) On note $S = \sum\limits_{n=1}^{\infty} p_n$, montrer que $\lim\limits_{n\to\infty} Q_n = \dfrac{S}{S+1}$. Quelle est la signification de ce résultat ?

Solution :

1. a) La fréquence sera de $1/3$ si sur les $3n$ achats, n ont été effectués sur le Net, soit :

$$p_n = \binom{3n}{n} p^n (1-p)^{2n}.$$

b) Il vient :

$$\frac{p_{n+1}}{p_n} = \frac{(3n+3)(3n+2)(3n+1)}{(n+1)(2n+1)(2n+2)} p(1-p)^2 \leqslant 3 \times \frac{3}{2} \times \frac{3}{2} p(1-p)^2 = \frac{27}{4} p(1-p)^2$$

Donc $p_{n+1} \leqslant \frac{27}{4} p(1-p)^2 p_n$ et une récurrence immédiate donne :

$$\forall\, n \geqslant 1, p_n \leqslant \Big(\frac{27}{4} p(1-p)^2\Big)^{n-1} p_1$$

c) La fonction $\varphi : x \mapsto \frac{27}{4} x(1-x)^2$ est maximale pour $x = 1/3$, le maximum étant égal à 1. Donc lorsque $p \neq 1/3$, $\sup |\varphi(x)| < 1$ et, par la règle de majoration des séries à termes positifs, la série $\sum p_n$ converge.

2. a) Au bout de $3n$ achats, la fréquence vaut $1/3$ si et seulement si l'un des événements incompatibles suivants est réalisé :

« $1/3$ des achats ont été effectués sur le Net pour la première fois au bout de $3k$ achats et lors des $3(n-k)$ achats suivants, la fréquence des achats sur le Net est de $1/3$ », k décrivant $[\![1, n]\!]$ et la deuxième partie de l'événement disparaissant si $k = n$.

Ce qui donne

$$p_1 = q_1 \text{ et pour } n \geqslant 2, p_n = \sum_{k=1}^{n-1} q_k p_{n-k} + q_n$$

b) Ainsi :

$$\sum_{k=1}^{n}(p_k-q_k)=\sum_{k=2}^{n}(p_k-q_k)=\sum_{k=2}^{n}\Big(\sum_{i=1}^{k-1}q_ip_{k-i}\Big)=\sum_{2\leqslant i+j\leqslant n}p_iq_j$$

or

$$\Big(\sum_{i=1}^{n}p_i\Big)\times\Big(\sum_{i=1}^{n}q_i\Big)=\sum_{2\leqslant i+j\leqslant 2n}p_iq_j$$

Donc, par positivité des objets concernés :

$$\sum_{k=1}^{n}(p_k-q_k)\leqslant\Big(\sum_{i=1}^{n}p_i\Big)\times\Big(\sum_{i=1}^{n}q_i\Big)\leqslant\sum_{k=1}^{2n}(p_k-q_k)$$

3. a) On a immédiatement $Q_n=q_1+q_2+\cdots+q_n$.

b) Comme Q_n est une probabilité, $\sum_{i=1}^{n}q_i\leqslant 1$, ce qui entraîne la convergence de la série $\sum q_i$.

c) Posons $S=\sum_{n=1}^{\infty}p_n$ et $Q=\sum_{n=1}^{\infty}q_n$.

Par le résultat de la question précédente, $S-Q=SQ$ et $Q=\dfrac{S}{S+1}$, donc $Q<1$.

Ainsi, si $p\neq 1/3$, il n'est pas quasi-certain que l'on ait au moins à un moment un tiers des achats effectués sur le Net.

Exercice 3.23.

Soit n un entier naturel non nul. Une urne contient initialement n boules indiscernables au toucher et numérotées de 1 à n. On effectue une succession de tirages d'une boule de cette urne selon le protocole suivant :

tant que l'urne n'est pas vide, si à un rang quelconque on obtient la boule portant le numéro k, alors on enlève de l'urne toutes les boules dont le numéro est supérieur ou égal à k avant de procéder au tirage suivant.

On note X_n la variable aléatoire égale au nombre de tirages juste nécessaires pour vider l'urne de toutes ses boules et on note u_n l'espérance de X_n.

1. a) Déterminer les lois de X_1, X_2 et X_3.

b) Calculer u_1, u_2 et u_3.

2. Montrer que pour $n\geqslant 2$, on a $u_n=\dfrac{1}{n}\sum_{i=1}^{n-1}u_i+1$.

3. En déduire que pour $n\geqslant 2$, $u_n=u_{n-1}+\dfrac{1}{n}$.

4. Donner un équivalent de u_n au voisinage de l'infini.

Solution :

1. On a $X_1(\Omega)=\{1\}$, $P(X_1=1)=1$.

De même $X_2(\Omega)=\{1,2\}$, $P(X_2=1)=\dfrac{1}{2}=P(X_2=2)=2$.
Ainsi :

$$E(X_1)=1,\ E(X_2)=\frac{3}{2}.$$

Enfin, $X_3(\Omega) = \{1, 2, 3\}$ et $P(X_3 = 1) = \dfrac{1}{3}$, $P(X_3 = 3) = \dfrac{1}{6}$ et $P(X_3 = 2) = \dfrac{1}{2}$. Ainsi :

$$E(X_3) = \frac{11}{6}.$$

2. On sait que $X_n(\Omega) = \{1, 2, \dots, n\}$ et que $P(X_n = 1) = \dfrac{1}{n}$.

Notons T la variable aléatoire représentant le numéro de la boule sortie au premier tirage. La variable T suit la loi uniforme sur $[\![1, n]\!]$, et, pour $k \geqslant 2$:

$$\begin{aligned} P(X_n = k) &= \sum_{i=2}^{n} P(X_n = k/T = i)P(T = i) = \sum_{i=2}^{n} \frac{1}{n} \times P(X_n = k/T = i) \\ &= \sum_{i=2}^{n} \frac{1}{n} \times P(X_{i-1} = k - 1) = \sum_{i=k-1}^{n-1} \frac{1}{n} \times P(X_i = k - 1) \end{aligned}$$

D'où :

$$\begin{aligned} u_n &= \sum_{k=1}^{n} kP(X_n = k) = \frac{1}{n} + \sum_{k=2}^{n} k\Big(\sum_{i=k-1}^{n-1} \frac{1}{n} \times P(X_i = k - 1)\Big) \\ &= \frac{1}{n} + \sum_{k=2}^{n} \sum_{i=k-1}^{n-1} \frac{k}{n} P(X_i = k - 1) = \frac{1}{n} + \sum_{i=1}^{n-1} \sum_{k=1}^{i} \frac{k+1}{n} P(X_i = k) \\ &= \frac{1}{n} + \frac{1}{n} \sum_{i=1}^{n-1} E(X_i + 1) = \frac{1}{n} \sum_{i=1}^{n-1} u_i + 1 \end{aligned}$$

3. On a :

$$nu_n = n + \sum_{i=1}^{n-1} u_i, \quad \text{et} \quad (n-1)u_{n-1} = (n-1) + \sum_{i=1}^{n-2} u_i$$

d'où, en soustrayant : $u_n = u_{n-1} + \dfrac{1}{n}$.

4. Ainsi $u_k - u_{k-1} = \dfrac{1}{k}$, et en sommant, $u_n - u_1 = \displaystyle\sum_{k=2}^{n} \frac{1}{k}$, donc $u_n \sim \ln n$, cette équivalence classique se démontrant par comparaison série–intégrale.

Exercice 3.24.

Une particule se déplace dans un plan rapporté à un repère orthonormé d'origine O, de façon rectiligne.

— Elle part de la position initiale $P_0 = (0, 0)$. Elle parcourt une distance d dans une direction faisant un angle aléatoire θ_1 par rapport à l'axe des abscisses, θ_1 suivant la loi uniforme sur le segment $[-\pi, \pi]$. Elle atteint alors la position $P_1 = (x_1, y_1)$

— Ensuite, l'opération recommence : la particule repart de P_1 avec les mêmes règles, elle parcourt une distance d dans une direction faisant un angle aléatoire θ_2 par rapport à l'axe des abscisses, θ_2 suivant la loi uniforme sur le segment $[-\pi, \pi]$.

— Et ainsi de suite.

Soient X_n, Y_n et ρ_n, les variables aléatoires correspondant respectivement à l'abscisse, à l'ordonnée de la particule et à la distance OP_n.

On suppose enfin que les variables aléatoires θ_i sont mutuellement indépendantes, et suivent toutes la même loi uniforme sur $[-\pi, \pi]$

1. Montrer que :

$$X_n = d(\cos(\theta_1) + \cos(\theta_2) + \ldots + \cos(\theta_n))$$
$$Y_n = d(\sin(\theta_1) + \sin(\theta_2) + \ldots + \sin(\theta_n))$$
$$\rho_n{}^2 = X_n^2 + Y_n^2$$

2. Montrer que : $\rho_n{}^2 = d^2(n + \sum\limits_{i\neq j} \cos(\theta_i)\cos(\theta_j) + \sum\limits_{i\neq j} \sin(\theta_i)\sin(\theta_j))$.

Dans la suite, on pose $T_n = \sum\limits_{i\neq j} \cos(\theta_i)\cos(\theta_j)$ et $U_n = \sum\limits_{i\neq j} \sin(\theta_i)\sin(\theta_j))$.

3. a) Calculer l'espérance et la variance de X_n et de Y_n.

b) Calculer l'espérance de T_n et de U_n.

c) En déduire l'espérance de $\rho_n{}^2$.

4. En appliquant le théorème de la limite centrée, montrer que la probabilité de l'événement $[-d\sqrt{2n} \leqslant X_n \leqslant d\sqrt{2n}]$ a une limite peu différente de 0.9555 lorsque n tend vers $+\infty$.
Même question pour Y_n.
[On donne $\Phi(2) \simeq 0.97725$, Φ désignant la fonction de répartition d'une variable suivant la loi normale centrée réduite.]

Solution :

1. Notons $\overrightarrow{OP_n} = \sum\limits_{i=1}^{n-1} \overrightarrow{OP_i}$; alors X_n (resp. Y_n) est l'abscisse (resp. l'ordonnée) de $\overrightarrow{OP_n}$. Comme
$$\overrightarrow{P_{n-1}P_n} = d(\cos\theta_n \overrightarrow{\imath} + \sin\theta_n \overrightarrow{\jmath})$$
il vient :
$$X_n = d \times \sum_{i=1}^{n} \cos\theta_i \text{ et } Y_n = d \times \sum_{i=1}^{n} \sin\theta_i$$
et :
$$\rho_n^2 = d^2(X_n^2 + Y_n^2)$$

2. Immédiatement :
$$\Big(\sum_{i=1}^{n} \cos\theta_i\Big)^2 + \Big(\sum_{i=1}^{n} \sin\theta_i\Big)^2 = \sum_{i=1}^{n} (\cos^2\theta_i + \sin^2\theta_i) + \sum_{i\neq j} \cos\theta_i\cos\theta_j + \sum_{i\neq j} \sin\theta_i \sin\theta_j$$

Comme $\cos^2 + \sin^2$ est la fonction constante égale à 1, on en déduit le résultat demandé.

3. a) Par le théorème de transfert :
$$E(\cos\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos t\, dt = 0, E(\sin\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin t\, dt = 0$$
Par linéarité de l'espérance :
$$E(X_n) = E(Y_n) = 0.$$
De même :
$$E(\cos^2\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos^2 t\, dt = \frac{1}{2}, E(\sin^2\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sin^2 t\, dt = \frac{1}{2}$$
et
$$V(\cos\theta) = V(\sin\theta) = \frac{1}{2}$$

Donc, par indépendance :

$$V(X_n) = d^2 n V(\cos\theta) = \frac{d^2}{2} = V(Y_n)$$

b) Par indépendance :

$$E(T_n) = \sum_{i\neq j} E(\cos\theta_i)E(\cos\theta_j) = 0, \text{ et } E(U_n) = \sum_{i\neq j} E(\sin\theta_i)E(\sin\theta_j) = 0$$

et

$$E(\rho_n^2) = nd^2$$

4. Par le théorème de la limite centrée, pour tout $a < b$:

$$P\Big(a < \frac{\frac{X_n}{n}}{\frac{d\sqrt{2}}{2\sqrt{n}}} < b\Big) \underset{n\to\infty}{\longrightarrow} \Phi(b) - \Phi(a)$$

d'où :

$$P\Big(- bd\frac{\sqrt{2n}}{2} < X_n < bd\frac{\sqrt{2n}}{2}\Big) \underset{n\to\infty}{\longrightarrow} 2\Phi(b) - 1$$

et $2\Phi(2) - 1 \approx 0.9555$.

Les variables aléatoires X_n et Y_n ayant même espérance et même variance, on obtient le même résultat pour Y_n.

Exercice 3.25.

Soit (A, B) un couple de variables aléatoires discrètes, définies sur un espace probabilisé $(\Omega, \mathcal{B}, P)$, à valeurs dans $\mathbb{N}\times\mathbb{N}$, telles que, pour tout (i, j) de $\mathbb{N}^2$, on a :

$$P[(A = i)\bigcap(B = j)] = C\times\frac{\mathrm{e}^{-i}}{j^2 + 3j + 2}.$$

1. a) Montrer que $C = 1 - \mathrm{e}^{-1}$.

b) Déterminer la loi de A, préciser son espérance et sa variance.

c) Les variables aléatoires A et B sont-elles indépendantes ?

2. Soit Z la variable aléatoire égale à $5A + 7B$.

a) Ecrire un programme turbo-pascal qui, n étant donné, calcule $P(Z = n)$.

b) Calculer la probabilité de l'événement $(Z = 23)$.

c) Montrer que, pour $n > 23$, l'événement $(Z = n)$ n'est jamais quasi-impossible.

Solution :

1. a) On a :

$$1 = C\sum_{j=0}^{\infty}\frac{1}{j^2 + 3j + 2}\Big(\sum_{i=0}^{\infty}\mathrm{e}^{-i}\Big) = \frac{C}{1 - \mathrm{e}^{-1}}\sum_{j=0}^{\infty}\Big(\frac{1}{j+1} - \frac{1}{j+2}\Big)$$

et comme $\sum_{j=0}^{n}\Big(\frac{1}{j+1} - \frac{1}{j+2}\Big) = 1 - \frac{1}{n+2} \underset{n\to\infty}{\longrightarrow} 0$, il vient :

$$C = 1 - \mathrm{e}^{-1}.$$

b) La loi de A est donnée par, pour tout $i \in \mathbb{N}$:

$$P(A = i) = \sum_{j=0}^{\infty} P(A = i \cap B = j) = \mathrm{e}^{-i}(1 - \mathrm{e}^{-1})$$

Donc $A+1$ suit la loi géométrique de paramètre $p = 1 - \mathrm{e}^{-1}$, et :

$$E(A) = \frac{1}{\mathrm{e}-1} \text{ et } V(A) = \frac{\mathrm{e}}{(\mathrm{e}-1)^2}.$$

c) On a $B(\Omega) = \mathbb{N}$ et pour tout $j \in \mathbb{N}$:

$$P(B = j) = \frac{1}{j^2 + 3j + 2}$$

ce qui entraîne que A et B sont indépendantes.

2. a) Définissons d'abord la fonction f sur $\mathbb{N}$ telle que $f(k) = \mathrm{e}^k$. Par exemple :

```
function f(n : integer) : real ;
const e= 2.71828183
var k : integer ; x : real ;
Begin
x := 1 ;
For k=0 to n do x :=x*e ;
f := x
end ;
```

puis

```
function pz( n : integer) : real ;
var b : integer ; p : real ;
Begin
p := 0 ;
For b := 0 to int(n/7) do
If int((n-7*b)/4)=(n-7*b)/4 then
p :=p+(1-1/e)*f((n-7*b)/4)/(b*b+3*b+2) ;
pz := p
end ;
```

b) On a $P(Z = 23) = 0$: il suffit d'essayer les valeurs possibles de B qui sont $0, \dots, 3$, pour voir que $23 - 7B$ n'est pas un multiple de 5.

c) On remarque que :

$$24 = 5{\times}2 + 7{\times}2\,;\ 25 = 5{\times}5 + 7{\times}0\,;\ 26 = 5{\times}1 + 7{\times}3$$
$$27 = 5{\times}4 + 7{\times}1\,;\ 28 = 5{\times}0 + 7{\times}4$$

Soit $n \geqslant 29$, supposons que pour tout $k \in [\![24, n]\!]$, on a $P(Z = k) > 0$, alors, comme

$$5a + 7b = n - 4 \implies 5(a+1) + 7b = n + 1,$$

on a $P(Z = n+1) \geqslant P(Z = n-4) > 0$.
Donc $n+1$ est aussi une valeur accessible pour Z. On conclut par le principe de récurrence.

Exercice 3.26.

1. Soit m un entier strictement positif et X une variable aléatoire à valeurs dans $\{0, \dots, m\}$ définie sur un espace probabilisé $(\Omega, \mathcal{B}, P)$. On appelle fonction génératrice de X la fonction de la variable réelle $t \in \mathbb{R}^+$ définie par :

$$G_X(t) = \sum_{k=0}^{m} t^k P(X = k).$$

Justifier la formule $G_X(t) = E(t^X)$, où E désigne l'opérateur « espérance ». Montrer que l'on a :

$$E(X) = G'_X(1) \text{ et } V(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2,$$
où V désigne l'opérateur « variance ».

2. Soient $(X_k)_{1\leqslant p\leqslant n}$, n variables aléatoires indépendantes, définies sur Ω, de même loi et toutes à valeurs dans $\{0,\dots,\ell\}$ avec $\ell \in \mathbb{N}^*$. On considère une variable aléatoire N, définie sur Ω, indépendante des variables X_k et à valeurs dans $\{1,\dots,n\}$. On définit une variable aléatoire Y par :

$$Y(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega)$$

pour tout $\omega \in \Omega$.

a) Déterminer la fonction génératrice G_Y de Y en fonction de celles de N et de $X = X_1$ (on pourra utiliser la formule de l'espérance totale).

b) En déduire l'expression de l'espérance et de la variance de Y en fonction de celles de N et de X_1.

3. 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, avec $0 < p < 1$. 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 piles obtenus.

Solution :

1. Par le théorème de transfert, on a $E(t^X) = G_X(t)$.

La fonction G_X étant polynomiale, elle est de classe C^∞ et, après dérivations :
$$E(X) = G'_X(1), E(X(X-1)) = G''_X(1)$$
d'où :
$$E(X) = G'_X(1), V(X) = G''_X(1) + G'_X(1) - G'^2_X(1)$$

2. a) La famille $(N = k)_{1\leqslant k\leqslant n}$ forme un système complet d'événements. On lui applique la formule de l'espérance totale :
$$G_Y(t) = \sum_{k=1}^{n} E(t^Y/N = k)P(N = k)$$
Par indépendance de la variable N et des variables $(X_1,\dots,X_n)$, on a :
$$E(t^Y/N = k) = E(t^{X_1+\cdots+X_k}) = \prod_{i=1}^{k} E(t^{X_i}) = (E(t^{X_1}))^k = (G_X(t))^k$$
Finalement :
$$G_Y(t) = \sum_{k=1}^{n} [G_X(t)]^k P(N = k) = G_N(G_X(t))$$
b) Par la formule précédente :
$$E(Y) = G'_Y(1) = G'_N(G_X(1))\times G'_X(1) = E(N)\times E(X)$$
D'autre part : $G''_Y(t) = G''_N(G_X(t))(G'_X(t))^2 + G'_N(G_X(t))G''_X(t)$, donne en 1 :
$$G''_Y(1) = G''_N(1)(G'_X(1))^2 + G'_N(1)G''_X(1)$$
et comme $V(Y) = G''_Y(1) + G'_Y(1) - G'^2_Y(1)$, il vient :

$$V(Y) = E(N(N-1))[E(X)]^2 + E(N)E(X(X-1)) + E(N)E(X) - [E(N)]^2[E(X)]^2$$

Soit, après calculs :

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

3. Notons N la variable aléatoire représentant le numéro du jeton tiré, et $X_1, \ldots, X_n$, n variables aléatoires associées au lancer de la pièce, qui suivent donc la loi de Bernoulli de paramètre p.

Le nombre de piles obtenu est donné par la variable aléatoire $Y = \sum\limits_{k=1}^{N} X_k$.

La variable N suit la loi uniforme sur $[\![1, n]\!]$ et :

$$E(Y) = E(N) \times E(X) = \frac{p(n+1)}{2}$$

$$V(Y) = E(N) \times V(X) + V(N) \times E(X) = \frac{(n+1)pq}{2} + \frac{(n^2-1)p^2}{12}$$

soit :

$$V(Y) = \frac{p(n+1)(6+np-7p)}{12}$$

Exercice 3.27.

A. Soit $n \in \mathbb{N}^*$. On considère une variable aléatoire réelle X_n de loi exponentielle de paramètre $1/n$, et on définit la variable aléatoire Y_n, par $Y_n = \lfloor X_n \rfloor$, où $\lfloor x \rfloor$ désigne la partie entière du réel x.

1. Préciser les valeurs prises par Y_n et déterminer la loi de Y_n.

2. Calculer l'espérance $E(Y_n)$ et la variance $V(Y_n)$.

B. Pour tout n de $\mathbb{N}^*$, on pose $Z_n = X_n - Y_n$. On définit ainsi une suite $(Z_n)_{n \geqslant 1}$ de variables aléatoires réelles.

1. Montrer que pour tout $n \in \mathbb{N}^*$, Z_n est une variable aléatoire à densité et déterminer une densité de Z_n.

2. Calculer l'espérance $E(Z_n)$. Montrer que la suite $(E(Z_n))_{n \geqslant 1}$ admet une limite que l'on déterminera.

3. Montrer que la suite $(Z_n)_{n \geqslant 1}$ converge en loi vers une variable aléatoire U dont on donnera la loi.

Solution :

A. 1. On a $Y_n(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}$:

$$P(Y_n = k) = P(k \leqslant X_n < k+1) = \int_k^{k+1} \frac{1}{n} \, \mathrm{e}^{-x/n} \, dx = \mathrm{e}^{-k/n}(1 - \mathrm{e}^{-1/n})$$

2. Les résultats précédents montrent que $Y_n + 1$ suit la loi géométrique de paramètre $1 - \mathrm{e}^{-1/n}$ et donc :

$$E(Y_n) = \frac{\mathrm{e}^{-1/n}}{1 - \mathrm{e}^{-1/n}}, \quad V(Y_n) = \frac{\mathrm{e}^{-1/n}}{(1 - \mathrm{e}^{-1/n})^2}$$

B. 1. On a $Z_n(\Omega) = [0,1[$ et, pour $z \in [0,1[$:

$$(Z_n \leqslant z) = \bigcup_{k=0}^{\infty} (k \leqslant X_n \leqslant k+z)$$

Par incompatibilité :

$$P(Z_n \leqslant z) = \sum_{k=0}^{\infty} P(k \leqslant X_n \leqslant z+k) = \sum_{k=0}^{\infty} \left(\int_k^{k+z} \frac{1}{n} \mathrm{e}^{-x/n}\, dx \right)$$
$$= \sum_{k=0}^{\infty} \left(\mathrm{e}^{-k/n} - \mathrm{e}^{-(k+z)/n} \right) = (1 - \mathrm{e}^{-z/n}) \sum_{k=0}^{\infty} \mathrm{e}^{-k/n} = \frac{1 - \mathrm{e}^{-z/n}}{1 - \mathrm{e}^{-1/n}}$$

et, bien évidemment, pour $z < 0$, $P(Z_n \leqslant z) = 0$ et pour $z > 1$, $P(Z_n \leqslant z) = 1$.

Finalement :

$$F(z) = P(Z_n \leqslant z) = \begin{cases} 0 & \text{si } z < 0 \\ \dfrac{1 - \mathrm{e}^{-z/n}}{1 - \mathrm{e}^{-1/n}} & \text{si } z \in [0,1[\\ 1 & \text{si } z > 1 \end{cases}$$

On vérifie ensuite facilement que la fonction $F : z \mapsto P(Z_n \leqslant z)$ est une fonction de répartition en étudiant chacune des propriétés définissant une telle fonction.

Une densité est alors :

$$f(z) = \begin{cases} 0 & \text{si } z \notin [0,1[\\ \dfrac{\mathrm{e}^{-z/n}}{n(1 - \mathrm{e}^{-1/n})} & \text{si } z \in [0,1[\end{cases}$$

2. Un calcul immédiat donne :

$$E(Z_n) = \frac{1}{1 - \mathrm{e}^{-1/n}} \int_0^1 \frac{z.\mathrm{e}^{-z/n}}{n}\, dz = n - \frac{\mathrm{e}^{-1/n}}{1 - \mathrm{e}^{-1/n}}$$

$$n - \frac{\mathrm{e}^{-1/n}}{1 - \mathrm{e}^{-1/n}} = \frac{n - n\mathrm{e}^{-1/n} - \mathrm{e}^{-1/n}}{1 - \mathrm{e}^{-1/n}}$$
$$= \frac{n - n(1 - \frac{1}{n} + \frac{1}{2n^2} + o(\frac{1}{n^2})) - (1 - \frac{1}{n} + o(\frac{1}{n}))}{\frac{1}{n} + o(\frac{1}{n})}$$
$$= \frac{\frac{1}{2n} + o(\frac{1}{n})}{\frac{1}{n} + o(\frac{1}{n})} \underset{n \to \infty}{\longrightarrow} \frac{1}{2}$$

Soit :

$$\lim_{n \to +\infty} E(Z_n) = \frac{1}{2}.$$

3. Pour tout $z \in [0,1[$ fixé, en utilisant à nouveau les développements limités (ici il s'agit simplement d'équivalents) :

$$\lim_{n \to +\infty} P(Z_n \leqslant z) = \lim_{n \to +\infty} \frac{1 - \mathrm{e}^{-z/n}}{1 - \mathrm{e}^{-1/n}} = z$$

On vient de montrer que $(Z_n)_n$ converge en loi vers une variable qui suit la loi uniforme sur $[0,1[$.

Exercice 3.28.

Une grande boule transparente contient des paillettes roses et bleues bien mélangées en proportion $\frac{1}{3}$ de roses et $\frac{2}{3}$ de bleues.

Un jeu consiste à faire tourner la boule, ce qui déclenche, après quelques tours, une sortie aléatoire de N paillettes.

N est une variable aléatoire à valeurs dans $\mathbb{N}$ et on note pour tout $n \in \mathbb{N}$, $\alpha_n = P(N = n)$.

Le joueur sépare ensuite les paillettes selon leur couleur : R représente le nombre de paillettes roses et $B = N - R$ représente le nombre de paillettes bleues.

1. Dans cette question les paillettes sont «grosses» et N suit la loi uniforme sur l'ensemble $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.

Quelle est la probabilité que le joueur obtienne exactement deux fois plus de paillettes bleues que de paillettes roses ? (on pourra remarquer que l'événement $(B = 2R)$ est inclus dans l'événement (N est multiple de 3)).

2. Dans cette question les paillettes sont «petites» et notre modèle suppose que N suit une loi discrète infinie : $N(\Omega) = \mathbb{N}$ et pour tout $a \in \mathbb{N}$, $P(N > a) \neq 0$.

L'animateur prétend que B et R sont indépendantes. Il s'agit de vérifier si cela est possible.

a) Justifier que pour tout $h \in \mathbb{N}$, $P(B = h) = \left(\frac{2}{3}\right)^h \frac{1}{h!} \sum_{n=h}^{+\infty} \frac{n!\alpha_n}{(n-h)!} \left(\frac{1}{3}\right)^{n-h}$

b) On pose :

- pour tout $n \in \mathbb{N}$, $u_n = n!\alpha_n$.
- Pour tout $h \in \mathbb{N}$, $\varphi(h) = \sum_{n=h}^{+\infty} \frac{u_n}{(n-h)!} \left(\frac{1}{3}\right)^{n-h}$.
- Pour tout $k \in \mathbb{N}$, $\psi(k) = \sum_{n=k}^{+\infty} \frac{u_n}{(n-k)!} \left(\frac{2}{3}\right)^{n-k}$.

Démontrer que l'indépendance de R et B se traduit par :

$$\forall (h, k) \in \mathbb{N}^2, u_{h+k} = \varphi(h) \times \psi(k)$$

c) En déduire que la suite (u_n) doit être géométrique de raison $\lambda = \dfrac{\varphi(1)}{\varphi(0)}$, puis que la loi de N doit être la loi de Poisson de paramètre λ.

d) On suppose que N suit la loi de Poisson de paramètre λ. Vérifier que l'animateur est crédible.

Solution :

1. On a l'égalité :

$(B = 2R) = ((N = 0) \cap (R = 0)) \cup ((N = 3) \cap (R = 1)) \cup ((N = 6) \cap (R = 2))$

Ces trois événements sont incompatibles. On calcule chacune des probabilités en utilisant la loi conditionnelle de R conditionnée par $(N = k)$, qui est la loi binomiale $\mathcal{B}(k, 1/3)$. Ainsi :

$$P(B = 2R) = \sum_{j=0}^{2} P(N = 3j) \binom{3j}{j} \left(\frac{1}{3}\right)^j \left(\frac{2}{3}\right)^{2j} = \frac{431}{2187}$$

2. a) La famille $(N = n)_{n \geqslant 0}$ est un système complet d'événements. Comme ci-dessus :

$$P(R = r / N = n) = \binom{n}{r} \left(\frac{1}{3}\right)^r \left(\frac{2}{3}\right)^{n-r}$$

et

$$P(B = h/N = n) = \binom{n}{h}\left(\frac{2}{3}\right)^h\left(\frac{1}{3}\right)^{n-h}$$

Ainsi, pour tout $h \in \mathbb{N}$:

$$P(B = h) = \sum_{n=0}^{\infty} P(N = n)\binom{n}{h}\left(\frac{2}{3}\right)^h\left(\frac{1}{3}\right)^{n-h} = \sum_{n=h}^{\infty} \alpha_n \frac{n!}{h!(n-h)!}\left(\frac{2}{3}\right)^h\left(\frac{1}{3}\right)^{n-h}$$

$$P(B = h) = \left(\frac{2}{3}\right)^h \frac{1}{h!} \sum_{n=h}^{\infty} \frac{n!\alpha_n}{(n-h)!}$$

b) De manière identique, pour tout $k \in \mathbb{N}$:

$$P(R = k) = \left(\frac{1}{3}\right)^k \frac{1}{k!} \sum_{n=k}^{\infty} \frac{n!\alpha_n}{(n-k)!}$$

Or on a $((B = h) \cap (R = k)) = ((N = h + k) \cap (B = h))$, donc :

$$\begin{aligned} P((B = h) \cap (R = k)) &= P(N = h + k)\binom{h+k}{h}\left(\frac{2}{3}\right)^h\left(\frac{1}{3}\right)^k \\ &= \frac{u_{h+k}}{h!k!}\left(\frac{2}{3}\right)^h\left(\frac{1}{3}\right)^k \end{aligned}$$

Donc l'indépendance des variables aléatoires R et B se traduit par :

$$\forall (h, k), u_{h+k} = \varphi(h) \times \psi(k)$$

c) En particulier pour $h = 0$ et $k = n$, il vient $u_n = \varphi(0)\psi(n)$ et pour $h = 1$ et $k = n$, il vient $u_{n+1} = \varphi(1)\psi(n)$.

Par l'hypothèse « pour tout $a \in \mathbb{N}$, $P(N > a) \neq 0$ », les fonctions φ et ψ sont strictement positives et, pour tout $n \in \mathbb{N}$:

$$\frac{u_{n+1}}{u_n} = \frac{\varphi(1)}{\psi(0)} = \lambda$$

Le suite (u_n) est donc géométrique, et pour tout $n \in \mathbb{N}$:

$$P(N = n) = \alpha_n = u_0 \frac{\lambda^n}{n!}$$

La condition $\sum_{n=0}^{\infty} P(N = n) = 1$ impose que $u_0 = \mathrm{e}^{-\lambda}$, et N suit la loi de Poisson de paramètre λ. B et R sont indépendantes.

d) Réciproquement, si N suit la loi de Poisson de paramètre λ, les formules ci-dessus montrent que R suit la loi de Poisson de paramètre $\frac{\lambda}{3}$ et que B suit la loi de Poisson de paramètre $\frac{2\lambda}{3}$, et que pour tout $(h, k) \in \mathbb{N}^2$,

$$P((B = h) \cap (R = k)) = P(B = h) \times P(R = k) :$$

B et R sont indépendantes.

Exercice 3.29.

Soit $n \geqslant 2$ et $p \in]0, 1[$. Soient X, Y, Z trois variables aléatoires indépendantes définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et suivant chacune la loi binomiale $\mathcal{B}(n, p)$.

Pour tout $\omega \in \Omega$, on note $M(\omega)$ la matrice $M(\omega) = \begin{pmatrix} X(\omega) & X(\omega) & X(\omega) \\ Y(\omega) & Y(\omega) & Y(\omega) \\ Z(\omega) & Z(\omega) & Z(\omega) \end{pmatrix}$.

1. La matrice M peut-elle être inversible ?

2. Déterminer la probabilité que M soit la matrice d'un projecteur.

3. Soit T la variable aléatoire qui à tout $\omega \in \Omega$ associe le cardinal de l'ensemble des valeurs propres de $M(\omega)$.

a) Montrer que $T(\Omega) = \{1, 2\}$.

b) Déterminer la loi de T et son espérance.

c) Calculer la probabilité que la matrice M soit diagonalisable.

4. On suppose dans cette question que $p = 1/2$.

Déterminer la probabilité pour qu'une ligne de M soit égale à la somme des deux autres lignes de M.

Solution :

1. Pour tout $\omega \in \Omega$, la matrice $M(\omega)$ est une matrice de rang 1. Donc

$$P(M(\omega) \text{ inversible}) = 0.$$

2. La matrice $M(\omega)$ est celle d'un projecteur si et seulement si $M^2(\omega) = M(\omega)$. Or :

$$M^2(\omega) = (X + Y + Z)(\omega)M(\omega)$$

Donc la probabilité cherchée est celle que $X + Y + Z = 1$.
Par indépendance des variables aléatoires, $X + Y + Z$ suit la loi binomiale $\mathcal{B}(3n, p)$ et :

$$P(M^2(\omega) = M(\omega)) = 3n{\times}p(1 - p)^{3n-1}$$

3. a) A priori $T(\Omega) \subseteq \{0, 1, 2, 3\}$. Mais comme la matrice $M(\omega)$ est de rang 1, on sait que 0 est valeur et que le sous-espace propre associé est de dimension 2. Donc $T(\Omega) \subseteq \{1, 2\}$.

b) On a vu dans la question précédente que $M^2(\omega) = (X+Y+Z)(\omega)M(\omega)$. Donc

• Si $(X + Y + Z)(\omega) = 0$. Dans ce cas, $M^2(\omega) = 0$ et $T(\omega) = 1$ (la seule valeur propre est 0).

• Si $(X + Y + Z)(\omega) \neq 0$. Dans ce cas, $T(\omega) = 2$ (la matrice $M(\omega)$ admet deux valeurs propres qui sont 0 et $(X + Y + Z)(\omega)$) et :

$$\begin{cases} P(T = 1) = P(X + Y + Z = 0) = p^{3n} \\ P(T = 2) = P(X + Y + Z \neq 0) = 1 - p^{3n} \end{cases}$$

Finalement $E(T) = 2 + p^{3n}$.

c) La matrice $M(\omega)$ est diagonalisable si et seulement si $T(\omega) = 2$ ou $(T(\omega) = 1$ et $M(\omega) = 0)$. Ainsi :

$P(M(\omega) \text{ diagonalisable }) = 1-p^{3n}+P(X = Y = Z = 0) = 1-p^{3n}+(1-p)^{3n}$

4. L'événement $X = Y + Z$ est l'événement $\bigcup\limits_{k=0}^{n} ((K = k) \cap (Y + Z = k))$.
Donc :

$$\begin{aligned} P(X = Y + Z) &= \sum_{k=0}^{n} P(X = k)P(Y + Z = n - k) \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \binom{2n}{k} p^k q^{n-k} \end{aligned}$$

$$= \left(\frac{1}{2}\right)^{2n} \sum_{k=0}^{n} \binom{n}{k}\binom{2n}{k} = \left(\frac{1}{2}\right)^{2n} \binom{3n}{n}$$

La dernière égalité (de Vandermonde) se démontrant en comparant les coefficients de x^n dans $(1+x)^{3n}$ et $(1+x)^n(1+x)^{2n}$.

Exercice 3.30.

Soient $X_1, X_2, \ldots, X_n$, n variables aléatoires indépendantes suivant la même loi exponentielle de paramètre 1.

On pose $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $X_n^* = \sqrt{n}(\overline{X_n} - 1)$.

1. a) Donner l'espérance et la variance de $\overline{X_n}$.
   b) Montrer que $\overline{X_n}$ converge en probabilité vers 1.

2. On note F_n la fonction de répartition de X_n^* .
   a) Quelle est, pour $x \in \mathbb{R}$, la limite de $F_n(x)$ lorsque n tend vers $+\infty$?
   b) Calculer une valeur approchée de $P(1 - \frac{2}{\sqrt{n}} \leqslant \overline{X_n} \leqslant 1 + \frac{2}{\sqrt{n}})$, pour n assez grand.

(Si Φ la fonction de répartition de la loi normale réduite, on donne $\Phi(2) \simeq 0{,}977$).

3. Étude des densités.
   a) Quelle est la loi de $Y_n = \sum_{i=1}^{n} X_i$? En donner une densité g_n.
   b) En déduire une densité h_n de $\overline{X_n}$.
   c) En déduire une densité f_n de X_n^*.
4. On admet la formule de Stirling : $n! \underset{(+\infty)}{\sim} n^n \mathrm{e}^{-n}\sqrt{2\pi n}$.

Déterminer, pour $x \in \mathbb{R}$, la limite de $f_n(x)$ lorsque n tend vers $+\infty$.

Solution :

1. a) D'après le cours $E(\overline{X_n}) = 1$ et $V(\overline{X_n}) = \frac{1}{n}$.
   b) Par la loi faible des grands nombres, $(\overline{X_n})_n$ converge en probabilité vers 1.

2. a) D'après le théorème de la limite centrée, pour tout x réel, la suite $(F_n(x))_n$ converge vers $\Phi(x)$, où Φ est la fonction de répartition de la loi normale centrée réduite.
   b) Il vient :

$$P\left(1 - \frac{2}{\sqrt{n}} \leqslant \overline{X_n} \leqslant 1 + \frac{2}{\sqrt{n}}\right) = P(-2 \leqslant X_n^\star \leqslant 2) = 2\Phi(2) - 1 \approx 0.954$$

3. a) Les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes et suivent la même loi $\Gamma(1,1)$. D'après le cours, la variable aléatoire Y_n suit la loi $\Gamma(1,n)$, de densité g_n définie sur $\mathbb{R}$ par :

$$g_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{x^{n-1}\mathrm{e}^{-x}}{(n-1)!} & \text{si } x > 0 \end{cases}$$

b) Pour tout x réel, $P(\overline{X_n} \leqslant x) = P(Y_n \leqslant nx)$; donc si h_n désigne une densité de $\overline{X_n}$, il vient $h_n(x) = ng_n(nx)$, soit :

$$h_n(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{n^n x^{n-1}\mathrm{e}^{-nx}}{(n-1)!} & \text{si } x > 0 \end{cases}$$

c) Pour tout x réel, on a $F_n(x) = P(\sqrt{n}(X_n - 1) \leqslant x) = P\big(\overline{X_n} \leqslant 1 + \dfrac{x}{\sqrt{n}}\big)$.

Donc :

$$f_n(x) = \begin{cases} 0 & \text{si } x \leqslant -\sqrt{n} \\ \dfrac{n^n(1+\frac{x}{\sqrt{n}})^{n-1}\mathrm{e}^{-n(1+\frac{x}{\sqrt{n}})}}{\sqrt{n}(n-1)!} & \text{si } x > -\sqrt{n} \end{cases}$$

4. Soit x réel fixé. Il existe $n_0 \in \mathbb{N}^*$ tel que pour tout $n \geqslant n_0$, on a $x > -\sqrt{n}$. Et alors :

$$f_n(x) = \frac{n^n\sqrt{n}\mathrm{e}^{-n}}{n!}\big(1 + \frac{x}{\sqrt{n}}\big)\mathrm{e}^{n\ln(1+\frac{x}{\sqrt{n}}) - \sqrt{n}x}$$

La formule de Stirling admise donne

$$\lim_{n\to+\infty} \frac{n^n\sqrt{n}e^{-n}}{n!} = \frac{1}{\sqrt{2\pi}}$$

et un développement limité au voisinage de 0 donne

$$\mathrm{e}^{n\ln(1+\frac{x}{\sqrt{n}}) - \sqrt{n}x} \sim \mathrm{e}^{-\frac{x^2}{2} + o(1)}$$

Finalement

$$\lim_{n\to+\infty} f_n(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$$

Exercice 3.31.

Soit X une variable aléatoire de densité $f(x) = \begin{cases} \frac{2}{\theta}\big(1 - \frac{x}{\theta}\big) & \text{si } 0 \leqslant x \leqslant \theta \\ 0 & \text{sinon} \end{cases}$, où θ est un paramètre inconnu strictement positif.

On cherche à estimer θ à partir d'un échantillon $(X_1, X_2, \ldots, X_n)$ de variables aléatoires indépendantes de même loi que X.

On pose $\overline{X_n} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ et $M_n = \max(X_1, X_2, \ldots, X_n)$.

1. Déterminer la fonction de répartition F de X, $E(X)$ et $V(X)$.

2. a) Déterminer la constante a pour que $T_n = a\overline{X_n}$ soit un estimateur sans biais du paramètre θ.

b) Calculer le risque quadratique de T_n.

3. a) Déterminer la loi de M_n.

On note $I_n = \displaystyle\int_0^1 (1 - u^2)^n\, du$ et on admet que $I_n \underset{(+\infty)}{\sim} \sqrt{\dfrac{\pi}{4n}}$.

b) Montrer que : $E(M_n) = \theta(1 - I_n)$. Que peut-on en déduire pour M_n en tant qu'estimateur de θ ?

c) Montrer (ou admettre) que : $V(M_n) = \theta^2\big(\dfrac{1}{n+1} - I_n^2\big)$.

d) Calculer le risque quadratique de M_n. Entre T_n et M_n , quel estimateur choisiriez-vous, et pourquoi ?

4. a) Donner un estimateur sans biais de θ de la forme $M'_n = a_n M_n$, a_n étant un réel dépendant de n.

b) Entre M'_n et T_n , quel estimateur choisiriez-vous, et pourquoi ?

Solution :

1. Un calcul simple donne :

$$F(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \dfrac{2}{\theta}\big(x - \dfrac{x^2}{2\theta}\big) & \text{si } x \in [0,\theta] \\ 1 & \text{si } x \geqslant \theta \end{cases}$$

et on obtient également sans problèmes :

$$E(X) = \frac{\theta}{3}, \qquad V(X) = \frac{\theta^2}{18}$$

2. a) Comme $E(\overline{X_n}) = \dfrac{\theta}{3}$, $T_n = 3\overline{X_n}$ est un estimateur sans biais de θ.

b) Son risque quadratique est alors :

$$r_{T_n}(\theta) = V(T_n) = 9V(\overline{X_n}) = \frac{\theta^2}{2n}$$

3. a) Les variables $X_1, \ldots, X_n$ étant indépendantes et de même loi, on obtient : $F_{M_n}(x) = F^n(x)$ et $f_{M_n}(x) = nf(x)F^{n-1}(x)$, soit :

$$f_{M_n}(x) = \begin{cases} n\big(\dfrac{2}{\theta}\big)^n\big(x - \dfrac{x^2}{2\theta}\big)^{n-1}\big(1 - \dfrac{x}{\theta}\big) & \text{si } x \in [0,\theta] \\ 0 & \text{sinon} \end{cases}$$

b) L'espérance de M_n est alors, à l'aide du changement de variable $u = 1 - \dfrac{x}{\theta}$:

$$\begin{aligned} E(M_n) &= \int_0^\theta 2n\big(\frac{x}{\theta}\big)^n\big(2 - \frac{x}{\theta}\big)^{n-1}\big(1 - \frac{x}{\theta}\big)\,dx \\ &= \theta\int_0^1 2nu(1-u^2)^{n-1}(1-u)\,du \end{aligned}$$

Une intégration par parties donne ensuite

$$E(M_n) = \theta\Big(\big[-(1-u^2)^n(1-u)\big]_0^1 - \int_0^1 (1-u^2)^n\,du\Big) = \theta(1 - I_n)$$

On sait que $\lim\limits_{n\to+\infty} I_n = 0$; donc M_n est un estimateur asymptotiquement sans biais de θ.

c) Une méthode identique à la précédente (même changement de variable) conduit à :

$$E(M_n^2) = \theta^2\big(1 + \frac{1}{n+1} - 2I_n\big) \text{ et } V(M_n) = \theta^2\big(\frac{1}{n+1} - I_n^2\big)$$

d) On a alors $r_{M_n(\theta)} = \dfrac{\theta^2}{n+1}$.

Ainsi $r_{M_n(\theta)} > r_{T_n(\theta)}$; donc T_n est un estimateur « préférable » à M_n. En plus il est sans biais.

4. a) On choisit $M'_n = \dfrac{1}{1-I_n} M_n$.

b) On a :

$$V(M'_n) = \frac{1}{(1-I_n)^2} V(M_n)$$

et

$$\frac{V(M'_n)}{V(T_n)} = \frac{1}{(1-I_n)^2}\Big(\frac{2n}{n+1} - 2nI_n^2\Big)$$

ainsi :

$$\lim_{n\to+\infty} \frac{V(M'_n)}{V(T_n)} = 2 - \frac{\pi}{2} < 1$$

Ainsi, pour n « grand », M'_n est un « meilleur » estimateur que M_n.

Exercice 3.32.

Soit α un réel strictement positif.
On considère la fonction f définie par :

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

1. Montrer que f est une densité de probabilité.

2. Soit X une variable aléatoire réelle de densité f. Étudier l'existence des moments (non centrés) de X. Les calculer lorsqu'ils existent.

3. On considère une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires réelles mutuellement indépendantes de même loi, de densité commune f.
Pour tout $n \in \mathbb{N}^*$, et pour tout $i \in [\![1, n]\!]$, on pose :

$$Y_i = \ln(X_i), \quad Z_n = \frac{1}{X_1 X_2 \cdots X_n}$$

a) Déterminer une densité de Y_i, $i \in [\![1, n]\!]$.

b) En déduire une densité de Z_n.

4. Pour tout $n \geqslant 1$, on pose $U_n = (Z_n)^{1/n}$.

a) Montrer que la suite $(\ln(U_n))_{n\geqslant 1}$ converge en probabilité vers une variable aléatoire constante C que l'on déterminera.

b) Montrer que la suite $(U_n)_{n\geqslant 1}$ converge en loi vers e^C.

Solution :

1. Il suffit de vérifier que f est positive, continue sur $\mathbb{R}$ privé de 1, et de remarquer que

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_1^{+\infty} \alpha x^{-\alpha-1}\,dx = 1$$

2. Soit $k \in \mathbb{N}$, le moment $E(X^k)$ existe si et seulement si la fonction $x \mapsto x^{k-\alpha-1}$ admet une intégrale convergente sur $[1, +\infty[$.
Cette fonction étant continue sur $[1, +\infty[$, les intégrales de Riemann montrent que $E(X^k)$ existe si et seulement si $\alpha + 1 - k > 1$ soit $0 \leqslant k \leqslant \lfloor \alpha - 1 \rfloor$.
Dans ce cas, un calcul élémentaire donne $E(X^k) = \dfrac{\alpha}{\alpha - k}$.

3. a) On a immédiatement $Y_i(\Omega) =]0, +\infty[$ et pour $y > 0$:

$$P(Y_i \leqslant y) = P(\ln X_i \leqslant y) = P(X_i \leqslant \mathrm{e}^y) = F_{X_i}(\mathrm{e}^y) = 1 - \mathrm{e}^{-\alpha y}$$

Une densité f_i de Y_i est ainsi :

$$f_i(y) = \begin{cases} 0 & \text{si } y < 0 \\ \alpha\,\mathrm{e}^{-\alpha y} & \text{si } y \geqslant 0 \end{cases}$$

Ainsi Y_i suit la loi exponentielle de paramètre α.

b) De même, $Z_n(\Omega) =]0,1]$, et pour tout $z \in]0,1]$:

$$P(Z_n \leqslant z) = P\big(-\sum_{i=1}^{n} \ln X_i \leqslant \ln z\big) = P\big(\sum_{i=1}^{n} \ln X_i \geqslant -\ln z\big)$$

Par stabilité des lois Γ pour la somme, on sait que $\sum\limits_{i=1}^{n} \ln X_i$ suit la loi $\Gamma\big(\frac{1}{\alpha}, n\big)$, d'où pour tout $z \in]0,1]$:

$$F_{Z_n}(z) = \int_{-\ln z}^{1} \frac{\mathrm{e}^{-\alpha t}(\alpha t)^{n-1}}{(n-1)!}\,dt$$

Une densité de Z_n est alors :

$$f_{Z_n}(z) = \begin{cases} 0 & \text{si } z > \text{ ou } z > 1 \\ \dfrac{\alpha^n}{(n-1)!}(-\ln z)^{n-1} z^{\alpha-1} & \text{si } z \in]0,1] \end{cases}$$

4. a) Pour tout $n \geqslant 1$

$$\ln(U_n) = \frac{1}{n}\ln(Z_n) = -\frac{1}{n}\sum_{i=1}^{n} Y_i$$

Par la question précédente, les variables aléatoires (Y_i) sont indépendantes de même loi exponentielle $\mathcal{E}(\alpha)$.

Par la loi faible des grands nombres, la suite $(-\ln(U_n))$ converge en probabilité vers la variable constante égale à $\dfrac{1}{\alpha}$ et la suite $(\ln(U_n))$ converge en probabilité vers la variable constante égale) $-\dfrac{1}{\alpha}$

b) Considérons la variable aléatoire constante égale à $C = \mathrm{e}^{-1/\alpha}$ caractérisée par sa fonction de répartition :

$$F_C(x) = \begin{cases} 0 & \text{si } x < \mathrm{e}^{-1/\alpha} \\ 1 & \text{si } x \geqslant \mathrm{e}^{-1/\alpha} \end{cases}$$

Or, on sait que pour tout $\varepsilon > 0$:

$$\lim_{n\to+\infty} P\big(-\frac{1}{\alpha} - \varepsilon \leqslant \ln(U_n) \leqslant -\frac{1}{\alpha} + \varepsilon\big) = 1$$

ou

$$\lim_{n\to+\infty} P\big(\mathrm{e}^{-\frac{1}{\alpha}}\mathrm{e}^{-\varepsilon} \leqslant U_n \leqslant \mathrm{e}^{-\frac{1}{\alpha}}\mathrm{e}^{\varepsilon}\big) = 1$$

soit, pour tout $\varepsilon > 0$:

$$\lim_{n\to+\infty} [F_{U_n}\big(\mathrm{e}^{-\frac{1}{\alpha}}\mathrm{e}^{\varepsilon}\big) - F_{U_n}\big(e^{-\frac{1}{\alpha}}\mathrm{e}^{-\varepsilon}\big)] = 1$$

Comme une fonction de répartition prend ses valeurs entre 0 et que 1, cela entraîne que $\lim\limits_{n\to+\infty} F_{U_n}\big(\mathrm{e}^{-\frac{1}{\alpha}}\mathrm{e}^{-\varepsilon}\big) = 0$ et que $\lim\limits_{n\to+\infty} F_{U_n}\big(\mathrm{e}^{-\frac{1}{\alpha}}\mathrm{e}^{\varepsilon}\big) = 1$.

ce qui signifie que pour $t > \mathrm{e}^{-\frac{1}{\alpha}}$, on a : $\lim\limits_{n\to\infty} F_{U_n}(t) = 1$, tandis que pour $t < \mathrm{e}^{-\frac{1}{\alpha}}$ cette limite est nulle. Par conséquent :

$$(U_n) \text{ tend en loi vers } C.$$

Exercice 3.33.

Soit r un entier non nul. On dit qu'une variable aléatoire X suit la loi du χ^2 à r degrés de liberté si X suit la loi Γ de paramètres 2 et $\frac{r}{2}$.

1. Déterminer l'espérance et la variance d'une variable X suivant la loi du χ^2 à r degrés de liberté.

2. a) Montrer que pour tout $\lambda > 0$, pour tout entier $n \geqslant 1$:

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

b) Soit Y_1 une variable aléatoire suivant une loi de Poisson de paramètre λ et X_{2n} une variable aléatoire suivant la loi du χ^2 à $2n$ degrés de liberté.

Montrer que $P(X_{2n} > 2\lambda) = P(Y_1 < n)$.

c) Écrire une fonction Pascal qui permet de simuler la fonction de répartition de la variable aléatoire X_{2n}.

3. Soit k un entier non nul et $X_1, \dots, X_k$ des variables aléatoires indépendantes suivantes toutes la loi gaussienne centrée réduite.

a) Déterminer la loi de X_1^2.

b) En déduire la loi de $X_1^2 + \cdots + X_k^2$.

Solution :

1. D'après le cours, on sait que si X suit une loi $\Gamma(2, r/2)$, son espérance vaut r et sa variance $2r$.

2. a) La fonction exponentielle étant de classe C^{∞}, la formule de Taylor avec reste intégral à l'ordre n donne :

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

Le changement de variable affine $t \mapsto \lambda - t$ donne la formule désirée.

b) On sait que :

$$\begin{aligned} P(X_{2n} > 2\lambda) &= \frac{1}{\Gamma(n) 2^n} \int_{2\lambda}^{+\infty} x^{n-1} \mathrm{e}^{-x/2} dx = \frac{1}{\Gamma(n) 2^n} \int_{\lambda}^{+\infty} 2^{n-1} t^{n-1} \mathrm{e}^{-t} 2 dt \\ &= \int_{\lambda}^{+\infty} \frac{t^{n-1}}{(n-1)!} \, \mathrm{e}^{-t} dt \end{aligned}$$

et :

$$\begin{aligned} P(Y_1 < n) &= \sum_{k=0}^{n-1} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda} \big[\mathrm{e}^{\lambda} - \mathrm{e}^{\lambda} \int_0^{\lambda} \frac{t^{n-1}}{(n-1)!} \, \mathrm{e}^{-t} \, dt \big] \\ &= 1 - \int_0^{\lambda} \frac{t^{n-1}}{(n-1)!} \, \mathrm{e}^{-t} \, dt \\ &= \int_0^{+\infty} \frac{t^{n-1}}{(n-1)!} \, \mathrm{e}^{-t} \, dt - \int_0^{\lambda} \frac{t^{n-1}}{(n-1)!} \, \mathrm{e}^{-t} \, dt \\ &= \int_{\lambda}^{+\infty} \frac{t^{n-1}}{(n-1)!} \, \mathrm{e}^{-t} \, dt = P(X_{2n} > 2\lambda) \end{aligned}$$

c) La question précédente montre que $P(X_{2n} > x) = P(Y_{x/2} < n)$. Une fonction Pascal possible est donc

```
Function P(n : integer ; x : real) : real ;
Var lam, pois, prob : real ;
   k : integer ;
Begin
lam := x/2 ;
pois : = exp(-lam) ;
for k := 1 to n-1 do
begin
      pois := pois*lam/k ;
      prob := prob + pois
end ;
P := prob
end ;
```

3. a) Soit $x \geqslant 0$. Alors :

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

Une densité de X^2 est donc :

$$f_{X^2}(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{1}{\Gamma(1/2)\sqrt{2}} x^{1/2-1} \mathrm{e}^{-x/2} & \text{si } x \geqslant 0 \end{cases}$$

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

b) Les variables aléatoires étant indépendantes, la somme $\displaystyle\sum_{k=1}^{k} X_i^2$ suit la loi $\Gamma(2, k/2)$ (loi du χ^2 à k degrés de liberté).

Exercice 3.34.

Soit k un réel strictement positif,

1. Déterminer le réel α pour que la fonction f_k de $\mathbb{R}$ dans $\mathbb{R}$ définie par :

$$f_k(x) = \max(0, \alpha - k|x|)$$

soit une densité de probabilité.

2. Soit X une variable aléatoire réelle telle que :

- $E(X) = m$;
- $X - E(X)$ admet la fonction f_k pour densité de probabilité.

a) Calculer les moments centrés $E[(X - m)^n]$, pour $n \in \mathbb{N}$.

b) En déduire la variance $V(X)$ de la variable aléatoire X.

3. On dispose d'un n-échantillon $(X_1, X_2, \ldots, X_n)$ de la loi de X à partir duquel on souhaite estimer le paramètre k.

Montrer que $T_n = \dfrac{1}{n} \sum\limits_{i=1}^{n} (X_i - m)^2$ est un estimateur de $V(X)$ sans biais et convergent.

4. On suppose en outre $E(X) = m = 0$. Soit Y une variable aléatoire de loi uniforme sur $[0, 2]$ indépendante de X.
Dans quel intervalle est-on sûr de trouver $X^2 + Y$ avec un risque inférieur à 5% ?

Solution :

1. La fonction f_k est paire non nulle sur $\left]-\frac{\alpha}{k}, \frac{\alpha}{k}\right[$, nulle ailleurs, et est continue sur $\mathbb{R}$. Ainsi :

- $f_k(x) \geqslant 0$ entraîne que $\alpha \geqslant 0$.
- $\displaystyle\int_{-\infty}^{+\infty} f_k(t)\,dt = 1$ entraîne que $\alpha^2 = k$.

L'unique solution est donc $\alpha = \sqrt{k}$.

2. a) Une densité de $X - m$ étant une fonction paire, les moments d'ordre impair sont nuls et pour $n = 2p$:

$$E((X-m)^n) = 2\int_0^{1/\sqrt{k}} t^n(\sqrt{k} - kt)\,dt = \frac{2}{(n+1)(n+2)}k^{-p}$$

b) Pour $n = 2$, il vient $V(X) = E((X - E(X))^2) = \dfrac{1}{6k}$.

3. La linéarité de l'espérance donne facilement $E(T_n) = V(X)$, ce qui montre que T_n est un estimateur sans biais de $V(X)$.

La loi faible des grands nombres appliquée à la suite de variables aléatoires identiquement distribuées $((X_i - m)^2)$ montre ensuite que la suite (T_n) converge en probabilité vers $V(X)$, ce qui permet d'écrire que la suite d'estimateurs (T_n) est convergente.

4. Ici $m = 0$. On applique l'inégalité de Bienaymé-Cebisheff à la variable aléatoire $Z = X^2 + Y$. Ainsi, pour tout $\varepsilon > 0$

$$P(|Z - E(Z)| < \varepsilon) \geqslant 1 - \frac{V(Z)}{\varepsilon^2}$$

Or $E(Z) = \dfrac{7}{6}$ et $V(Z) = \dfrac{67}{180}$.

On veut que $1 - \dfrac{V(Z)}{\varepsilon^2} \geqslant 0.95$, soit $\varepsilon^2 \geqslant \dfrac{V(Z)}{0.05}$ et donc $\varepsilon \geqslant \dfrac{\sqrt{67}}{3}$.

L'intervalle cherché est donc $\left]\dfrac{7-2\sqrt{67}}{6}, \dfrac{7+2\sqrt{67}}{6}\right[$.