PC de Probabilités - Cours MA 105
Formation SUPAERO

Institut Supérieur de l'Aéronautique et de l'Espace (ISAE)

Jérémie Bigot

17 mars 2014

Remerciements

L'ensemble de ces exercices et leurs corrigés sont inspirés du document "Exercices corrigés de probabilités" par Manuel Samuelides, SUPAERO, 1998, ainsi que des exercices de probabilités préparés par Claudie Chabriac qui sont disponibles à l'adresse :

`https://sites.google.com/site/cchabriac/`

Table des matières

1 Variables aléatoires discrètes - PC1

__Énoncés des exercices__

__**Exercice 1.1**__ : *Loi géométrique*
On considère l'expérience qui consiste à jeter en l'air une pièce de monnaie et à s'arrêter dès que l'on obtient pile pour la première fois. On note, les événements élémentaires : $F =$ "Face" $=$, $\Pi =$ "Pile", et leur probabilité : $\mathbb{P}(F) = p$, $\mathbb{P}(\Pi) = q$ avec $p \in]0, 1[$ et $q = 1 - p$. On note N la variable aléatoire $N =$" nombre de lancers effectués dans une partie ".

1. Déterminer la loi de probabilité de N. La loi de N est appelée 'loi géométrique'.
2. Calculer l'espérance et la variance de N i.e. $\mathbb{E}(N)$ et $\mathrm{Var}(N)$.

__**Exercice 1.2**__ : *Loi de Bernoulli*
Un joueur joue à pile ou face avec une pièce. On note les évènements $F =$ "Face" $=$, $\Pi =$ "Pile", et leur probabilité : $\mathbb{P}(F) = p$, $\mathbb{P}(\Pi) = q$ avec $p \in]0, 1[$ et $q = 1 - p$. Soit X la variable aléatoire qui prend la valeur 0 quand on obtient pile et la valeur 1 lorsqu'on obtient face. Calculer la loi de X, son espérance et sa variance. La loi de X est appelée "loi de Bernoulli".

__**Exercice 1.3**__ : *Loi binomiale*
Un joueur joue à pile ou face avec une pièce. Il lance la pièce en l'air n fois. D'après l'exercice précédent, chaque lancer i peut être modélisé par une variable aléatoire X_i qui suit une loi de Bernoulli. Soit la variable aléatoire $S_n = \sum_{i=1}^{n} X_i$.

1. Calculer $\mathbb{P}(S_n = k) = \mathbb{P}($"$k$ fois face sur n lancers"$)$. La loi de S_n est appelée "loi binomiale".
2. Calculer l'espérance et la variance de S_n.
3. <u>Application 1</u> : Un homme attendant un rendez-vous décide de marcher en jouant à pile ou face. Si il obtient pile, il fait 10 pas vers la droite, et dans le cas contraire (face), 10 pas vers la gauche. On suppose $p = q = 1/2$. Quelle est la probabilité qu'après avoir fait 100 pas :
   – il soit revenu à son point de départ.
   – il se retrouve à 20 pas de son point de départ.
4. <u>Application 2</u> : La proportion de pièces défectueuses dans une fabrication est de 1/1000. Quelle est la probabilité pour que l'on ait effectivement exactement une pièce défectueuse dans un lot de 1000 pièces ?
5. <u>Application 3</u> : Statistiquement 5% des réservations aériennes ne sont pas utilisées. Une compagnie décide donc de faire du 'surbooking' et vend 100 billets pour 97 places. Quelle est la probabilité pour que tous les passagers qui se présentent aient une place ?

__**Exercice 1.4**__ : *Fonction génératrice*
Une urne est composée en proportion de p boules blanches et de q boules noires. On effectue un tirage avec remise pour constituer un échantillon de n boules, et on note X la variable aléatoire égale au nombre de boules blanches dans l'échantillon de taille n.

1. Calculer la probabilité d'avoir exactement k boules blanches i.e. $\mathbb{P}(X = k)$ dans un échantillon. Quelle est la loi suivie par X ?

2. Soit $z \in \mathbb{R}$. Déterminer $G_X(z) = \mathbb{E}(z^X)$ la fonction génératrice de X. Quand le rayon de convergence de la série $G_X(z)$ est supérieur à 1, vérifier les relations suivantes : $G_X(1) = 1$, $G'_X(1) = \mathbb{E}(X)$ et $G''_X(1) = \mathbb{E}(X(X-1))$.
3. A l'aide la fonction génératrice, déterminer l'espérance et la variance de X.
4. En considérant $\mathbb{E}\left((-1)^X\right)$, déterminer la probabilité pour que X soit impair.

<u>Exercice 1.5</u> : *Loi de Poisson*
On considère une variable aléatoire discrète X, dont la loi de probabilité est définie par :

$$\forall n \in \mathbb{N},\ \mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda},$$

où $\lambda > 0$ est un paramètre positif.

1. Vérifier qu'il s'agit bien d'une loi de probabilité.
2. Calculer l'espérance et la variance de X.
3. Montrer que X a plus de chance d'être paire qu'impaire.
4. Calculer $\mathbb{E}\left(\dfrac{1}{1+X}\right)$

<u>Exercice 1.6</u> : *Loi uniforme et fonction de répartition d'une variable aléatoire discrète*
Une roue de loterie est divisée en 10 secteurs semblables numérotés de 1 à 10. On fait tourner la roue 2 fois. On note X le plus grand numéro obtenu, Y le plus petit et X_i (i=1 ou 2) le i-ème numéro obtenu.

1. Déterminer pour $k = 1, \dots, 10$, $\mathbb{P}(X_i = k)$ puis $\mathbb{E}(X_i)$. En déduire que X_i suit une loi uniforme sur l'ensemble $\{1, 2, \dots, 10\}$.
2. Déterminer la fonction de répartition de la variable aléatoire X, $F_X(k) = \mathbb{P}(X \leqslant k)$ pour $k = 1, \dots, 10$. En déduire que $\mathbb{P}(X = k) = \dfrac{2k-1}{100}$ et déterminer $\mathbb{E}(X)$.
3. Déterminer $\mathbb{E}(Y)$.

Corrigés des exercices

Exercice 1.1 : *Loi géométrique*

1. Si on appelle ω_n l'événement correspondant à un tirage de $(n-1)$ fois Face suivies de Pile (avec $n \in \mathbb{N}^*$) :

$$\omega_n = \{\underbrace{F, F, \ldots\ldots, F}_{(n-1)\text{ fois}}, \Pi\}.$$

On a $\mathbb{P}(N = n) = \mathbb{P}(\omega_n) = p^{n-1}q = q(1-q)^{n-1}$ car les n lancers sont indépendants. On vérifie bien que

$$\sum_{n=1}^{+\infty} \mathbb{P}(N = n) = \sum_{n=1}^{+\infty} q(1-q)^{n-1} = q \sum_{n=1}^{+\infty} p^{n-1} = q\frac{1}{1-p} = 1.$$

On a donc bien défini une loi de probabilités. On dit que N suit une loi géométrique de paramètre q.

2. Par définition (avec $P(N = 0) = 0$) :

$$\mathbb{E}(N) = \sum_{n=0}^{+\infty} n\mathbb{P}(N = n) = q \sum_{n=1}^{+\infty} np^{n-1},$$

qui est donc la dérivée de la série $\displaystyle\sum_{n=0}^{+\infty} p^n = \frac{1}{1-p}$. Ainsi, comme $q = 1-p$, on obtient que

$$\mathbb{E}(N) = q\frac{1}{(1-p)^2} = \frac{1}{q}.$$

Notons maintenant que

$$\mathrm{Var}(N) = \mathbb{E}(N^2) - (\mathbb{E}(N))^2 = \mathbb{E}(N(N-1)) + \mathbb{E}(N) - (\mathbb{E}(N))^2 = \mathbb{E}(N(N-1)) + \frac{1}{q} - \frac{1}{q^2}.$$

Il suffit maintenant de remarquer que

$$\mathbb{E}(N(N-1)) = \sum_{n=0}^{+\infty} n(n-1)\mathbb{P}(N = n) = q \sum_{n=0}^{+\infty} n(n-1)p^{n-1} = qp \sum_{n=1}^{+\infty} n(n-1)p^{n-2},$$

qui est donc la dérivée seconde de la série $\displaystyle\sum_{n=0}^{+\infty} p^n = \frac{1}{1-p}$. D'où

$$\mathbb{E}(N(N-1)) = qp\frac{2}{(1-p)^3} = \frac{2p}{q^2} = \frac{2(1-q)}{q^2}.$$

Finalement, on obtient que

$$\mathrm{Var}(N) = \frac{2(1-q)}{q^2} - \frac{1-q}{q^2} = \frac{1-q}{q^2}.$$

Exercice 1.2 : *Loi de Bernoulli*

X est une variable aléatoire discrète qui peut prendre les valeurs 0 ou 1. Par définition, sa loi est telle que $\mathbb{P}(X=0) = \mathbb{P}(\Pi) = q = 1-p$ et $\mathbb{P}(X=1) = \mathbb{P}(F) = p$. On dit que X suit une loi de Bernoulli de paramètre p. On obtient alors que

$$\mathbb{E}(X) = 0\mathbb{P}(X=0) + 1\mathbb{P}(X=1) = p,$$

et

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \mathbb{E}(X^2) - p^2.$$

Or $\mathbb{E}(X^2) = 0^2\mathbb{P}(X=0) + 1^2\mathbb{P}(X=1) = p$ et donc $\mathrm{Var}(X) = p - p^2 = p(1-p)$.

Exercice 1.3 : *Loi binomiale*

1. Le nombre de combinaisons possibles permettant d'obtenir k fois F (et donc $(n-k)$ fois Π) sur n lancers est C_n^k. En effet, cela revient à chercher toutes les combinaisons permettant de positionner k événements F sur n positions possibles, c'est-à-dire toutes les combinaisons de k éléments dans un ensemble de n éléments. La probabilité de chacune de ses combinaisons est $p^k(1-p)^{n-k}$ quelque soit l'ordre envisagé (indépendance des lancers). On a donc que

$$\mathbb{P}(S_n = k) = C_n^k p^k (1-p)^{n-k}.$$

On vérifie facilement (à l'aide de la formule du binôme) que

$$\sum_{k=0}^{n} \mathbb{P}(S_n = k) = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p - (1-p))^n = 1.$$

On a donc bien défini une loi de probabilité appelée loi binomiale de paramètres n et p et notée $\mathcal{B}(n,p)$.

2. Par définition,

$$\begin{aligned}
\mathbb{E}(S_n) &= \sum_{k=0}^{n} k\mathbb{P}(S_n = k) \\
&= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} = np \sum_{k=0}^{n-1} C_{n-1}^{k} p^k (1-p)^{(n-1)-k} \\
&= np(p + (1-p))^{n-1} = np.
\end{aligned}$$

On peut également procéder de façon plus simple pour la calcul de l'espérance. En utilisant le fait que $S_n = \sum_{i=1}^{n} X_i$, que $\mathbb{E}(X_i) = p$ et la linéarité de l'espérance, on obtient que

$$\mathbb{E}(S_n) = \sum_{i=1}^{n} \mathbb{E}(X_i) = np.$$

Procédons maintenant au calcul de la variance. On a que

$$\begin{aligned}\mathrm{Var}(S_n) &= \mathbb{E}(S_n^2) - (\mathbb{E}(S_n))^2 = \mathbb{E}(S_n(S_n-1)) + \mathbb{E}(S_n) - (\mathbb{E}(S_n))^2 \\ &= \mathbb{E}(S_n(S_n-1)) + np(1-np)\end{aligned}$$

Il suffit ensuite de remarquer que

$$\begin{aligned}\mathbb{E}(S_n(S_n-1)) &= \sum_{k=0}^{n} k(k-1)\mathbb{P}(S_n=k) \\ &= \sum_{k=0}^{n} k(k-1)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= n(n-1)p^2\sum_{k=2}^{n} C_{n-2}^{k-2}p^{k-2}(1-p)^{(n-2)-(k-2)} = n(n-1)p^2\sum_{k=0}^{n-2} C_{n-2}^{k}p^{k}(1-p)^{(n-2)-k} \\ &= n(n-1)p^2(p+(1-p))^{n-2} = n(n-1)p^2.\end{aligned}$$

Ainsi, on obtient que

$$\mathrm{Var}(S_n) = n(n-1)p^2 + np(1-np) = np((n-1)p+1-np) = np(1-p).$$

On peut également procéder de façon plus simple pour le calcul de la variance. Etant donné que les lancers sont indépendants, on a que que $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes telles que $\mathrm{Var}\,(X_i) = p(1-p)$. Ainsi,

$$\mathrm{Var}(S_n) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\,(X_i) = np(1-p).$$

3. Pour se retrouver à son point de départ il a obtenu 5 fois Pile et 5 fois Face. La probabilité recherchée est donc :

$$C_{10}^5 p^5(1-p)^5 \approx 0,246 \text{ si } p = 1/2.$$

La probabilité qu'il se retrouve à 20 pas (à gauche ou à droite) de son point de départ est :

$$C_{10}^4 p^6(1-p)^4 + C_{10}^4 p^4(1-p)^6 \approx 0,41 \text{ si } p = 1/2.$$

4. On peut modéliser le nombre de pièces défectueuses à l'aide d'une loi binomiale de paramètres $n = 1000$ et $p = 1/1000$ comme somme de 1000 lois de Bernoulli, chaque pièce ayant 2 états possibles (correcte ou défectueuse). D'où la probabilité d'avoir exactement une pièce défectueuse sur un lot de 1000 pièces :

$$C_{1000}^1 p^1(1-p)^{999} = \left(\frac{999}{1000}\right)^{999} \approx 0,368.$$

5. La probabilité qu'un passager se présente à l'embarquement est $p = 0,95$. Pour que la compagnie soit en défaut, il faut que plus de 97 passagers sur les 100 réservés se présentent. Le problème peut se modéliser par une loi binomiale, S_{100}, de paramètres $n = 100$ et $p = 0,95$. La probabilité P pour que la compagnie soit en défaut peut alors s'écrire :

$$P = \mathbb{P}(S_{100} > 97) = \mathbb{P}(S_{100} = 98) + \mathbb{P}(S_{100} = 99) + \mathbb{P}(S_{100} = 100),$$

soit

$$P = C_{100}^{98} p^{98}(1-p)^2 + C_{100}^{99} p^{99}(1-p) + p^{100} \approx 0,118.$$

Exercice 1.4 : *Fonction génératrice*

Pour un tirage notons : $p_B = \dfrac{p}{p+q}$ et $p_N = \dfrac{q}{p+q}$, les probabilités respectives de tirer une boule blanche et une boule noire. On a $p_B + p_N = 1$.

1. Il est clair que X suit une loi binomiale $\mathcal{B}(n, p_B)$ d'espérance $\mathbb{E}(X) = np_B$ et de variance $\mathrm{Var}(X) = np_B(1 - p_B$.

2. Par définition de l'espérance, on obtient que

$$G_X(z) = \mathbb{E}(z^X) = \sum_{k=0}^{n} z^k \mathbb{P}(X = k) = \sum_{k=0}^{n} z^k C_n^k p_B^k p_N^{n-k} = (zp_B + p_N)^n = ((z-1)p_B + 1)^n.$$

On vérifie ensuite les relations (valables pour toute variable aléatoire discrète à valeurs dans $\{0, \ldots, n\}$)

$$G_X(1) = \sum_{k=0}^{n} \mathbb{P}(X = k) = 1,$$

$$G'_X(1) = \sum_{k=0}^{n} k\mathbb{P}(X = k) = \mathbb{E}(X),$$

et

$$G''_X(1) = \sum_{k=0}^{n} k(k-1)1\mathbb{P}(X = k) = \mathbb{E}(X(X-1)).$$

3. Dans le cas où X suit une loi binomiale $\mathcal{B}(n, p_B)$, on a que $G_X(z) = ((z-1)p_B + 1)^n$. Donc, on obtient que $G'_X(z) = np_B((z-1)p_B + 1)^{n-1}$ et $G''_X(z) = n(n-1)p_B^2((z-1)p_B + 1)^{n-2}$. Ainsi,

$$G'_X(1) = np_B = \mathbb{E}(X).$$

et

$$G''_X(1) = n(n-1)p_B^2.$$

De plus, en remarquant que

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2 \\ &= G''_X(1) + \mathbb{E}(X) - (\mathbb{E}(X))^2. \end{aligned}$$

on obtient que

$$\mathrm{Var}(X) = n(n-1)p_B^2 + np_B(1 - np_B) = np_B(1 - p_B).$$

4. On a que

$$\mathbb{E}\left((-1)^X\right) = G_X(-1) = (1 - 2p_B)^n.$$

Or,

$$\mathbb{P}(\{X \text{ est pair}\}) = \mathbb{P}(X = 0) + \mathbb{P}(X = 2) + \ldots = \frac{1}{2}\sum_{k=0}^{n}\Big(\mathbb{P}(X = k) + (-1)^k\mathbb{P}(X = k)\Big) = \frac{1}{2}\left(1 + G_X(-1)\right)$$

D'où,

$$\mathbb{P}(\{X \text{ est pair}\}) = \frac{1}{2}\left(1 + (1 - 2p_B)^n\right).$$

Exercice 1.5 : *Loi de Poisson*

1. On a que
$$\sum_{n=0}^{+\infty} \mathbb{P}(X=n) = \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1.$$

2. Par définition
$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} n\mathbb{P}(X=n) = e^{-\lambda} \sum_{n=1}^{+\infty} \frac{\lambda^n}{(n-1)!} = e^{-\lambda}\lambda \sum_{n=1}^{+\infty} \frac{\lambda^{n-1}}{(n-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$

et

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2 = \mathbb{E}(X(X-1)) + \lambda - \lambda^2.$$

Or
$$\mathbb{E}(X(X-1)) = \sum_{n=0}^{+\infty} n(n-1)\mathbb{P}(X=n) = e^{-\lambda}\lambda^2 \sum_{n=2}^{+\infty} \frac{\lambda^{n-2}}{(n-2)!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2,$$

et donc $\mathrm{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

3. Il suffit de remarquer tout d'abord que

$$\mathbb{P}(\{X \text{ est pair}\}) = \frac{1}{2} \sum_{n=0}^{+\infty} \left(\mathbb{P}(X=n) + (-1)^n \mathbb{P}(X=n)\right) = \frac{1}{2}\left(1 + G_X(-1)\right),$$

et

$$\mathbb{P}(\{X \text{ est impair}\}) = \frac{1}{2} \sum_{n=0}^{+\infty} \left(\mathbb{P}(X=n) - (-1)^n \mathbb{P}(X=n)\right) = \frac{1}{2}\left(1 - G_X(-1)\right),$$

où $G_X(z) = \mathbb{E}(z^X) \sum_{n=0}^{+\infty} z^n \mathbb{P}(X=n)$ est la fonction génératrice de X. On a alors que

$$G_X(z) = \sum_{n=0}^{+\infty} z^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{(\lambda z)^n}{n!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}.$$

Et donc

$$\mathbb{P}(\{X \text{ est pair}\}) = \frac{1}{2}\left(1 + e^{-2\lambda}\right) > \frac{1}{2}\left(1 - e^{-2\lambda}\right) = \mathbb{P}(\{X \text{ est impair}\}).$$

4. Par définition

$$\mathbb{E}\left(\frac{1}{1+X}\right) = \sum_{n=0}^{+\infty} \frac{1}{1+n} \mathbb{P}(X=n) = \frac{e^{-\lambda}}{\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^{n+1}}{(n+1)!} = \frac{e^{-\lambda}}{\lambda}\left(e^{\lambda} - 1\right) = \frac{1 - e^{-\lambda}}{\lambda}.$$

Exercice 1.6 : *Loi uniforme et fonction de répartition d'une variable aléatoire discrète*

1. Par définition, on a que pour tout $k \in \{1, \ldots, 10\}$

$$\mathbb{P}(X_1 = k) = \mathbb{P}(X_2 = k) = \frac{1}{10}, \text{ (définition d'une loi uniforme sur l'ensemble } \{1, 2, \ldots, 10\}),$$

et donc

$$\mathbb{E}(X_1) = \mathbb{E}(X_2) = \sum_{k=1}^{10} k \frac{1}{10} = \frac{10 \times 11}{10 \times 2} = \frac{11}{2}.$$

2. En utilisant l'indépendance de X_1 et X_2, on a que

$$F_X(k) = \mathbb{P}(X \leqslant k) = \mathbb{P}(\{X_1 \leqslant k\} \cap \{X_2 \leqslant k\}) = \mathbb{P}(X_1 \leqslant k)\mathbb{P}(X_2 \leqslant k) = \frac{k^2}{100},$$

car

$$\mathbb{P}(X_1 \leqslant k) = \mathbb{P}(X_2 \leqslant k) = \frac{k}{10}.$$

Ensuite, étant donné que X est une variable aléatoire discrète, on obtient que

$$\mathbb{P}(X = k) = \mathbb{P}(X \leqslant k) - \mathbb{P}(X \leqslant k - 1) = \frac{k^2 - (k-1)^2}{100} = \frac{2k - 1}{100},$$

et donc

$$\mathbb{E}(X) = \sum_{k=1}^{10} k\mathbb{P}(X = k) = \frac{2}{100} \sum_{k=1}^{10} k^2 - \frac{1}{100} \sum_{k=1}^{10} k = \frac{1}{100} \left(2\frac{10 \times 11 \times 12}{6} - \frac{10 \times 11}{2} \right) = \frac{429}{60} = 7,15.$$

3. Pour calculer $\mathbb{E}(Y)$ il suffit de remarquer que $X + Y = X_1 + X_2$. Donc, par linéarité de l'espérance, on obtient que

$$\mathbb{E}(Y) = \mathbb{E}(X_1) + \mathbb{E}(X_2) - \mathbb{E}(X) = 11 - \frac{429}{60} = 3,85.$$

2 Variables aléatoires continues - PC2

Énoncés des exercices

Exercice 2.1 : *Loi gaussienne ou loi Normale*

On dit qu'une variable aléatoire X suit une loi gaussienne centrée normée, notée $\mathcal{N}(0,1)$, si sa densité est

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},\ x \in \mathbb{R}.$$

1. Montrer que f_X est bien une densité de probabilité. Montrer que X et $-X$ ont même loi.
2. Calculer l'espérance et la variance de X.
3. On pose $Y = m + \sigma X$ avec $m \in \mathbb{R}$ et $\sigma > 0$. Déterminer la densité de Y, son espérance et sa variance. On dit que Y suit une loi gaussienne ou normale de paramètres m et σ, notée $N(m, \sigma)$.
4. Exemple : la longueur L d'une pièce produite en chaîne est une variable aléatoire gaussienne $\mathcal{N}(1000, 1)$. Déterminer la probabilité pour que $998 < L < 1002$ (on admet que $\mathbb{P}(-2 < X < 2) \approx 0,96$ si X suit la loi $\mathcal{N}(0,1)$.

Exercice 2.2 : *Loi uniforme*

U est une variable aléatoire sur $[0,1]$ telle que la probabilité que U appartienne à un intervalle de longueur a, est proportionnelle à a et indépendante de la position de cet intervalle dans $[0,1]$.

1. Déterminer la fonction de répartition de U, sa densité de probabilité, son espérance et sa variance.
2. Généralisation pour un intervalle $[a,b]$.

Exercice 2.3 : *Loi exponentielle*

Soit T une variable aléatoire positive de densité de probabilité :

$$f_T(t) = ke^{-\lambda t},\ t \in [0, +\infty[.$$

1. Déterminer k pour f_T soit effectivement une densité de probabilité.
2. Calculer l'espérance et la variance de T. Que peut-on en conclure ?
3. Déterminer $\alpha > 0$ tel que $\mathbb{P}(T < \alpha) = \mathbb{P}(T > \alpha)$.

Exercice 2.4 : *Slalom*

Un skieur passe un slalom de 20 portes. S'il ne tombe pas son temps de parcours suit une loi normale de moyenne $50s$ et d'écart-type $2s$. A chaque porte il peut tomber avec un probabilité de $1/20$, ce qui lui fait perdre $4s$. On admet que les chutes sont indépendantes les unes des autres. Quelle est la probabilité de décrocher la flèche de platine (temps de parcours inférieur à $52s$) ?

Exercice 2.5 : *Loi Gamma*

On dit qu'une variable aléatoire X positive suit une loi Gamma de paramètres $\alpha > 0$ et $\lambda > 0$, notée $G(\alpha, \lambda)$, si sa densité est

$$f_X(x) = \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, \ x \in [0, +\infty[, \text{ avec } \Gamma(\alpha) = \int_0^{+\infty} e^{-t} t^{\alpha-1} dt$$

On rappelle que $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ pour tout $\alpha > 0$.

1. Montrer que f_X est bien une densité de probabilité.
2. Calculer l'espérance et la variance de X.

Corrigés des exercices

Exercice 2.1 : *Loi gaussienne ou loi Normale*

1. On peut remarquer que (à l'aide d'un changement de variables en coordonnées polaires)

$$\left(\int_{-\infty}^{+\infty} f_X(x)dx\right)^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}}dxdy = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{+\infty} re^{-\frac{r^2}{2}}drd\theta = \frac{1}{2\pi}\times 2\pi = 1.$$

Pour montrer que X et $-X$ ont même loi, il suffit (par exemple) de montrer que leur densité de probabilité f_X et f_{-X} sont les mêmes. Pour cela, on peut dériver la fonction de répartition $F_{-X}(x) = \mathbb{P}(-X \leqslant x) = \mathbb{P}(X \geqslant -x) = 1 - F_X(-x)$ de $-X$, ce qui donne (pour $x \in \mathbb{R}$)

$$f_{-X}(x) = f_X(-x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} = f_X(x).$$

2. Par définition,

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf_X(x)dx = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}xe^{-\frac{x^2}{2}}dx.$$

La fonction $x \mapsto xe^{-\frac{x^2}{2}}$ étant impaire et intégrable sur $[0,+\infty[$ on obtient que $\mathbb{E}(X) = 0$. Ensuite, on a donc que

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X) = \mathbb{E}(X^2) = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}x^2e^{-\frac{x^2}{2}}dx.$$

Ainsi, à l'aide d'une intégration par parties, on obtient que

$$\mathrm{Var}(X) = \left[-\frac{1}{\sqrt{2\pi}}xe^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = 1.$$

3. Calculons tout d'abord la fonction de répartition de Y (avec $y \in \mathbb{R}$) :

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}\left(X \leqslant \frac{y-m}{\sigma}\right) = F_X\left(\frac{y-m}{\sigma}\right).$$

Donc la densité de Y est donné par

$$f_Y(y) = F_Y'(y) = \frac{1}{\sigma}f_X\left(\frac{y-m}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-m)^2}{2\sigma^2}}.$$

Par linéarité de l'espérance,

$$\mathbb{E}(Y) = m + \sigma\mathbb{E}(X) = m.$$

De plus,

$$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \mathbb{E}(m^2 + \sigma^2X^2 + 2m\sigma X) - m^2 = \sigma^2\mathbb{E}(X^2) = \sigma^2.$$

---

3. En utilisant le fait que $X = L - 1000$ suit une loi $\mathcal{N}(0,1)$, on obtient que

$$\mathbb{P}(998 < L < 1002) = \mathbb{P}(-2 < L - 1000 < 2) \approx 0,96.$$

Exercice 2.2 : *Loi uniforme*

1. D'après la définition de U on a que la fonction de répartition de U est telle que pour $t \in [0,1]$

$$F_U(t) = \mathbb{P}(U \leqslant t) = \mathbb{P}(U \in [0,t]) = Ct,$$

où $C > 0$ est une constante à déterminer. Pour $t < 0$, $F_U(t) = \mathbb{P}(U \leqslant t) = 0$ et pour $t > 1$, $F_U(t) = \mathbb{P}(U \leqslant t) = 1$. Ainsi, la densité de U est donné par

$$f_U(t) = C\mathbb{1}_{[0,1]}(t),$$

et la contrainte $\displaystyle\int_{-\infty}^{+\infty} f_U(t)dt = 1$ implique que $C = 1$. Par définition,

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} tf_U(t)dt = \int_0^1 tdt = \left[\frac{t^2}{2}\right]_0^1 = \frac{1}{2},$$

et

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \int_0^1 t^2 dt - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

2. Soit V une variable aléatoire de loi uniforme sur $[a,b]$, i.e. telle que la probabilité que V appartienne à un intervalle de longueur α, est proportionnelle à α et indépendante de la position de cet intervalle dans $[a,b]$. Ainsi, on obtient que la fonction de répartition de V est telle que pour $t \in [a,b]$

$$F_V(t) = \mathbb{P}(V \leqslant t) = \mathbb{P}(V \in [a,t]) = C(t-a),$$

où $C > 0$ est une constante à déterminer. Pour $t < a$, $F_V(t) = \mathbb{P}(V \leqslant t) = 0$ et pour $t > b$, $F_V(t) = \mathbb{P}(V \leqslant t) = 1$. Ainsi, la densité de V est donné par

$$f_V(t) = C\mathbb{1}_{[a,b]}(t),$$

et la contrainte $\displaystyle\int_{-\infty}^{+\infty} f_U(t)dt = 1$ implique que $C = \dfrac{1}{b-a}$. Pour calculer l'espérance et la variance de V, on peut passer par l'utilisation de la densité f_V, mais il est plus simple de remarquer que V peut s'écrire sous la forme

$$V = a + (b-a)U,$$

où U est une variable aléatoire de loi uniforme sur $[0,1]$. On obtient alors que

$$\mathbb{E}(V) = a + (b-a)\mathbb{E}(U) = \frac{a+b}{2} \text{ et } \mathrm{Var}(V) = (b-a)^2\mathrm{Var}(U) = \frac{(b-a)^2}{12}.$$

Exercice 2.3 : *Loi exponentielle*

1. La contrainte $\displaystyle\int_0^{+\infty} f_T(t)dt = 1$ implique que $k = \lambda$ car

$$\int_0^{+\infty} e^{-\lambda t}dt = \left[-\frac{e^{-\lambda t}}{\lambda}\right]_0^{+\infty} = \frac{1}{\lambda}.$$

2. Par définition,

$$\mathbb{E}(T) = \int_0^{+\infty} t f_T(t)dt = \int_0^{+\infty} t\lambda e^{-\lambda t}dt = \lambda\left[-t\frac{e^{-\lambda t}}{\lambda}\right]_0^{+\infty} + \lambda\int_0^{+\infty}\frac{e^{-\lambda t}}{\lambda}dt = \frac{1}{\lambda},$$

et

$$\mathrm{Var}(T) = \mathbb{E}(T^2) - (\mathbb{E}(T))^2 = \int_0^{+\infty} t^2\lambda e^{-\lambda t}dt - \frac{1}{\lambda^2} = \lambda\left[-t^2\frac{e^{-\lambda t}}{\lambda}\right]_0^{+\infty} + \int_0^{+\infty} 2te^{-\lambda t}dt - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

On peut remarque que l'écart-type $\sqrt{\mathrm{Var}(T)} = \mathbb{E}(T)$ est d'autant plus grand que λ est petit. La loi exponentielle est donc fortement dispersée.

3. Soit $F_T(t) = \mathbb{P}(T \leqslant t) = \int_0^t f_T(u)du = \int_0^t \lambda e^{-\lambda u}du = 1 - e^{-\lambda t}$, la fonction de répartition de T. On cherche $\alpha > 0$ tel que $\mathbb{P}(T < \alpha) = \mathbb{P}(T > \alpha)$ ce qui correspond à l'équation

$$\begin{aligned} F_T(\alpha) &= 1 - F_T(\alpha) \\ 1 - e^{-\lambda\alpha} &= e^{-\lambda\alpha}, \end{aligned}$$

ce qui donne $\alpha = -\frac{1}{\lambda}\ln\left(\frac{1}{2}\right) = \frac{\ln(2)}{\lambda}$.

<u>Exercice 2.4</u> : *Slalom*

Si T est le temps du parcours, on a $T = X + 4 \times N$, où X suit (approximativement) la loi normale $\mathcal{N}(50, 4)$, et N la loi binomiale $\mathcal{B}\left(20, \frac{1}{20}\right)$ (il y a 20 portes, chacune ayant la probabilité $p = \frac{1}{20}$ d'être accrochée et entrainer une chute). On cherche $\mathbb{P}([T < 52])$. On a :

$$\begin{aligned} \mathbb{P}([T < 52]) &= \mathbb{P}([X + 4N < 52]) = \mathbb{P}\left(\bigcup_{n=0}^{20}[X + 4N < 52] \cap [N = n]\right) \\ &= \mathbb{P}\left(\bigcup_{n=0}^{20}[X + 4n < 52] \cap [N = n]\right) = \mathbb{P}\left(\bigcup_{n=0}^{20}[X < 52 - 4n] \cap [N = n]\right) \\ &= \sum_{n=0}^{20}\mathbb{P}([X < 52 - 4n] \cap [N = n]). \end{aligned}$$

Comme X et N sont indépendantes, on a

$$\mathbb{P}([T < 52]) = \sum_{n=0}^{20}\mathbb{P}([X < 52 - 4n])\mathbb{P}([N = n]).$$

X suit la loi $\mathcal{N}(50, 4)$ donc $\frac{X - 50}{2}$ suit la loi $\mathcal{N}(0, 1)$, de fonction de répartition notée

$$\Phi(x) = F_Y(x) = \mathbb{P}(Y \leqslant x) \text{ pour } Y \text{ qui suit une loi } \mathcal{N}(0, 1),$$

et qui vérifie, pour $x < 0$, $\Phi(x) = 1 - \Phi(-x)$). Ainsi :

$$\mathbb{P}([X < 52 - 4n]) = \mathbb{P}\left(\left[\frac{X - 50}{2} < \frac{52 - 4n - 50}{2}\right]\right) = \Phi(1 - 2n)$$

et $\mathbb{P}([N = n]) = \binom{20}{n}p^n(1 - p)^{20-n} = \binom{20}{n}\left(\frac{1}{20}\right)^n\left(\frac{19}{20}\right)^{20-n}$ et donc :

$$\mathbb{P}([T < 52]) = \sum_{n=0}^{20} \Phi(1-2n) \binom{20}{n} \frac{19^{20-n}}{20^{20}}.$$

Pour $n = 0$, $\Phi(1) \approx 0,841$, pour $n = 1$, $\Phi(-1) \approx 0,159$, pour $n = 2$, $\Phi(-3) \approx 0,001$ et pour $n \geqslant 3$, $\Phi(1-2n)$ est négligeable. On peut donc négliger les valeurs de $n = 3$ à $n = 20$, ce qui allège agréablement les calculs, puisqu'il ne reste que 3 termes à calculer.

$$\mathbb{P}([T < 52]) \approx \Phi(1) \left(\frac{19}{20}\right)^{20} + 20\Phi(-1)\frac{19^{19}}{20^{20}} + 190\Phi(-3)\frac{19^{18}}{20^{20}} \approx 0,37.$$

Exercice 2.5 : *Loi Gamma*

1. A l'aide du changement de variable $t = \lambda x$, on obtient que

$$\int_{-\infty}^{+\infty} f_X(x)dx = \int_0^{+\infty} \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} dx = \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} e^{-t} t^{\alpha-1} dt = 1.$$

2. Par définition et en utilisant la relation $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f_X(x)dx = \int_0^{+\infty} \frac{\lambda^\alpha e^{-\lambda x} x^\alpha}{\Gamma(\alpha)} dx = \frac{\alpha}{\lambda} \int_0^{+\infty} \frac{\lambda^{\alpha+1} e^{-\lambda x} x^\alpha}{\Gamma(\alpha+1)} dx = \frac{\alpha}{\lambda}.$$

De même, on peut calculer

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x)dx = \int_0^{+\infty} \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha+1}}{\Gamma(\alpha)} dx = \frac{\alpha(\alpha+1)}{\lambda^2} \int_0^{+\infty} \frac{\lambda^{\alpha+2} e^{-\lambda x} x^{\alpha+1}}{\Gamma(\alpha+2)} dx = \frac{\alpha(\alpha+1)}{\lambda^2},$$

ce qui donne

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}.$$

3 Couple de variables aléatoires / Vecteurs aléatoires - PC3

__Énoncés des exercices__

__Exercice 3.1__ : *Problème de symétrie*
On considère une urne contenant N boules blanches et 2 boules noires. On tire les boules sans remise, les unes après les autres et on définit les variables aléatoires X et Y égales aux numéros de tirage des deux boules noires (avec $X < Y$).

1. Déterminer la loi de probabilité de la variable aléatoire X, ainsi que son espérance. On rappelle que
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$
2. Par un raisonnement simple, déterminer la loi du couple (X, Y). Dessiner le support du vecteur aléatoire (X, Y).
3. Déterminer la loi et l'espérance de Y.

__Exercice 3.2__ : *La rencontre*
Deux personnes A et B arrivent aléatoirement à un rendez-vous entre $5h$ et $6h$. Elles ne s'attendent pas plus de 10 minutes.

1. Quelle est la probabilité pour qu'elles se rencontrent ?
2. On note T_A et T_B les instants d'arrivée respectifs de A et B sur l'intervalle du rendez-vous de durée $d = 1$ heure. On note $T = |T_A - T_B|$ la variable aléatoire temps d'attente. Quelle est la loi de probabilité de T ?

__Exercice 3.3__ : *Covariance et indépendance*
Soit θ une variable aléatoire de loi uniforme sur $[0, 2\pi]$. On pose $X = \cos(\theta)$ et $Y = \sin(\theta)$.

1. Montrer que X et Y sont dépendantes.
2. Calculer la covariance $\text{Cov}(X, Y)$.

__Exercice 3.4__ : *Optimisation d'un multisenseur*
On considère deux dispositifs de mesure pour mesurer un même paramètre p. Les résultats des deux mesures sont des variables aléatoires X_1 et X_2 telles que
$$X_1 = p + U_1 \text{ et } X_2 = p + U_2,$$
où les variables aléatoires U_i représentent des erreurs de mesure indépendantes centrées et de variances σ_i^2 (niveau de bruit).

1. Déterminer l'estimateur du paramètre inconnu p sous la forme de la variable aléatoire $\widehat{p} = aX_1 + bX_2$, vérifiant $\mathbb{E}(\widehat{p}) = p$ ($\widehat{p}$ est alors dit sans biais) et ayant la plus petite variance possible, où a et b sont des réels à optimiser.

__Exercice 3.5__ : *Comptage de séries*
On effectue une suite de tirages successifs avec remise dans une urne contenant des jetons gagnants (portant le numéro 1), en proportion p ($p \in]0, 1[$), et des jetons perdants (portant le

numéro 0), en proportion $q = 1 - p$.
Soit X la longueur de la première "série" et Y la longueur de la deuxième "série". Ainsi, par exemple, pour la suite de tirages $111001011\cdots$, on a $X = 3$ et $Y = 2$ et pour $01101\cdots$, on a $X = 1$ et $Y = 2$. [*On rappelle que* $\displaystyle\sum_{k=1}^{+\infty} kx^k = \frac{x}{(1-x)^2}$ *pour* $|x| < 1$].

1. Déterminer la loi de la variable aléatoire X.
2. Montrer que X admet une espérance $\mathbb{E}(X) = \dfrac{p}{1-p} + \dfrac{1-p}{p}$. Montrer que $\mathbb{E}(X)$ est minimale lorsque $p = \dfrac{1}{2}$, et calculer cette valeur minimale.
3. Montrer, pour tout $(i,j) \in (\mathbb{N}^*)^2$, $\mathbb{P}([X = i] \cap [Y = j]) = p^{i+1}q^j + q^{i+1}p^j$.
4. En déduire la loi de la variable aléatoire Y et montrer que $\mathbb{E}(Y) = 2$.
5. Etablir que, si $p \neq \dfrac{1}{2}$, les variables aléatoires X et Y ne sont pas indépendantes. Démontrer que, si $p = \dfrac{1}{2}$, les variables aléatoires X et Y sont indépendantes.
6. Etablir que $\text{cov}(X,Y) = -\dfrac{(1-2p)^2}{p(1-p)}$.

Exercice 3.6 : *Covariance d'un signal numérique*
Soit un signal aléatoire stationnaire binaire modélisé par une suite double

$$(\ldots, X_{-n}, \ldots, X_{-1}, X_0, X_1, \ldots, X_n, \ldots)$$

de variables aléatoires indépendantes de même loi de Bernoulli de paramètre $p \in]0,1[$. Un montage d'électronique logique délivre en sortie le signal binaire $Y_n = \max(X_{n-1}, X_n)$.

1. Calculer la loi de Y_n, son espérance et sa variance.
2. Calculer la fonction de covariance de la suite Y_n : $C_n = \text{Cov}(Y_0, Y_n)$.

Corrigés des exercices

Exercice 3.1 : *Problème de symétrie*

Soit $\Omega = \{(x_1, x_2, \cdots, x_{N+2})$ avec 2 "B" et N "N"$\}$ l'ensemble des réalisations possibles de l'expérience, dont les éléments correspondent à une succession de boules blanches ou noires. Il y a C_{N+2}^2 choix de places pour les 2 boules blanches "B" donc card$\Omega = C_{N+2}^2 = \dfrac{(N+2)(N+1)}{2}$.

1. X est le numéro de la première boule blanche apparue donc $X \in [\![1, N+1]\!]$. L'évènement $[X = k]$ est réalisé si la première boule blanche apparaît au rang k et la deuxième à l'un des rangs suivants : il y a donc $N+2-k$ choix équiprobables pour la position de la deuxième boule blanche. Ainsi,

$$\mathbb{P}([X = k]) = \frac{2(N+2-k)}{(N+1)(N+2)}.$$

Par définition, on a alors que

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=1}^{N+1} k\mathbb{P}([X=k]) = \frac{1}{C_{N+2}^2}\left[(N+2)\sum_{k=1}^{N+1} k - \sum_{k=1}^{N+1} k^2\right] \\
&= \frac{1}{C_{N+2}^2}\left[(N+2)\frac{(N+1)(N+2)}{2} - \frac{(N+1)(N+2)(2N+3)}{6}\right] \\
&= N+2-\frac{2N+3}{3} = \frac{3N+6-2N-3}{3} = \frac{N+3}{3}.
\end{aligned}$$

Finalement, $\mathbb{E}(X) = \dfrac{1}{3}(N+3)$.

2. Le couple (X, Y) détermine un unique tirage complet et, comme tous ces tirages ont même probabilité et qu'il y en a C_{N+2}^2, la loi du couple (X, Y) est définie par :

$$\mathbb{P}([(X,Y) = (i,j)]) = \begin{cases} \dfrac{2}{(N+1)(N+2)} & \text{si } 1 \leqslant i < j \leqslant N+2 \\ 0 \text{ sinon.} & \end{cases}.$$

3. Connaissant la loi du couple (X, Y), on peut calculer la loi marginale

$$\mathbb{P}([Y = j]) = \sum_{i=1}^{j-1} \mathbb{P}([(X,Y) = (i,j)]) = \frac{2(j-1)}{(N+1)(N+2)},$$

pour $2 \leqslant j \leqslant N+2$.
Remarque : On peut aussi trouver Y directement en procédant comme pour X. Y est le numéro de la deuxième boule blanche apparue donc $Y(\Omega) \in [\![2, N+2]\!]$. L'évènement $[Y = k]$ est réalisé si la deuxième boule blanche apparaît au rang k et la première à l'un des rangs précédents : il y a donc $k-1$ choix équiprobables pour la position de la première boule blanche. Ainsi,

$$\mathbb{P}([Y = k]) = \frac{2(k-1)}{(N+1)(N+2)}.$$

Par définition,

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{k=1}^{N+1} k\mathbb{P}([Y=k]) = \frac{1}{C_{N+2}^2}\left[\sum_{k=2}^{N+2} k(k-1)\right] = \frac{1}{C_{N+2}^2}\left[\sum_{j=1}^{N+1} j(j+1)\right] \\
&= \frac{1}{C_{N+2}^2}\left[\frac{(N+1)(N+2)(2N+3)}{6} + \frac{(N+1)(N+2)}{2}\right] \\
&= \frac{2N+3}{3} + 1 = \frac{2N+3+3}{3} = \frac{2N+6}{3}.
\end{aligned}$$

Finalement, $\mathbb{E}(Y) = \frac{2}{3}(N+3) = 2\mathbb{E}(X)$.

Remarque : On peut aussi remarquer que pour $1 \leqslant k \leqslant N+1$, $\mathbb{P}([N+3-Y=k]) = \mathbb{P}([Y=N+3-k]) = \frac{N+2-k}{C_{N+2}^2} = \mathbb{P}([X=k])$, et donc les variables aléatoires $N+3-Y$ et X ont même loi. Ainsi on obtient que $\mathbb{E}(X) = \mathbb{E}(N+3-Y) = N+3-\mathbb{E}(Y)$ d'où $\mathbb{E}(Y) = N+3-\mathbb{E}(X) = \frac{2}{3}(N+3)$.

<u>Exercice 3.2</u> : *La rencontre*

1. On peut représenter l'espace des états, couples des temps d'arrivée de chaque personne, par le carré $[5,6]\times[5,6]$. Soit T_A et T_B les instants d'arrivée respectifs de A et B sur l'intervalle du rendez-vous de durée $d=1$ heure. L'événement "les personnes ne s'attendent pas plus de 10 minutes" correspond à l'évènement $|T_A - T_B| \leqslant \frac{1}{6}$ qui est une partie hachurée autour de la diagonale du carré $[5,6]\times[5,6]$, de surface $\frac{11}{36}$, ce qui donne

$$\mathbb{P}(\text{ "Les personnes ne s'attendent pas plus de 10 minutes" }) = \frac{11}{36}.$$

2. Afin de simplifier les écritures, on prend l'origine des temps à 5 heures et on se ramène ainsi au cas $(T_A, T_B) \in [0,1]\times[0,1]$. On va utiliser le fait que la densité de la somme de deux variables aléatoires indépendantes est égale au produit de convolution de leurs densités respectives. La variable aléatoire $T_A - T_B$ prend ses valeurs dans $[-1,1]$, et sa densité est donnée par

$$f_{T_A-T_B}(t) = \int f_{T_A}(s) f_{T_B}(s-t)\,ds = \int \mathbb{1}_{[0,1]}(s)\,\mathbb{1}_{[0,1]}(s-t)\,ds = \int \mathbb{1}_{[0,1]\cap[t,t+1]}(s)\,ds.$$

Deux cas se présentent :
- si $t \in [-1,0]$, alors $[0,1]\cap[t,t+1] = [0,t+1]$ et $f_{T_A-T_B}(t) = 1+t$
- si $t \in [0,1]$, alors $[0,1]\cap[t,t+1] = [t,1]$ et $f_{T_A-T_B}(t) = 1-t$.

D'où

$$f_{T_A-T_B}(t) = (1+t)\mathbb{1}_{[-1,0]}(t) + (1-t)\mathbb{1}_{[0,1]}(t).$$

Or, pour toute variable aléatoire réelle X, $f_{|X|}(x) = (f_X(x) + f_X(-x))\mathbb{1}_{[0,+\infty[}(x)$, d'où :

$$f_T(t) = 2(1-t)\mathbb{1}_{[0,1]}(t),$$

pour $T = |T_A - T_B|$.
On a alors que $\mathbb{P}\left(|T_A - T_B| \leqslant \frac{1}{6}\right) = 2\int_0^{1/6} (1-t)\,dt = \left[-(1-t)^2\right]_0^{1/6} = 1 - 25/36$, et on retrouve bien que

$$\mathbb{P}\left(|T_A - T_B| \leqslant \frac{1}{6}\right) = \frac{11}{36}.$$

<u>Exercice 3.3</u> : *Covariance et indépendance*

Remarquons tout d'abord que la densité de θ est $f_\theta(u) = \frac{1}{2\pi}\mathbb{1}_{[0,2\pi]}(u)$.

1. Comme $X^2 + Y^2 = 1$, il est clair que X et Y ne sont pas indépendantes car, par exemple,

$$\mathbb{P}\left([|X| < 1/\sqrt{2}] \cap [|Y| < 1/\sqrt{2}]\right) = 0 \neq \mathbb{P}\left(|X| < 1/\sqrt{2}\right)\mathbb{P}\left(|Y| < 1/\sqrt{2}\right).$$

2. $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$. Or,

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \cos(u) f_\theta(u) du = \frac{1}{2\pi}\int_0^{2\pi} \cos(u) du = 0.$$

De même

$$\mathbb{E}(Y) = \int_{-\infty}^{+\infty} \sin(u) f_\theta(u) du = \frac{1}{2\pi}\int_0^{2\pi} \sin(u) du = 0.$$

De plus

$$\mathbb{E}(XY) = \int_{-\infty}^{+\infty} \cos(u)\sin(u) f_\theta(u) du = \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{2}\sin(2u) du = 0.$$

On a donc que $\mathrm{Cov}(X,Y) = 0$ alors que X et Y ne sont pas indépendantes.

<u>Exercice 3.4</u> : *Optimisation d'un multisenseur*

1. L'espérance mathématique étant linéaire, on a

$$\mathbb{E}(\hat{p}) = a\mathbb{E}(X_1) + b\mathbb{E}(X_2) = ap + bp = (a+b)p\ ;$$

si $\hat{p}$ est sans biais, alors $a + b = 1$.
D'autre part, les variables X_1 et X_2 sont indépendantes (car les variables U_1 et U_2 le sont). On a alors

$$\mathrm{Var}(\hat{p}) = a^2\mathrm{Var}(X_1) + b^2\mathrm{Var}(X_2) = a^2\sigma_1^2 + b^2\sigma_2^2.$$

Il s'agit donc de minimiser $\phi(a,b) = a^2\sigma_1^2 + b^2\sigma_2^2$ sous la contrainte $a + b = 1$, ce qui revient à chercher le minimum de la fonction $\psi : \mathbb{R} \to \mathbb{R}$, $a \mapsto \psi(a) = a^2\sigma_1^2 + (1-a)^2\sigma_2^2$.

$\psi'(a) = 2a\sigma_1^2 - 2(1-a)\sigma_2^2$ est nul pour $a(\sigma_1^2 + \sigma_2^2) = \sigma_2^2$, soit $a = \dfrac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$ et $b = \dfrac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$.

Exercice 3.5 : *Comptage de séries*

On notera A_k l'événement "le k^{ieme} jeton tiré est gagnant".

1. Notons que $X \in N^*$ et que l'évènement $[X = k]$ signifie que la première série a pour longueur k, c'est-à-dire que les k premiers jetons sont gagnants et le $(k+1)^{ieme}$ perdant ou bien les k premiers jetons sont perdants et le $(k+1)^{ieme}$ gagnant. On a donc que :
$[X = k] = (A_1 \cap \cdots \cap A_k \cap \overline{A}_{k+1}) \cup (\overline{A}_1 \cap \cdots \cap \overline{A}_k \cap A_{k+1})$ et donc

$$\mathbb{P}([X = k]) = p^k q + q^k p, \text{ pour } k \in N^*.$$

2. Par définition

$$\mathbb{E}(X) = \sum_{k=1}^{+\infty} kP([X = k]) = \sum_{k=1}^{+\infty} kp^k(1-p) + \sum_{k=1}^{+\infty} k(1-p)^k p.$$

En utilisant les espérances de lois géométriques, on a alors $\mathbb{E}(X) = \dfrac{p}{1-p} + \dfrac{1-p}{p}$.

Soit $f(x) = \dfrac{x}{1-x} + \dfrac{1-x}{x}$: $f'(x) = \dfrac{1-x+x}{(1-x)^2} - \dfrac{1}{x^2} = \dfrac{x^2-(1-x)^2}{x^2(1-x)^2} = \dfrac{2x-1}{x^2(1-x)^2}$.
f' est négative sur $]0, \dfrac{1}{2}[$, nulle en $\dfrac{1}{2}$ et positive sur $]\dfrac{1}{2}, 1[$, et $\mathbb{E}(X) = f(p)$. On a donc :

$$\mathbb{E}(X) \text{ minimale lorsque } p = \frac{1}{2}, \text{ et cette valeur minimale est } f(\frac{1}{2}) = 2.$$

3. L'évènement $[X = i] \cap [Y = j]$ signifie que la première série est de longueur i et la deuxième de longueur j. On a donc, en notant $p_{ij} = P([X = i] \cap [Y = j])$,

$$p_{ij} = P\left[\left(A_1 \cap \cdots \cap A_i \cap \overline{A}_{i+1} \cap \cdots \overline{A}_{i+j} \cap A_{i+j+1}\right) \cup \left(\overline{A}_1 \cap \cdots \cap \overline{A}_i \cap A_{i+1} \cap \cdots A_{i+j} \cap \overline{A}_{i+j+1}\right)\right]$$

et, toujours en utilisant l'indépendance des tirages, on obtient :

$$\text{pour tout } (i,j) \in (N^*)^2,\ P([X = i] \cap [Y = j]) = p^{i+1}q^j + q^{i+1}p^j.$$

4. La loi marginale de Y est donnée par $P([Y = j]) = \sum_{i=1}^{+\infty} P([X = i] \cap [Y = j]) = \sum_{i=1}^{+\infty} p^{i+1}q^j + \sum_{i=1}^{+\infty} q^{i+1}p^j$ avec $\sum_{i=1}^{+\infty} p^{i+1} = \sum_{i=2}^{+\infty} p^i = \dfrac{p^2}{q}$ et $\sum_{i=1}^{+\infty} q^{i+1} = \dfrac{q^2}{p}$. On en déduit alors que

$$P([Y = j]) = p^2 q^{j-1} + q^2 p^{j-1} \text{ pour } j \in N^*.$$

Par définition

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

soit $\mathbb{E}(Y) = 2$.

5. Notons que $P([X = 1] \cap [Y = 1]) = p^2q + pq^2 = pq(p+q) = pq$. D'autre part $P([X = 1]) = 2pq$ et $P([Y = 1]) = p^2 + q^2 = 2p^2 - 2p + 1$. On a donc :

$$P([X = 1])P([Y = 1]) - P([X = 1] \cap [Y = 1]) = pq(4p^2 - 4p + 2 - 1) = pq(2p-1)^2.$$

Si $p \neq \frac{1}{2}$, $P([X = 1] \cap [Y = 1]) \neq P([X = 1])P([Y = 1])$ et donc X et Y ne sont pas indépendantes.

Si $p = q = \frac{1}{2}$, alors pour tout $(i, j) \in (\mathbb{N}^*)^2$, on a que $P([X = i] \cap [Y = j]) = \frac{1}{2^{i+j}}$, $P([X = i]) = \frac{1}{2^i}$ et $P([Y = j]) = \frac{1}{2^j}$, donc $P([X = i] \cap [Y = j]) = P([X = i])P([Y = j])$ et donc

$$\text{si } p = \frac{1}{2}, \text{ les variables aléatoires } X \text{ et } Y \text{ sont indépendantes.}$$

6. On peut remarquer que $\mathbb{E}(XY) = \sum_{i,j\geqslant 1} ijp_{ij} = \left(\sum_{i=1}^{+\infty} ip^{i+1}\right)\left(\sum_{j=1}^{+\infty} jq^j\right) + \left(\sum_{i=1}^{+\infty} ip^i\right)\left(\sum_{j=1}^{+\infty} jq^{j+1}\right) = \frac{p^2}{q^2}\frac{q}{p^2} + \frac{p}{q^2}\frac{q^2}{p^2}$ donc $\mathbb{E}(XY) = \frac{1}{q} + \frac{1}{p} = \frac{1}{pq}$ et $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{pq} - 2\frac{p^2 + q^2}{pq}$, soit

$$\operatorname{cov}(X, Y) = -\frac{(1 - 2p)^2}{p(1 - p)}.$$

Exercice 3.6 : *Covariance d'un signal numérique*
Soit un signal aléatoire stationnaire binaire modélisé par une suite double

$$(\ldots, X_{-n}, \ldots, X_{-1}, X_0, X_1, \ldots, X_n, \ldots)$$

de variables aléatoires indépendantes de même loi de Bernoulli de paramètre $p \in]0, 1[$. Un montage d'électronique logique délivre en sortie le signal binaire $Y_n = \max(X_{n-1}, X_n)$.

1. Calculer la loi de Y_n, son espérance et sa variance.
2. Calculer la fonction de covariance de la suite Y_n : $C_n = \operatorname{Cov}(Y_0, Y_n)$.

1. Etant donné que X_n suit une loi de Bernoulli de paramètre $p \in]0, 1[$, on a que

$$\mathbb{P}(X_n = 1) = p,\ \mathbb{P}(X_n = 0) = 1 - p,\ \mathbb{E}(X_n) = p,\ \operatorname{Var}(X_n) = p(1 - p).$$

Pour $Y_n = \max(X_{n-1}, X_n) \in \{0, 1\}$, on a que (par indépendance de X_{n-1} et X_n)

$$\mathbb{P}(Y_n = 0) = \mathbb{P}\left([X_{n-1} = 0] \cap [X_n = 0]\right) = \mathbb{P}\left(X_{n-1} =\right)\mathbb{P}\left(X_n = 0\right) = (1 - p)^2,$$

et

$$\mathbb{P}(Y_n = 1) = 1 - \mathbb{P}(Y_n = 0) = 1 - (1 - p)^2 = 2p - p^2.$$

On obtient alors que

$$\mathbb{E}(Y_n) = 0 \times \mathbb{P}(Y_n = 0) + 1 \times \mathbb{P}(Y_n = 1) = 2p - p^2,$$

et

$$\operatorname{Var}(Y_n) = \mathbb{E}(Y_n^2) - (\mathbb{E}(Y_n))^2 = 0^2 \times \mathbb{P}(Y_n = 0) + 1^2 \times \mathbb{P}(Y_n = 1) - (1 - (1 - p)^2)^2 = (2p - p^2)(1 - p)^2.$$

2. Pour $n \geqslant 2$, il est clair que Y_n et Y_0 sont indépendantes, et donc $C_n = \operatorname{Cov}(Y_0, Y_n) = 0$. Comme $C_0 = \operatorname{Cov}(Y_0, Y_0) = \operatorname{Var}(Y_0) = (2p - p^2)(1 - p)^2$, il reste à calculer

$$C_1 = \operatorname{Cov}(Y_0, Y_1) = \mathbb{E}(Y_0Y_1) - \mathbb{E}(Y_0)\mathbb{E}(Y_1) = \mathbb{E}(Y_0Y_1) - (2p - p^2)^2.$$

On a que $\mathbb{E}(Y_0Y_1) = 1 \times \mathbb{P}\left([Y_0 = 1] \cap [Y_1 = 1]\right)$ car si Y_0 ou Y_1 est nul alors le produit Y_0Y_1 est nul. Il suffit ensuite de remarquer que

$$\begin{aligned}
\mathbb{P}\left([Y_0 = 1] \cap [Y_1 = 1]\right) &= \mathbb{P}\left(\{[X_{-1} = 1] \cup [X_0 = 1]\} \cap \{[X_0 = 1] \cup [X_1 = 1]\}\right) \\
&= \mathbb{P}(X_{-1} = 1) + \mathbb{P}\left(\{[X_{-1} = 1] \cap [X_1 = 1]\}\right) - \mathbb{P}\left(\{[X_{-1} = 1] \cap [X_0 = 1] \cup [X_1 = 1]\}\right) \\
&= \mathbb{P}(X_{-1} = 1) + \mathbb{P}(X_{-1} = 0)\mathbb{P}\left(X_{-1} = 1\right)\mathbb{P}\left(X_1 = 1\right) \\
&= p + (1-p)p^2.
\end{aligned}$$

Et donc $C_1 = p + (1-p)p^2 - (2-p)p^2 = p(1-p)$.

4 Calcul de lois - PC4 & PC5

Énoncés des exercices

Exercice 4.1 : *Max et Min de lois uniformes*
Soient U et V deux variables aléatoires réelles indépendantes de loi uniforme sur $[0, 1]$.

1. Déterminer la loi de $S = \max(U, V)$, l'espérance et la variance de S.
2. Déterminer la loi de $X = \min(U, V)$, l'espérance et la variance de X.

Exercice 4.2 : *Loi de Cauchy*
Soit X une variable aléatoire continue de loi uniforme sur $]-\pi/2, \pi/2[$.

1. Déterminer la loi de $Y = \tan(X)$.

Exercice 4.3 : *Loi exponentielle*
Soit V une variable aléatoire de loi uniforme sur $]0, 1]$.

1. Déterminer la loi de $U = -\ln(V)$.
2. Déterminer l'espérance et la variance de U.

Exercice 4.4 : *Simulation*
Soit X une variable aléatoire réelle continue de fonction de répartition F_X et de densité $f_X(x) > 0$ pour tout $x \in \mathbb{R}$. On pose $Y = F_X(X)$.

1. Calculer la loi de Y.
2. En déduire une méthode permettant de simuler n'importe quelle variable aléatoire réelle à partir de tirages aléatoires uniformes sur $[0, 1]$.
3. Application : simuler la réalisation d'une variable aléatoire X de loi exponentielle de paramètre $\lambda > 0$.

Exercice 4.5 : *Loi d'une combinaison linéaire*
Soit X une variable aléatoire continue sur $\mathbb{R}$ de densité de probabilité f_X. Soient a, b deux réels avec $a > 0$.

1. Quelle est la densité de probabilité de la variable aléatoire $Y = aX + b$?

Exercice 4.6 : *Loi de X^2*
Soit X une variable aléatoire continue sur $\mathbb{R}$ de densité de probabilité f_X.

1. Quelle est la densité de probabilité de la variable aléatoire $Y = X^2$?
2. Application à la loi normale centrée réduite.

Exercice 5.1 : *Loi de $X + Y$*
Soient X et Y deux variables aléatoires continues et indépendantes de densité respective f_X et f_Y.

1. Déterminer la loi de $S = X + Y$.
2. Exemple 1 : X et Y suivent une loi uniforme sur $[0, 1]$.
3. Exemple 2 : X et Y suivent une loi exponentielle de paramètre $\lambda = 1$.
4. Exemple 3 : X et Y suivent la première loi de Laplace de densité $\frac{1}{2}e^{-|t|}$, $t \in \mathbb{R}$.

Exercice 5.2 : *Sommes de variables aléatoires de loi exponentielle*
Soient X et Y deux variables aléatoires réelles indépendantes et de même loi exponentielle de paramètre $\lambda = 1$.

1. Déterminer la loi du vecteur aléatoire (U, V) où $U = X + Y$ et $V = \dfrac{X}{X+Y}$.
2. En déduire les lois marginales de U et de V.
3. Quelle est la loi de la somme de n v.a. exponentielles indépendantes de paramètre λ ?

Exercice 5.3 : *Lois uniformes*
Soient X et Y deux variables aléatoires réelles indépendantes de loi uniforme sur $[0, 1]$.

1. Quelle est la loi de $T = \dfrac{X}{Y}$?
2. T a-t-elle une espérance mathématique ?

Exercice 5.4 : *Min et Max*
Soient $T_1, T_2, T_3, \ldots T_n$ des variables aléatoires continues indépendantes et de même loi de fonction de répartition F_T et de densité f_T . On note

$$T_{min} = \min_{1 \leqslant i \leqslant n} \{T_i\} \text{ et } T_{max} = \max_{1 \leqslant i \leqslant n} \{T_i\}.$$

1. Exprimer en fonction de F_T les fonctions de répartition de T_{max} et T_{min}.
2. En déduire les densités de probabilités de T_{max} et T_{min}.
3. Application : les durées de vie de 2 composants C_1 et C_2 sont des variables aléatoires indépendantes T_1 et T_2 de même loi exponentielle de paramètre λ. Déterminer la durée de vie Y du modèle série (les 2 éléments sont nécessaires) et celle Z du modèle parallèle (un au moins des 2 éléments est nécessaire) en exprimant les lois de probabilités associées.

Exercice 5.5 : *Produit de variables aléatoires mixtes*
Soit les variables aléatoires indépendantes : X normale centrée réduite et B définie par $\mathbb{P}(B = 1) = \mathbb{P}(B = -1) = 1/2$. Posons $Y = BX$.

1. Montrer que Y est normale centrée réduite.
2. X et Y sont-elles dépendantes ? Décorrélées ?
3. Calculer $\mathbb{P}(X + Y = 0)$. La loi de $X + Y$ est-elle absolument continue ?

Exercice 5.6 : *Coordonnées sphériques*
Soit le vecteur normal (X_1, X_2, X_3) de composantes indépendantes, chacune de loi $\mathcal{N}(0,1)$. Soit T la bijection :

$$(x_1, x_2, x_3) \in \mathbb{R}^3 \quad \longrightarrow \quad T(x_1, x_2, x_3) = (r, \phi, \theta) \in [0, +\infty[\times[0, \pi[\times[-\pi, \pi[$$

avec

$$\begin{cases} x_1 = r \sin\phi \cos\theta \\ x_2 = r \sin\phi \sin\theta \\ x_3 = r \cos\phi \end{cases}$$

Soit (R, Φ, Θ) le vecteur aléatoire image de (X_1, X_2, X_3) par T.

1. Calculer la densité conjointe de (R, Φ, Θ), ainsi que les densités marginales. En déduire le caractère isotrope du vecteur (X_1, X_2, X_3).
2. Calculer la loi de R.
3. Calculer la loi de R^2.

Corrigés des exercices

Exercice 4.1 : *Max et Min de lois uniformes*

Etant donné que U et V dont de loi uniforme sur $[0,1]$ on a que

$$F_U(u) = \mathbb{P}(U \leqslant u) = u \text{ et } f_U(u) = 1,\ u \in [0,1],$$

et

$$F_V(v) = \mathbb{P}(V \leqslant v) = v \text{ et } f_V(v) = 1,\ v \in [0,1].$$

1. Calculons la fonction de répartition de S. Par l'indépendance de U et V on a que

$$F_S(s) = \mathbb{P}(S \leqslant s) = \mathbb{P}([U \leqslant s] \cap [V \leqslant s]) = \mathbb{P}(U \leqslant s)\mathbb{P}(V \leqslant s) = s^2,\ s \in [0,1].$$

et donc

$$f_S(s) = 2s\mathbb{1}_{[0,1]}(s).$$

On a alors que

$$\mathbb{E}(S) = \int_{-\infty}^{+\infty} s f_S(s)ds = \int_0^1 2s^2 ds = \frac{2}{3},$$

et

$$\mathrm{Var}(S) = \mathbb{E}(S^2) - (\mathbb{E}(S))^2 = \int_{-\infty}^{+\infty} s^2 f_S(s)ds - \frac{4}{9} = \int_0^1 2s^3 ds = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

2. Calculons la fonction de répartition de X. Par l'indépendance de U et V on a que

$$F_X(x) = \mathbb{P}(X \leqslant x) = 1-\mathbb{P}(X > x) = 1-\mathbb{P}([U > x]\cap[V > x]) = 1-\mathbb{P}(U > x)\mathbb{P}(V > x) = 1-(1-x)^2,\ x \in [0,1].$$

et donc

$$f_X(x) = 2(1-x)\mathbb{1}_{[0,1]}(x).$$

On a alors que

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f_X(x)dx = \int_0^1 2x(1-x)dx = \frac{1}{3},$$

et

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \int_{-\infty}^{+\infty} x^2 f_X(x)dx - \frac{1}{9} = \int_0^1 2x^2(1-x)ds = 2\left(\frac{1}{3} - \frac{1}{4}\right) - \frac{1}{9} = \frac{1}{18}.$$

Exercice 4.2 : *Loi de Cauchy*
Soit X une variable aléatoire continue de loi uniforme sur $]-\pi/2, \pi/2[$.

1. Déterminer la loi de $Y = \tan(X)$.

Etant donné que X est de loi uniforme sur $]-\pi/2, \pi/2[$, on a que

$$f_X(x) = \frac{1}{\pi}\mathbb{1}_{]-\pi/2,\pi/2[}(x),\ \text{et } F_X(x) = \frac{x}{\pi} + 1/2,\ x \in [0,1].$$

1. Calculons la fonction de répartition de Y :

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(\tan(X) \leqslant x) = \mathbb{P}(X \leqslant \arctan(y)) = \frac{\arctan(y)}{\pi} + 1/2,$$

et donc

$$f_Y(y) = F'_Y(y) = \frac{1}{\pi}\frac{1}{(1+y^2)},\ y \in \mathbb{R}.$$

On peut remarquer que la fonction $y \mapsto \dfrac{y}{(1+y^2)}$ n'est pas intégrable sur $\mathbb{R}$, et donc l'espérance de Y n'existe pas !

<u>Exercice 4.3</u> : *Loi exponentielle*

1. Calculons la fonction de réparation de U :

$F_U(u) = \mathbb{P}(U \leqslant u) = \mathbb{P}(-\ln(V) \leqslant u) = \mathbb{P}(\ln(V) \geqslant -u) = 1-\mathbb{P}(V \leqslant e^{-u}) = 1-F_V(e^{-u}) = 1-e^{-u}$,

car V est de loi uniforme sur $[0,1]$. Ainsi,

$$f_U(u) = F'_U(u) = e^{-u}\mathbb{1}_{[0,+\infty[}(u),$$

et donc on obtient que U est de loi exponentielle de paramètre $\lambda = 1$.

2. Lors des exercices précédents, on a montré que

$$\mathbb{E}(X) = \frac{1}{\lambda} = 1 \text{ et } \mathrm{Var}(X) = \frac{1}{\lambda^2} = 1.$$

<u>Exercice 4.4</u> : *Simulation*

1. Déterminons la fonction de répartition de Y :

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(F_X(X) \leqslant y) = \mathbb{P}(X \leqslant F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y,\ y \in [0,1].$$

Donc Y est une variable aléatoire de loi uniforme sur $[0,1]$.

2. Ainsi, si on dispose d'un générateur d'un nombre aléatoire Y qui suit une loi uniforme sur $[0,1]$ on peut générer n'importe quelle variable aléatoire X réelle dont on sait calculer la fonction de répartition en prenant $X = F_X^{-1}(Y)$.

3. Il suffit de remarquer que si X suit une loi exponentielle de paramètre $\lambda > 0$, alors $F_X(x) = 1 - e^{-\lambda x}$ et donc on peut facilement calculer son inverse donnée par $F_X^{-1}(y) = -\dfrac{1}{\lambda}\ln(1-y)$.

<u>Exercice 4.5</u> : *Loi d'une combinaison linéaire*

1. La fonction de répartition de Y est :

$$F_Y(y) = \mathbb{P}(aX + b \leqslant y) = \mathbb{P}\left(X \leqslant \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$

Par dérivation on obtient donc que :

$$f_Y(y) = F'_Y(y) = \frac{1}{a}f_X\left(\frac{y-b}{a}\right)$$

Exercice 4.6 : *Loi de X^2*

1. Déterminons la fonction de réparation de Y :

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}),\ y \in [0, +\infty[,$$

et donc

$$f_Y(y) = F_Y'(y) = \frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right) \mathbb{1}_{[0,\infty[}(y).$$

1. Si X est une variable aléatoire de loi normale centrée réduite, alors $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ et donc

$$f_Y(y) = \frac{1}{2\sqrt{\pi y}} e^{-\frac{y}{2}} \mathbb{1}_{[0,\infty[}(y).$$

qui est la densité d'une variable aléatoire du Chi2 à 1 degré de liberté.

Exercice 5.1 : *Loi de $X + Y$*

1. Il a été vu en cours que

$$f_S(x) = f_X \star f_Y(x) = \int_{-\infty}^{+\infty} f_X(u) f_Y(x-u) du.$$

La densité de probabilité de la somme de 2 variables aléatoires indépendantes est le produit de convolution des densités de probabilité de ces 2 variables aléatoires.

Pour les trois exemples, il suffit de faire le calcul du produit de convolution des densités de X et de Y. On donne seulement les résultats :

2. $f_S(x) = s\mathbb{1}_{[0,1]}(s) + (2-s)\mathbb{1}_{]1,2]}(s)$.
3. $f_S(x) = se^{-s}\mathbb{1}_{[0,+\infty[}(s)$.
4. $f_S(x) = \frac{1}{4}e^{-|s|}(1+|s|)$.

Exercice 5.2 : *Sommes de variables aléatoires de loi exponentielle*

1. Etant donné que X et Ysont des variables aléatoires indépendantes de même loi exponentielle de paramètre $\lambda = 1$ la densité du couple (X, Y) est donné par

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) = \lambda^2 e^{-\lambda(x+y)} \mathbf{1}_{]0,+\infty[}(x) \mathbf{1}_{]0,+\infty[}(y).$$

Soit l'application $h : (x, y) \mapsto (u, v) = \left(x + y, \dfrac{x}{x+y}\right)$ définie sur $\mathbb{R}^2 \setminus \{(x, -x),\ x \in \mathbb{R}\}$ à valeurs dans $\mathbb{R}^2 \setminus (Oy)$ et son inverse $h^{-1} : (u, v) \mapsto (x, y) = (uv, u(1-v))$ pour $u \neq 0$.

$$f_{U,V}(u, v) = f_{X,Y}(h^{-1}(u, v)) \times |J_{h^{-1}}(u, v)|$$

avec $J_{h^{-1}}(u, v) = \begin{vmatrix} v & u \\ 1-v & -u \end{vmatrix} = -uv - u + uv = -u$. Il suffit maintenant de déterminer

$$\begin{aligned} f_{X,Y}(h^{-1}(u, v)) &= f_{X,Y}(uv, u(1-v)) \\ &= \lambda^2 e^{-\lambda u} \mathbf{1}_{]0,+\infty[}(uv)\, \mathbf{1}_{]0,+\infty[}(u(1-v)) \end{aligned}$$

avec $\begin{cases} uv > 0 \\ u(1-v) > 0 \end{cases}$ si et seulement si $\begin{cases} uv > 0 \\ u^2v(1-v) > 0 \\ u(1-v) > 0 \end{cases}$, soit $\begin{cases} v \in]0,1[\\ u > 0 \end{cases}$. Par la formule du changement de variables on a donc que :

$$f_{U,V}(u,v) = \lambda^2 e^{-\lambda u} u \mathbf{1}_{]0,+\infty[}(u)\, \mathbf{1}_{]0,1[}(v)$$

2. Les densités marginales de U et de V sont donc :

$$f_V(v) = \int f_{U,V}(u,v)\, du = \lambda^2 \mathbf{1}_{]0,1[}(v) \int_0^{+\infty} u\, e^{-\lambda u}\, du = \mathbf{1}_{]0,1[}(v).$$

Ainsi, V suit la loi uniforme sur $[0,1]$. Par ailleurs, on a immédiatement que :

$$f_U(u) = \lambda^2 e^{-\lambda u} u \mathbf{1}_{]0,+\infty[}(u).$$

Ainsi, U suit la loi Gamma $\gamma(\lambda, 2)$.

3. On vient de montrer qte si X_1 et X_2 sont des variables aléatoires indépendantes de loi exponentielle de paramètre λ, alors $X_1 + X_2$ suit une loi Gamma $\gamma(\lambda, 2)$. Pour pouvoir déduire la loi de la somme de n v.a. exponentielles indépendantes de paramètre λ, il faut pouvoir sommer deux variables aléatoires indépendantes X et Y, respectivement de loi $\gamma(\lambda, \alpha)$ et $\gamma(\lambda, \beta)$, dont la densité jointe est donné par :

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda(x+y)} x^{\alpha-1} y^{\beta-1} \mathbf{1}_{]0,+\infty[}(x) \mathbf{1}_{]0,+\infty[}(y).$$

En utilisant le même changement de variables que précédemment i.e. $h : (x,y) \mapsto (u,v) = \left(x+y, \dfrac{x}{x+y}\right)$ et son inverse $h^{-1} : (u,v) \mapsto (x,y) = (uv, u(1-v))$, on obtient que

$$f_{U,V}(u,v) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda u} v^{\alpha-1} (1-v)^{\beta-1} u^{\alpha+\beta-1} \mathbf{1}_{]0,+\infty[}(u)\, \mathbf{1}_{]0,1[}(v)$$

$$f_U(u) = \int f_{U,V}(u,v)\, dv = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda u} u^{\alpha+\beta-1} \mathbf{1}_{]0,+\infty[}(u) \int_0^1 v^{\alpha-1}(1-v)^{\beta-1}\, dv$$

$$f_V(v) = \int f_{U,V}(u,v)\, du = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} v^{\alpha-1}(1-v)^{\beta-1}\, \mathbf{1}_{]0,1[}(v) \int_0^{+\infty} u^{\alpha+\beta-1}\, e^{-\lambda u}\, du.$$

Or $\displaystyle\int_0^{+\infty} u^{\alpha+\beta-1}\, e^{-\lambda u}\, du \overset{(t=\lambda u)}{=} \int_0^{+\infty} t^{\alpha+\beta-1}\, e^{-t}\, dt \times \frac{1}{\lambda^{\alpha+\beta}} = \frac{\Gamma(\alpha+\beta)}{\lambda^{\alpha+\beta}}$ donc

$$f_V(v) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, v^{\alpha-1}(1-v)^{\beta-1}\, \mathbf{1}_{]0,1[}(v).$$

Ainsi, V suit la loi Bêta de paramètres α et β. Comme f_V est une densité, on a en particulier $\displaystyle\int f_V(v)\, dv = 1$ soit

$$\int_0^1 v^{\alpha-1}(1-v)^{\beta-1}\, dv = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

On a alors $f_U(u) = \dfrac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda u} u^{\alpha+\beta-1} \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \mathbf{1}_{]0,+\infty[}(u)$, soit

$$f_U(u) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha+\beta)}\, u^{\alpha+\beta-1}\, e^{-\lambda u}\, \mathbf{1}_{]0,+\infty[}(u).$$

Ainsi, $U = X + Y$ suit la loi Gamma $\gamma\,(\lambda, \alpha+\beta)$.

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes de même loi exponentielle de paramètre λ. Par récurrence, supposons maintenant que $Z_n = X_1 + \cdots + X_n$ suit la loi Gamma $\gamma(\lambda, n)$. Alors $Z_{n+1} = Z_n + X_{n+1}$ avec Z_n qui suit une loi Gamma $\gamma(\lambda, n)$, $P_{X_{n+1}} = \gamma(\lambda, 1)$, et Z_n et X_{n+1} indépendantes donc d'après ce qui précède $Z_n + X_{n+1}$ suit la loi Gamma $\gamma(\lambda, n+1)$ car la loi $\gamma(\lambda, 1)$ correspond à la loi exponentielle de paramètre λ. Ainsi, si les X_i sont indépendantes de loi exponentielle de paramètre λ, alors $X_1 + \cdots + X_n$ suit la loi $\gamma(\lambda, n)$.

Plus généralement, avec ce qui précède, on peut faire le même raisonnement avec la loi Gamma : si les X_i suivent la loi Gamma $\gamma(\lambda, \alpha)$ et sont indépendantes, alors $X_1 + \cdots + X_n$ suit la loi Gamma $\gamma(\lambda, n\alpha)$.

Exercice 5.3 : *Lois uniformes*

1. On commence par déterminer la loi de $h(X,Y) = \left(\dfrac{X}{Y}, Y\right)$ par la formule du changement de variables :

$$f_{h(X,Y)}(u,v) = f_{X,Y}(h^{-1}(u,v)) \times |J_{h^{-1}}(u,v)|$$

car h est une bijection de $\mathbb{R} \times \mathbb{R}^*$ dans $\mathbb{R} \times \mathbb{R}^*$. Pour déterminer h^{-1}, on a $h(x,y) = (u,v) = \left(\dfrac{x}{y}, y\right)$, donc $y = v$ et $x = uv$ et $h^{-1} : (u,v) \mapsto (uv, v)$, puis $J_{h^{-1}}(u,v) = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$ d'où

$$f_{\left(\frac{X}{Y}, Y\right)}(u,v) = f_{(X,Y)}(uv, v)|v|.$$

On a alors :

$$f_{X/Y}(u) = \int f_{X,Y}\,(uv, v)\,|v|\,dv = \int \mathbf{1}_{]0,1[}(uv)\mathbf{1}_{]0,1[}(v)\,v\,dv = \int \mathbf{1}_{]0,1/u[\cap]0,1[}(v)\,v\,dv.$$

Il faut distinguer deux cas :

- $0 < u \leqslant 1$ qui implique $f_{X/Y}(u) = \displaystyle\int_0^1 v\,dv = \frac{1}{2}$,
- $1 < u$ qui implique $f_{X/Y}(u) = \displaystyle\int_0^{1/u} v\,dv = \frac{1}{2u^2}$, et donc finalement

$$f_{X/Y}(u) = \frac{1}{2}\mathbf{1}_{]0,1]}(u) + \frac{1}{2u^2}\mathbf{1}_{]1,+\infty[}(u).$$

2. Etant donné que $u\,f_{X/Y}(u) \sim \dfrac{1}{2u}$ en $+\infty$, qui n'est pas intégrable, on obtient que X/Y n'admet pas d'espérance.

Exercice 5.4 : *Min et Max*

1. Calculons la fonction de répartition de T_{max}. Par indépendance de $T_1, \ldots, T_n$ on a que :

$$F_{T_{max}}(t) = \mathbb{P}([T_1 \leqslant t] \cap \cdots \cap [T_n \leqslant t]) = \prod_{i=1}^{n} \mathbb{P}(T_i \leqslant t) = (F_T(t))^n,\ t \in \mathbb{R},$$

et

$$F_{T_{min}}(t) = 1 - \mathbb{P}([T_1 > t] \cap \cdots \cap [T_n > t]) = 1 - \prod_{i=1}^{n} (1 - \mathbb{P}(T_i \leqslant t)) = 1 - (1 - F_T(t))^n, \ t \in \mathbb{R}.$$

2. Par dérivation, on alors que

$$f_{T_{max}}(t) = F'_{T_{max}}(t) = n f_T(t)(F_T(t))^{n-1}, \ t \in \mathbb{R},$$

et

$$f_{T_{min}}(t) = F'_{T_{min}}(t) = n f_T(t)(1 - F_T(t))^{n-1}, \ t \in \mathbb{R}.$$

3. On a que $Y = \min(T_1, T_2)$ et $Z = \max(T_1, T_2)$ et donc d'après ce qui précède :

$$f_Z(t) = 2\lambda e^{-\lambda t}(1 - e^{-\lambda t})\mathbb{1}_{[0,+\infty[}(t)$$

et

$$f_Y(t) = 2\lambda e^{-2\lambda t}\mathbb{1}_{[0,+\infty[}(t).$$

<u>Exercice 5.5</u> : *Produit de variables aléatoires mixtes*

1. Calculons la fonction de répartition de Y :

$$\begin{aligned}
F_Y(z) &= \mathbb{P}([Y \leqslant z]) = \mathbb{P}([XB \leqslant z]) \\
&= \mathbb{P}\left(([XB \leqslant z] \cap [B = -1]) \cup ([XB \leqslant z] \cap [B = 1])\right) \\
&= \mathbb{P}\left(([-X \leqslant z] \cap [B = -1]) \cup ([X \leqslant z] \cap [B = 1])\right) \\
&= \frac{1}{2}\left(\mathbb{P}([X \geqslant -z]) + \mathbb{P}([X \leqslant z])\right) \\
&= \frac{1}{2}\left(1 - F_X(-z) + F_X(z)\right),
\end{aligned}$$

d'où

$$f_Y(z) = F'_Y(z) = \frac{1}{2\sqrt{2\pi}}\left[e^{-\frac{(-z)^2}{2}} + e^{-\frac{z^2}{2}}\right] = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

donc Y suit aussi la loi $\mathcal{N}(0,1)$.

2. X et Y sont dépendantes par construction et pourtant décorrélées (ce qui n'est pas contradictoire) :

$$\operatorname{cov}(X,Y) = \mathbb{E}(X.BX) - \mathbb{E}(X).\mathbb{E}(BX) = \mathbb{E}(B)\left(\mathbb{E}(X^2) - \mathbb{E}(X)^2\right) = 0.$$

3. On a que :

$$\begin{aligned}
\mathbb{P}(X + Y = 0) &= \mathbb{P}(X(1 + B) = 0) \\
&= \mathbb{P}(X(1 + -1) = 0)\mathbb{P}(B = -1) + \mathbb{P}(X(1 + 1) = 0)\mathbb{P}(B = 1) \\
&= \frac{1}{2}\mathbb{P}(0 = 0) + \frac{1}{2}\mathbb{P}(2X = 0) = \frac{1}{2}
\end{aligned}$$

donc la loi de $X + Y$ n'est pas absolument continue.

Exercice 5.6 : *Coordonnées sphériques*

1. On va utiliser les coordonnées sphériques

$$\begin{cases} x_1 = r\sin\phi\cos\theta \\ x_2 = r\sin\phi\sin\theta \\ x_3 = r\cos\phi \end{cases},$$

et les applications $h : (x_1, x_2, x_3) \mapsto (r, \phi, \theta)$, et son inverse $h^{-1} : (r, \phi, \theta) \mapsto (x_1, x_2, x_3)$. Ici, h^{-1} s'exprime beaucoup plus facilement que h, ce qui nous permet de facilement calculer :

$$\begin{aligned} J_{h^{-1}}(r,\phi,\theta) &= \begin{vmatrix} \dfrac{\partial x_1}{\partial r} & \dfrac{\partial x_1}{\partial \phi} & \dfrac{\partial x_1}{\partial \theta} \\ \dfrac{\partial x_2}{\partial r} & \dfrac{\partial x_2}{\partial \phi} & \dfrac{\partial x_2}{\partial \theta} \\ \dfrac{\partial x_3}{\partial r} & \dfrac{\partial x_3}{\partial \phi} & \dfrac{\partial x_3}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin\phi\cos\theta & r\cos\phi\cos\theta & -r\sin\phi\sin\theta \\ \sin\phi\sin\theta & r\cos\phi\sin\theta & r\sin\phi\cos\theta \\ \cos\phi & -r\sin\phi & 0 \end{vmatrix} \\ &= r^2\cos^2\phi\cos^2\theta\sin\phi + r^2\sin^3\phi\sin^2\theta + r^2\cos^2\phi\sin^2\theta\sin\phi + r^2\sin^3\phi\cos^2\theta \\ &= r^2\cos^2\phi\sin\phi + r^2\sin^3\phi = r^2\sin\phi \end{aligned}$$

On peut remarquer que h est une bijection de $\mathbb{R}^3 \setminus (Oz)$ sur $]0,+\infty[\times]0,\pi[\times]-\pi,\pi]$ sur lequel $J_{h^{-1}} \neq 0$. Par la formule de changement de variables, on peut calculer la densité du vecteur $(R, \Phi, \Theta) = h(X_1, X_2, X_3)$

$$\begin{aligned} f_{h(X_1,X_2,X_3)}(r,\phi,\theta) &= f_{X_1,X_2,X_3}(h^{-1}(r,\phi,\theta)) \times |J_{h^{-1}}(r,\phi,\theta)|\, \mathbb{1}_{]0,+\infty[}(r)\, \mathbb{1}_{]0,\pi[}(\phi)\, \mathbb{1}_{]-\pi,\pi[}(\theta) \\ &= f_{X_1,X_2,X_3}(h^{-1}(r,\phi,\theta))\, r^2\sin\phi\, \mathbb{1}_{]0,+\infty[}(r)\, \mathbb{1}_{]0,\pi[}(\phi)\, \mathbb{1}_{]-\pi,\pi[}(\theta) \end{aligned}$$

avec $f_{X_1,X_2,X_3}(x_1,x_2,x_3) = f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3) = \dfrac{1}{(\sqrt{2\pi})^3}e^{-\frac{1}{2}(x_1^2+x_2^2+x_3^2)}$ par indépendance de $X1, X2, X3$.

Or $x_1^2 + x_2^2 + x_3^2 = r^2$, et donc :

$$f_{R,\Phi,\Theta}(r,\phi,\theta) = \frac{1}{(\sqrt{2\pi})^3}e^{-\frac{r^2}{2}}r^2\, \mathbb{1}_{]0,+\infty[}(r) \times \sin(\phi)\mathbb{1}_{]0,\pi[}(\phi) \times \mathbb{1}_{]-\pi,\pi[}(\theta).$$

Le calcul des lois marginales de R, Φ et Θ est immédiat, du fait de la décomposition de $f_{R,\Phi,\Theta}$ en le produit de trois densités qui dépendent respectivement uniquement de r, ϕ et θ.

2. On a en particulier que $f_R(r) = \displaystyle\int\int f_{R,\Phi,\Theta}(r,\phi,\theta)\,d\phi\,d\theta = \frac{1}{(\sqrt{2\pi})^3}e^{-\frac{r^2}{2}}r^2\, \mathbb{1}_{]0,+\infty[}(r) \times 2\pi\,[-\cos\phi]_0^\pi$, soit

$$f_R(r) = \sqrt{\frac{2}{\pi}}\, r^2\, e^{-\frac{r^2}{2}}\, \mathbb{1}_{]0,+\infty[}(r).$$

3. Calculons maintenant la fonction de répartition de R^2 : $F_{R^2}(u) = \mathbb{P}([R^2 \leqslant u])$. Si $u \leqslant 0$, alors $F_{R^2}(u) = 0$, et si $u > 0$, alors

$$F_{R^2}(u) = \mathbb{P}([-\sqrt{u} \leqslant R \leqslant \sqrt{u}]) = F_R(\sqrt{u}) - F_R(-\sqrt{u}) = F_R(\sqrt{u}).$$

On en déduit donc

$$f_{R^2}(u) = F'_{R^2}(u) = \frac{1}{2\sqrt{u}}f_R(\sqrt{u})\, \mathbb{1}_{]0,+\infty[}(u) = \frac{1}{2}\sqrt{\frac{2}{\pi}} \times \frac{1}{\sqrt{u}} \times ue^{-\frac{u}{2}}\, \mathbb{1}_{]0,+\infty[}(u)$$

soit $f_{R^2}(u) = \dfrac{1}{\sqrt{2\pi}}\sqrt{u}e^{-\frac{u}{2}}\,\mathbb{1}_{]0,+\infty[}(u)$. Donc R^2 sui la loi Gamma $\gamma\left(\dfrac{1}{2},\dfrac{3}{2}\right)$.

Remarque : $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$ donc $\Gamma\left(\dfrac{3}{2}\right) = \Gamma\left(\dfrac{1}{2}+1\right) = \dfrac{1}{2}\Gamma\left(\dfrac{1}{2}\right) = \dfrac{\sqrt{\pi}}{2}$ et

$$\frac{\left(\frac{1}{2}\right)^{3/2}}{\Gamma\left(\frac{3}{2}\right)} = \frac{1}{2\sqrt{2}} \times \frac{2}{\sqrt{\pi}} = \frac{1}{\sqrt{2\pi}}.$$

On a donc bien définit une densité de probabilité.

5 Conditionnement discret - PC6

Énoncés des exercices

Exercice 6.1 : *Trésor caché*
Un présentateur de jeu télévisé cache un prix derrière une porte (il y a 3 portes : A, B et C). Il invite un concurrent à choisir l'une des trois portes, sans l'ouvrir. Il ouvre ensuite l'une des deux portes qui n'a pas été choisie par le concurrent, en sachant que la voiture ne se trouve pas derrière. Il offre alors au concurrent la possibilité de remplacer la porte qu'il avait choisie par l'autre qui est restée fermée.

1. Quelle est la meilleure stratégie pour le joueur : maintenir son choix initial ou opter pour la troisième porte ?

Exercice 6.2 : *Composition d'une famille*
On suppose équiprobable l'ensemble des compositions possibles en garçons et filles d'une famille de deux enfants. Soient : B_1, l'événement "la famille a un garçon", B_2, l'événement "l'aîné est un garçon", et A, l'événement "les 2 enfants sont des garçons".

1. Calculer $\mathbb{P}(A|B_1)$ et $\mathbb{P}(A|B_2)$.

Exercice 6.3 : *Echec/réussite à un examen*
Le taux de réussite d'un examen donné est de 60 % pour les candidats issus de l'établissement A et de 80 % pour ceux issus de B. D'autre part, 55 % des candidats proviennent de A et 45% de B.

1. Quel est le taux d'échec à cet examen ?
2. Quelle est la probabilité pour qu'un candidat éliminé provienne de A ?

Exercice 6.4 : *Recherche de vérité*
Deux personnes mentent parfois : l'une dit la vérité trois fois sur quatre et l'autre quatre fois sur cinq.

1. Lorsqu'elles énoncent la même assertion, quelle est la probabilité pour que cette assertion soit vraie ?

Exercice 6.5 : *Conditionnement discret par la somme*
Les nombres de véhicules X et Y à destination pour les villes A et B, passant par un poste de contrôle commun, sont des v.a. indépendantes de Poisson d'intensités $a > 0$ et $b > 0$.

1. Quelle est la loi conditionnelle de X sachant $(X + Y)$?
2. Calculer l'espérance $\mathbb{E}(X \mid X + Y = n)$ et la variance $\mathrm{Var}(X \mid X + Y = n)$.

Exercice 6.6 : *Toujours en retard*
Pour se rendre à son domicile, un automobiliste passe par la ville 3 fois sur 4. Dans ce cas, il est pris dans un embouteillage 1 fois sur 3 ; quand il contourne la ville, il est retardé par un contrôle de police avec la probabilité 0,2.

1. Calculer la probabilité que l'automobiliste soit retardé
2. Sachant qu'il est en retard, quelle est la probabilité qu'il soit passé par la ville ?

Corrigés des exercices

Exercice 6.1 : *Trésor caché*

1. Notons A_1 l'évènement "le candidat a choisi la bonne porte dès le début". La stratégie qui consiste à maintenir son choix initial revient à calculer la probabilité de l'évènement A_1 (sans prendre en compte d'information supplémentaire) ce qui conduit à une probabilité de succès égale à $\mathbb{P}(A_1) = \dfrac{1}{3}$. Notons maintenant G l'évènement "le candidat gagne la voiture" en choisissant la deuxième stratégie qui consiste à changer de porte après l'indication. D'après la formule des probabilités totales on a que

$$\mathbb{P}(G) = \mathbb{P}(G|A_1)\mathbb{P}(A_1) + \mathbb{P}(G|\overline{A_1})\mathbb{P}(\overline{A_1}).$$

Etant donné que, si le candidat avait choisit la bonne porte initialement, alors il perd s'il change de porte après l'indication, on a que $\mathbb{P}(G|A_1) = 0$. De même, étant donné que, si le candidat avait choisit la mauvaise porte initialement, alors il gagne s'il change de porte après l'indication, on a que $\mathbb{P}(G|\overline{A_1}) = 1$. On a donc finalement que

$$\mathbb{P}(G) = 0 \times \mathbb{P}(A_1) + 1 \times \mathbb{P}(\overline{A_1}) = \mathbb{P}(\overline{A_1}) = \frac{2}{3}.$$

Son meilleur choix est donc de changer de porte !

Exercice 6.2 : *Composition d'une famille*

1. Calculer $\mathbb{P}(A|B_1)$ correspond à la probabilité que les 2 enfants soient des garçons sachant qu'il y a un garçon dans la famille. D'après la définition de la probabilité conditionnelle, on a que

$$\mathbb{P}(A|B_1) = \frac{\mathbb{P}(A \cap B_1)}{\mathbb{P}(B_1)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B_1)} = \frac{1/4}{3/4} = \frac{1}{3},$$

car $A \cap B_1 = A$.

Calculer $\mathbb{P}(A|B_2)$ correspond à la probabilité que les 2 enfants soient des garçons sachant que l'aîné est un garçon. D'après la définition de la probabilité conditionnelle, on a que

$$\mathbb{P}(A|B_2) = \frac{\mathbb{P}(A \cap B_2)}{\mathbb{P}(B_2)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B_2)} = \frac{1/4}{1/2} = \frac{1}{2},$$

car $A \cap B_2 = A$.

Exercice 6.3 : *Echec/réussite à un examen*

Soit les événements A : "être issu de A", B : "être issu de B", R : "réussir" et E : "être en échec".

1. Par la formule des probabilités totales on a que :

$$\mathbb{P}(R) = \mathbb{P}(R|A)\mathbb{P}(A) + \mathbb{P}(R|B)\mathbb{P}(B) = \frac{60 \times 55}{100^2} + \frac{80 \times 45}{100^2} = 69\% : \text{ taux de réussite.}$$

Et donc
$$\mathbb{P}(E) = 1 - \mathbb{P}(R) = 31\% : \text{ taux d'échec.}$$
2. Par définition de la probabilité conditionnelle,
$$\mathbb{P}(A|E) = \frac{\mathbb{P}(A \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|A)\mathbb{P}(A)}{\mathbb{P}(E)} = \frac{(40/100) \times (55/100)}{31/100} \approx 71\%.$$

<u>Exercice 6.4</u> : *Recherche de vérité*

1. Soit E l'événement "les 2 personnes énoncent la même assertion" et A_i l'événement "la i-ème personne ne ment pas", alors $E = (A_1 \cap A_2) \cup (\overline{A}_1 \cap \overline{A}_2)$ et la probabilité cherchée est $\mathbb{P}((A_1 \cap A_2)|E)$. Or $\mathbb{P}((A_1 \cap A_2) \cap E) = \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$ et $\mathbb{P}(E) = \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(\overline{A}_1 \cap \overline{A}_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) + \mathbb{P}(\overline{A}_1)\mathbb{P}(\overline{A}_2)$ car les 2 personnes agissent indépendamment. Avec $\mathbb{P}(A_1) = \frac{3}{4}$ et $\mathbb{P}(A_2) = \frac{4}{5}$, on obtient $\mathbb{P}(A_1 \cap A_2) = \frac{12}{20}$ et $\mathbb{P}(E) = \frac{12}{20} + \frac{1}{20} = \frac{13}{20}$, d'où :
$$\mathbb{P}((A_1 \cap A_2)|E) = \frac{\mathbb{P}((A_1 \cap A_2) \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(E)} = \frac{12}{13} \approx 0,92.$$

<u>Exercice 6.5</u> : *Conditionnement discret par la somme*

1. On a que $X \sim \mathcal{P}(a)$, $Y \sim \mathcal{P}(b)$ et X et Y sont indépendantes. On sait alors que $X + Y \sim \mathcal{P}(a+b)$ que l'on peut remontrer rapidement :
$$\begin{aligned}
\mathbb{P}([X+Y=n]) &= \sum_{k=0}^{n} \mathbb{P}([X=k] \cap [Y=n-k]) \\
&= \sum_{k=0}^{n} \mathbb{P}([X=k])\mathbb{P}([Y=n-k]) \\
&= \sum_{k=0}^{n} e^{-a}\frac{a^k}{k!}e^{-b}\frac{b^{n-k}}{(n-k)!} = \frac{e^{-(a+b)}}{n!}\sum_{k=0}^{n} C_n^k a^k b^{n-k} \\
&= e^{-(a+b)}\frac{(a+b)^n}{n!}.
\end{aligned}$$
Il est également possible de passer par les fonctions génératrices pour obtenir ce résultat, car
$$G_X(z) = \mathbb{E}(z^X) = \sum_{k=0}^{+\infty} z^k e^{-a}\frac{a^k}{k!} = e^{a(z-1)}, \text{ et } G_Y(z) = e^{b(z-1)}, \ z \in \mathbb{R}.$$
Ainsi, on obtient que la fonction génératice de $X+Y$ est (par indépendance)
$$G_{X+Y} = \mathbb{E}(z^{X+Y}) = \mathbb{E}(z^X)\mathbb{E}(z^Y) = e^{a(z-1)}e^{b(z-1)} = e^{(a+b)(z-1)}, \ z \in \mathbb{R},$$
ce qui montre que $X + Y \sim \mathcal{P}(a+b)$.
On a alors
$$\begin{aligned}
\mathbb{P}^{[X+Y=n]}([X=k]) &= \frac{\mathbb{P}([X=k] \cap [X+Y=n])}{\mathbb{P}([X+Y=n])} = \frac{\mathbb{P}([X=k] \cap [Y=n-k])}{\mathbb{P}([X+Y=n])} \\
&= \frac{\mathbb{P}([X=k])\mathbb{P}([Y=n-k])}{\mathbb{P}([X+Y=n])}
\end{aligned}$$

par indépendance de X et de Y. On a alors :
Si $k \notin [\![0, n]\!]$, $\mathbb{P}^{[X+Y=n]}([X = k]) = 0$, sinon :

$$\mathbb{P}^{[X+Y=n]}([X = k]) = \frac{e^{-a}\frac{a^k}{k!}e^{-b}\frac{b^{n-k}}{(n-k)!}}{e^{-(a+b)}\frac{(a+b)^n}{n!}} = C_n^k \frac{a^k b^{n-k}}{(a+b)^n} = C_n^k \left(\frac{a}{a+b}\right)^k \left(\frac{b}{a+b}\right)^{n-k}$$

donc $\mathbb{P}_X^{[X+Y=n]} \sim \mathcal{B}\left(n, \frac{a}{a+b}\right)$ (loi binomiale).

2. On obtient ainsi que

$$\mathbb{E}(X \mid X + Y = n) = n \times \frac{a}{a+b} \quad \text{et} \quad \mathrm{Var}(X \mid X + Y = n) = n \times \frac{a}{a+b} \times \frac{b}{a+b} = n\frac{ab}{(a+b)^2}$$

<u>Exercice 6.6</u> : *Toujours en retard*

Soient les évènements :
- V = “l’automobiliste passe par la ville ”
- R =“ l’automobiliste est retardé ”

A partir des informations données dans l’énoncé on a :

$$\mathbb{P}(V) = \frac{3}{4},\ \mathbb{P}(R|V) = \frac{1}{3},\ \mathbb{P}(R|\overline{V}) = 0.2 = \frac{1}{5}.$$

1. D’après la formule des probabilités totales on a :

$$\mathbb{P}(R) = \mathbb{P}(R|V)\mathbb{P}(V) + \mathbb{P}(R|\overline{V})\mathbb{P}(\overline{V}) = \frac{1}{3} \times \frac{3}{4} + \frac{1}{5} \times \frac{1}{4} = \frac{5}{20} + \frac{1}{20} = \frac{3}{10}.$$

2. On veut calculer $\mathbb{P}(V|R)$. D’après la formule de Bayes on peut écrire :

$$\mathbb{P}(V|R) = \frac{\mathbb{P}(R|V)\mathbb{P}(V)}{\mathbb{P}(R)} = \frac{1/4}{3/10} = \frac{5}{6} \approx 0.83.$$

6 Conditionnement continu - PC7

Énoncés des exercices

Exercice 7.1 : *Loi uniforme sur un carré*
Soit le vecteur aléatoire (X,Y) des coordonnées d'un point placé au hasard (de façon uniformément répartie) sur le support carré de sommets $(-1,0)$, $(0,-1)$, $(1,0)$, $(0,1)$.

1. Déterminer la loi conjointe de (X,Y), les lois marginales de X et Y, et les lois conditionnelles.
2. Les composantes X et Y sont-elles indépendantes ? corrélées ?

Exercice 7.2 : *Conditionnement et indépendance*
Soit (X,Y) un couple de variables aléatoires de densité de probabilité $f_{(X,Y)}(x,y) = e^{-y}$ sur le domaine $D = \{(x,y) : \ x \geqslant 0 \text{ et } y \geqslant x\}$ et qui prend la valeur zéro en dehors de D.

1. Déterminer les lois de X et de Y. X et Y sont-elles indépendantes ?
2. Calculer la densité conditionnelle de Y par rapport à X puis celle de X par rapport à Y.
3. Calculer la densité de probabilité de $Z = Y - X$. Est-ce que les variables aléatoires X et $Z = Y - X$ sont indépendantes ?
4. Calculer la densité de probabilité de $\dfrac{X}{Y}$. Est-ce que les variables aléatoires Y et $\dfrac{X}{Y}$ sont indépendantes ?

Exercice 7.3 : *Somme d'un nombre aléatoire de variables aléatoires*
Soit (X_n) une suite de variables aléatoires indépendantes et identiquement distribuées (iid), de densité f_X et N une variable aléatoire entière positive, indépendante des X_n. Posons

$$S_N = \sum_{i=1}^{N} X_i.$$

1. Déterminer la loi conditionnelle de N sachant que $S_N = x$, puis la loi de S_n où n est fixé.
2. Calculer l'espérance et la variance de S_N.
3. Application : $X \sim \mathcal{E}(\lambda)$ et N est la v.a. géométrique $\mathcal{G}(p)$ (on rappelle que la somme de n v.a. indépendantes $\mathcal{E}(\lambda)$ suit une loi Gamma $\gamma(\lambda, n)$).

Exercice 7.4 : *Loi binomiale de paramètre aléatoire*
Soit la variable aléatoire Y de loi $\mathcal{B}(n,X)$ où le paramètre X est une variable aléatoire de loi uniforme $\mathcal{U}([0,1])$. On a donc que :

$$\mathbb{P}(Y = k | X = x) = \begin{pmatrix} n \\ k \end{pmatrix} x^k (1-x)^{n-k}, \forall k \in \{0,1,2,\ldots,n\}$$

1. Déterminer la loi de Y.
2. Déterminer la loi de X conditionnée par Y.

Corrigés des exercices

Exercice 7.1 : *Loi uniforme sur un carré*

1. Par définition de la loi du couple (X,Y) on a que $f_{(X,Y)}(x,y) = \dfrac{1}{2}\mathbb{1}_D(x,y)$ où D est le carré de sommets $(-1,0)$, $(0,-1)$, $(1,0)$, $(0,1)$ d'aire égale à 2. Ainsi, on obtient que

$$f_X(x) = \int_{-1}^{1} \frac{1}{2}\mathbb{1}_D(x,y)\,dy = (1+x)\mathbb{1}_{[-1,0]}(x) + (1-x)\mathbb{1}_{[0,1]}(x).$$

Par analogie, $f_Y(y) = (1+y)\mathbb{1}_{[-1,0]}(y) + (1-y)\mathbb{1}_{[0,1]}(y)$.
Par définition de la densité conditionnelle, on a que (si $f_Y(y) \neq 0$) $f_{X/Y=y}(x) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}$?
Donc, si par exemple, $y \in [0,1]$, on obtient que

$$f_{X/Y=y}(x) = \frac{1}{2}\frac{\mathbb{1}_D(x,y)}{(1-y)} = \frac{\mathbb{1}_{[y-1\,;\,1-y]}(x)}{2(1-y)}.$$

2. Les composantes ne peuvent pas être indépendantes, puisque le support n'est pas un pavé rectangle. Déterminons la covariance de X et Y :

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

On a que

$$\mathbb{E}(X) = \int_{-1}^{1} x f_X(x)dx = 0,\ \mathbb{E}(Y) = \int_{-1}^{1} y f_Y(y)dy = 0,$$

et

$$\begin{aligned}
\mathbb{E}(XY) &= \int_D xy f_{(X,Y)}(x,y)dxdy \\
&= \frac{1}{2}\int_0^1 \left(\int_{y-1}^{1-y} xdx\right) ydy + \frac{1}{2}\int_{-1}^{0} \left(\int_{-y-1}^{1+y} xdx\right) ydy \\
&= \frac{1}{2}\int_0^1 \left(\int_{y-1}^{1-y} xdx\right) ydy - \frac{1}{2}\int_0^1 \left(\int_{v-1}^{1-v} xdx\right) vdv \text{ (changement de variable } v=-y) \\
&= 0.
\end{aligned}$$

Donc $\mathrm{Cov}(X,Y) = 0$ alors que X et Y ne sont pas indépendantes.

Exercice 7.2 : *Conditionnement et indépendance*

1. La loi de X est donnée par

$$f_X(x) = \int f_{(X,Y)}(x,y)dy = \int_x^{+\infty} e^{-y}dy = e^{-x}\,\mathbb{1}_{]0,+\infty[}(x),$$

donc X suit la loi exponentielle $\mathcal{E}(1)$. La loi de Y est donnée par

$$f_Y(y) = \int f_{(X,Y)}(x,y)dx = \int_0^y e^{-y}dx = ye^{-y}\,\mathbb{1}_{]0,+\infty[}(y)$$

donc Y suit la loi Gamma $\gamma(1,2)$. On a donc $\mathbb{E}(X) = \mathrm{var}(X) = 1$, et $\mathbb{E}(Y) = \mathrm{var}(Y) = 2$. Les v.a. X et Y ne sont pas indépendantes car le support de $f_{(X,Y)}$ n'est pas un rectangle de $\mathbb{R}^2$.

2. La densité conditionnelle de Y par rapport à X est donnée par

$$f_Y^{X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = e^{-(y-x)}\mathbb{1}_{[x,+\infty[}(y), \text{ (loi exponentielle d'origine } x) \ .$$

et la densité conditionnelle de X par rapport à Y est donnée par

$$f_X^{Y=y}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}\mathbb{1}_{[0,y]}(x), \text{ (loi uniforme sur } [0,y]) \ .$$

3. Considérons le changement de variable $h(x,y) = (u,v) = (x, y-x)$ pour $(x,y) \in D$. Pour $(u,v) \in]0,+\infty[\times]0,+\infty[$ on a que $h^{-1}(u,v) = (u, v+u)$ ce qui implique que

$$J_{h^{-1}}(u,v) = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1.$$

Par la formule de changement de variables on a donc que

$$f_{U,V}(u,v) = |J_{h^{-1}}(u,v)| f_{X,Y}(h^{-1}(u,v)) = f_{X,Y}(u, v+u) = e^{-(v+u)}\mathbb{1}_{[0,+\infty[}(u)\mathbb{1}_{[0,+\infty[}(v).$$

On a donc que

$$f_{U,V}(u,v) = f_U(u) f_V(v),$$

où $f_U(u) = e^{-u}\mathbb{1}_{[0,+\infty[}(u)$ est la densité de la v.a. $U = X$, et $f_V(v) = e^{-v}\mathbb{1}_{[0,+\infty[}(v)$ est la densité de la v.a. $V = Y - X$. Donc les v.a. X et $Z = Y - X$ sont indépendantes.

4. Considérons le changement de variables $h(x,y) = (u,v) = \left(\frac{x}{y}, y\right)$ pour $(x,y) \in D$. Pour $0 < u < 1$ et $v > 0$, on a donc que

$$h^{-1}(u,v) = (uv, v)\,,$$

ce qui implique que

$$J_{h^{-1}}(u,v) = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v.$$

Par la formule du changement de variables, on a donc que

$$f_{U,V}(u,v) = |J_{h^{-1}}(u,v)| f_{X,Y}(h^{-1}(u,v)) = ve^{-v}\mathbb{1}_{]0,1[}(u)\mathbb{1}_{]0,+\infty[}(v).$$

Ainsi

$$f_{U,V}(u,v) = f_U(u) f_V(v),$$

où $f_U(u) = \mathbb{1}_{]0,1[}(u)$ est la densité de la v.a. $U = \frac{X}{Y}$ (loi uniforme), et $f_V(v) = ve^{-v}\mathbb{1}_{]0,+\infty[}(v)$ est la densité de la v.a. $V = Y$. Donc les v.a. $\frac{X}{Y}$ et Y sont indépendantes.

<u>**Exercice 7.3**</u> : *Somme d'un nombre aléatoire de variables aléatoires*

Il s'agit du cas d'un couple mixte (S_N, N) où $S_N \in \mathbb{R}$ est une v.a. continue et $N \in \mathbb{N}^*$ une v.a. discrète. La loi jointe du couple (S_N, N) est définie par

$$\text{pour tout } (B, n) \in \mathcal{B}_{\mathbb{R}} \times \mathbb{N}^*,\ \mathbb{P}_{S_N,N}(B, n) = \int_B f_{S_N,N}(x, n)\, dx$$

où $f_{S_N,N} : \mathbb{R} \times \mathbb{N}^* \to \mathbb{R}_+$ est une fonction borélienne positive vérifiant

$$\sum_{n=1}^{+\infty} \int_{\mathbb{R}} f_{S_N,N}(x, n)\, dx = 1.$$

D'après l'énoncé, la densité conditionnelle $f_{S_N|N=n}(x)$ de S_N sachant $N = n$ est la densité de la v.a. $\sum_{i=1}^{n} X_i$. De plus on a la relation

$$f_{S_N|N=n}(x) = \frac{f_{S_N,N}(x, n)}{\mathbb{P}(N = n)}.$$

Les lois marginales sont ici :

$$f_{S_N}(x) = \sum_{n=1}^{+\infty} f_{S_N,N}(x, n) \text{ et } \mathbb{P}(N = n) = \int_{\mathbb{R}} f_{S_N,N}(x, n)\, dx.$$

1. Par définition, la loi conditionnelle de N sachant que $S_N = x$ est

$$\begin{aligned} \mathbb{P}(N = n|S_N = x) &= \frac{f_{S_N,N}(x, n)}{f_{S_N}(x)} = \frac{f_{S_N|N=n}(x)}{f_{S_N}(x)} \mathbb{P}(N = n) \\ &= \frac{f_{S_N|N=n}(x)\mathbb{P}(N = n)}{\sum_{k=1}^{+\infty} f_{S_N|N=k}(x)\mathbb{P}(N = k)} \end{aligned}$$

2. Conditionnellement à l'événement $\{N = n\}$, la loi de S_N est celle de la somme de n v.a. indépendantes. On a donc que

$$\mathbb{E}(S_N|N = n) = n\mathbb{E}(X) \text{ et } \mathrm{Var}(S_N|N = n) = n\mathrm{Var}(X).$$

Ainsi, $\mathbb{E}(S_N|N) = N\mathbb{E}(X)$, et donc par la relation $\mathbb{E}(S) = \mathbb{E}(\mathbb{E}(S_N|N))$ (théorème de l'espérance totale) on obtient que

$$\mathbb{E}(S) = \mathbb{E}(N)\mathbb{E}(X).$$

Par le théorème de la variance totale

$$\mathrm{Var}(S) = \mathbb{E}(\mathrm{Var}(S_N|N)) + \mathrm{Var}(\mathbb{E}(S_N|N)),$$

et donc

$$\mathrm{Var}(S) = \mathbb{E}(N\mathrm{Var}(X)) + \mathrm{Var}(N\mathbb{E}(X)) = \mathbb{E}(N)\mathrm{Var}(X) + (\mathbb{E}(X))^2\mathrm{Var}(N).$$

3. Dans le cas où $X \sim \mathcal{E}(\lambda)$, on a que $f_{S_N|N=n}$ est la convolution de n densités exponentielles $\mathcal{E}(\lambda)$ et est donc égale à la densité de la loi gamma $\gamma(\lambda, n)$ donnée par

$$f_{S_N|N=n}(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)} \mathbb{1}_{[0,+\infty[}(x).$$

On a donc pour tout $x \geqslant 0$

$$
\begin{aligned}
\mathbb{P}(N = n | S_N = x) &= \frac{\frac{\lambda^n x^{n-1} e^{-\lambda x} p(1-p)^{n-1}}{\Gamma(n)}}{\sum_{k=1}^{+\infty} \frac{\lambda^k x^{k-1} e^{-\lambda x} p(1-p)^{k-1}}{\Gamma(k)}} \\
&= \frac{\frac{\lambda^n x^{n-1} (1-p)^{n-1}}{\Gamma(n)}}{\lambda e^{\lambda(1-p)x}} = \frac{(\lambda(1-p)x)^{n-1} e^{-\lambda(1-p)x}}{(n-1)!}
\end{aligned}
$$

La v.a. $N-1$ conditionnée par $\{S_N = x\}$ est poissonnienne $\mathcal{P}(\lambda(1-p)x)$.

Exercice 7.4 : *Loi binomiale de paramètre aléatoire*

Il s'agit du cas d'un couple mixte (X, Y) où $X \in [0,1]$ est une v.a. continue de loi uniforme et $Y \in \{0, 1 \ldots, n\}$ une v.a. discrète. D'après l'énoncé, la loi jointe du couple (X, Y) est définie par

$$
\text{pour tout } (B, k) \in \mathcal{B}_{\mathbb{R}} \times \{0, 1, \ldots, n\},\ \mathbb{P}_{X,Y}(B, k) = \int_B f_{X,Y}(x, k)\, dx
$$

où $f_{X,Y} : \mathbb{R} \times \{0, 1 \ldots, n\} \to \mathbb{R}_+$ est une fonction borélienne positive vérifiant

$$
\sum_{k=0}^{n} \int_{\mathbb{R}} f_{X,Y}(x, k)\, dx = 1
$$

qui est définie par

$$
f_{X,Y}(x, k) = \binom{n}{k} x^k (1-x)^{n-k} \mathbb{1}_{[0,1]}(x).
$$

Les lois marginales sont ici :

$$
f_X(x) = \sum_{k=0}^{n} f_{X,Y}(x, k) \text{ et } \mathbb{P}_Y(\{k\}) = \int_{\mathbb{R}} f_{X,Y}(x, k)\, dx.
$$

1. Le loi marginale de Y est donnée par

$$
\mathbb{P}(Y = k) = \binom{n}{k} \int_0^1 x^k (1-x)^{n-k}\, dx = \frac{n!}{k!(n-k)!} \frac{(n-k)!k!}{(n+1)!} = \frac{1}{n+1}
$$

donc Y est de loi uniforme sur $\{0, 1, \cdots, n\}$. De plus, on retrouve bien que

$$
f_X(x) = \left(\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \right) \mathbb{1}_{[0,1]}(x) = \mathbb{1}_{[0,1]}(x),\ x \in \mathbb{R}
$$

est la densité de la loi uniforme sur $[0,1]$.

2. Par définition de la densité conditionnelle pour un couple mixte continu/discret on obtient que

$$
f_{X/Y=k}(x) = \frac{f_{X,Y}(x, k)}{\mathbb{P}_Y(\{k\})} = \frac{\binom{n}{k} x^k (1-x)^{n-k}}{\left(\frac{1}{n+1}\right)} \mathbb{1}_{[0,1]}(x).
$$

7 Espérance conditionnelle - PC8

Énoncés des exercices

Exercice 8.1 : *Analyse de la variance*
Soient X et Y deux variables aléatoires réelles. On note $\mathbb{E}^X(Y)$ l'espérance conditionnelle de Y sachant X, notée également $\mathbb{E}(Y|X)$. On définit, de même, la variance conditionnelle par

$$\mathrm{Var}^X(Y) = \mathbb{E}^X\left(Y - \mathbb{E}^X(Y)\right)^2.$$

1. En supposant que X et Y sont des variables aléatoires continues (i.e. qui admettent une densité de probabilité), montrer que $\mathbb{E}(Y) = \mathbb{E}\left(\mathbb{E}^X(Y)\right)$.
2. Montrer que $\mathrm{Var}(Y) = \mathbb{E}\left(\mathrm{Var}^X(Y)\right) + \mathrm{Var}(\mathbb{E}^X(Y))$ (formule de l'analyse de la variance).
3. On définit le facteur de corrélation de Y sachant X par $k_{Y|X} = \sqrt{\dfrac{\mathrm{Var}(\mathbb{E}^X(Y))}{\mathrm{Var}(Y)}}$. Interpréter ce coefficient dans les cas limites 0 ou 1.
4. Application : une étude statistique faite pour un magasin a montré que le nombre de clients par semaine est une variable aléatoire de moyenne 400 et d'écart-type 20, et que la dépense de chaque client est une variable aléatoire de moyenne 100 euros et d'écart-type 100 euros. Les dépenses des différents clients sont mutuellement indépendantes, et sont indépendantes du nombre de clients. Calculer la moyenne et l'écart-type du chiffre d'affaire réalisé par le magasin pendant une semaine (le chiffre d'affaire est la recette totale).

Exercice 8.2 : *Choix d'une stratégie optimale*
Soient deux variables aléatoires X et S définies sur $[0,1]$. La variable aléatoire S est caractérisée par sa densité de probabilité : $f_S(s) = 2s$ pour $s \in [0,1]$. D'autre part, la loi de X conditionnelle à $S = s$ est la loi uniforme sur $[0,s]$.

1. Calculer $\mathbb{E}(S|X)$.
2. Soit le jeu suivant : votre adversaire tire " au hasard " une réalisation du vecteur (X,S). Vous choisissez une valeur de a comprise entre 0 et 1. Votre gain G est déterminé par la règle du jeu : $G = a$ si $a + X \leqslant S$ et $G = 0$ si $a + X > S$. Calculer $\mathbb{E}(G|S = s)$ puis $\mathbb{E}(G)$.
3. Déterminer la valeur de a pour laquelle $\mathbb{E}(G)$ est maximum. Déterminez ce maximum.
4. Vous avez maintenant accès à la valeur s que prend S. Quelle valeur de $a = a(s)$ choisirez-vous ?
5. Vous retombez dans l'ignorance de S. Rendu prudent par des pertes de jeu, vous choisissez une stratégie plus sûre et décidez de choisir a pour optimiser $\mathbb{E}(G)$ sous la contrainte $\mathbb{P}(G > 0) \geqslant 0.5$. Quelle valeur de a choisissez-vous ?

Exercice 8.3 : *Estimation bayésienne*
On cherche à estimer la probabilité u de réalisation d'un événement aléatoire en répétant n fois une expérience aléatoire. On note X le nombre de réalisations aléatoires indépendantes observées de cet événement au cours des n expériences. On choisit ensuite une fonction $f : \mathbb{N} \to [0,1]$ de façon à ce que $T = f(X)$ soit un estimateur de u. Le biais d'un estimateur T la quantité

$$b = \mathbb{E}(T) - u.$$

1. Montrer que l'erreur quadratique moyenne $Q(T)$ associée à un estimateur T est la somme de sa variance et du carré de son biais c'est-à-dire :

$$Q(T) = \mathbb{E}\,(T-u)^2 = b^2 + \mathrm{Var}(T).$$

2. Quelle est la loi de X ? Montrer que $T_1 = \dfrac{X}{n}$ est un estimateur sans biais de u. Quelle est sa variance ?

3. L'événement dont on cherche à estimer la probabilité étant peu fréquent, on décide d'adopter un point de vue bayésien en choisissant de considérer que le paramètre u est une variable aléatoire U dont la loi a priori a une densité de probabilité $f_U(u) = 2(1-u)$ sur l'intervalle $[0,1]$.

   Déterminer, à partir de la loi X sachant $U = u$, trouvée à la question 2. et de la loi a priori de U, la loi a posteriori de U c'est à dire la loi conditionnelle de U sachant $X = k$. En déduire que l'estimateur des moindres carrés bayésien est $T_2 = \mathbb{E}(U|X)$.

   Nota :on rappelle l'intégrale suivante

$$\int_0^1 u^j (1-u)^k du = \frac{j!k!}{(j+k+1)!}.$$

4. Montrer que T_2 a en général un biais sauf si $u = 1/3$. Quel est le rapport de cette valeur de u avec la loi a priori de U ? Montrer que T_2 a une variance moins grande que T_1.

5. Application numérique : $u = 0,4$ et $n = 5$. Comparer $Q(T_1)$ et $Q(T_2)$.

Corrigés des exercices

Exercice 8.1 : *Analyse de la variance*

1. Par définition de l'espérance conditionnelle à partir de la loi conditionnelle de Y sachant $X = x$, on a que

$$\mathbb{E}(Y|X=x) = \int_{-\infty}^{+\infty} y f_Y^{X=x}(y)dy = \int_{-\infty}^{+\infty} y \frac{f_{X,Y}(x,y)}{f_X(x)} dy,$$

ce qui implique

$$\mathbb{E}(\mathbb{E}(Y|X)) = \int_{-\infty}^{+\infty} \mathbb{E}(Y|X=x) f_X(x)dx = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} y \frac{f_{X,Y}(x,y)}{f_X(x)} dy \right) f_X(x)dx,$$

et donc en utilisant le théorème de Fubini, on obtient que

$$\mathbb{E}(\mathbb{E}(Y|X)) = \int_{-\infty}^{+\infty} y \left(\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx \right) dy = \int_{-\infty}^{+\infty} y f_Y(y)dy = \mathbb{E}(Y),$$

que l'on peut également écrire $\mathbb{E}\left(\mathbb{E}^X(Y)\right) = \mathbb{E}(Y)$.

2. Par linéarité de l'espérance conditionnelle on obtient que

$$\begin{aligned}
\mathrm{Var}^X(Y) &= \mathbb{E}^X\left(Y^2 - 2Y\mathbb{E}^X(Y) + \left(\mathbb{E}^X(Y)\right)^2\right) \\
&= \mathbb{E}^X\left(Y^2\right) - 2\mathbb{E}^X\left(Y\mathbb{E}^X(Y)\right) + \mathbb{E}^X\left(\left(\mathbb{E}^X(Y)\right)^2\right).
\end{aligned}$$

Or en utilisant la caractérisation de l'espérance conditionnelle $\mathbb{E}(Z|X)$ comme la projection orthogonale de Z sur l'espace des v.a. mesurables par rapport à la tribu engendrée par X (i.e. les v.a. de la forme $h(X)$ où h est une fonction borélienne), on a les relations suivantes $\mathbb{E}^X\left(Y\mathbb{E}^X(Y)\right) = \left(\mathbb{E}^X(Y)\right)^2$ et $\mathbb{E}^X\left(\left(\mathbb{E}^X(Y)\right)^2\right) = \left(\mathbb{E}^X(Y)\right)^2$. Ainsi,

$$\mathrm{Var}^X(Y) = \mathbb{E}^X\left(Y^2\right) - \left(\mathbb{E}^X(Y)\right)^2.$$

Par le théorème de l'espérance totale, on en déduit donc

$$\mathbb{E}\left(\mathrm{Var}^X(Y)\right) = \mathbb{E}\left(\mathbb{E}^X\left(Y^2\right)\right) - \mathbb{E}\left(\left(\mathbb{E}^X(Y)\right)^2\right) = \mathbb{E}(Y^2) - \mathbb{E}\left(\left(\mathbb{E}^X(Y)\right)^2\right).$$

Par ailleurs, par définition de la variance d'une v.a. on a que

$$\mathrm{Var}(\mathbb{E}^X(Y)) = \mathbb{E}\left(\left(\mathbb{E}^X(Y)\right)^2\right) - \left(\mathbb{E}\left(\mathbb{E}^X(Y)\right)\right)^2 = \mathbb{E}\left(\left(\mathbb{E}^X(Y)\right)^2\right) - \left(\mathbb{E}(Y)\right)^2.$$

On obtient donc finalement que

$$\mathbb{E}\left(\mathrm{Var}^X(Y)\right) + \mathrm{Var}(\mathbb{E}^X(Y)) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = \mathrm{Var}(Y).$$

Cette égalité montre que $\mathrm{Var}(\mathbb{E}^X(Y)) \leqslant \mathrm{Var}(Y)$ ce qui illustre la réduction moyenne de la variance par le conditionnement.

3. Soit le facteur de corrélation de Y sachant X défini par $k_{Y|X} = \sqrt{\dfrac{\mathrm{Var}(\mathbb{E}^X(Y))}{\mathrm{Var}(Y)}}$. D'après l'inégalité $\mathrm{Var}(\mathbb{E}^X(Y)) \leqslant \mathrm{Var}(Y)$, on a que $k_{Y|X} \in [0,1]$. De plus, si $k_{Y|X} = 0$ alors $\mathrm{Var}(\mathbb{E}^X(Y)) =$

0 et donc $\mathbb{E}^X(Y)$ est une constante. Si $k_{Y|X} = 1$ alors d'après la formule de la variance totale $\mathbb{E}\left(\mathrm{Var}^X(Y)\right) = 0$ ce qui implique que $\mathrm{Var}^X(Y) = 0$ car $\mathrm{Var}^X(Y)$ est une v.a. positive. Dans ce dernier cas, la loi conditionnelle de Y sachant X est donc celle d'une constante.

4. Soient les variables aléatoires : N, "nombre de clients par semaine", et D, "dépense d'un client". On notera D_i les v.a. indépendantes de même loi que D et qui représentent la dépense pour chaque client (D_i = "dépense du i-ème client"). D'après l'énoncé, on a que

$$\mathbb{E}(N) = 400,\ \sqrt{\mathrm{Var}(N)} = 20,\ \mathbb{E}(D) = 100,\ \sqrt{\mathrm{Var}(D)} = 100.$$

Le chiffre d'affaire hebdomadaire est donné par la v.a.

$$C = \sum_{i=1}^{N} D_i.$$

On a que $\mathbb{E}(C|N=n) = \mathbb{E}\left(\sum_{i=1}^{n} D_i\right) = n\mathbb{E}(D)$. Donc $\mathbb{E}(C|N) = N\mathbb{E}(D)$ et ainsi on obtient que $\mathbb{E}(C) = \mathbb{E}(N)\mathbb{E}(D) = 40\,000$ euros.
Par la formule de la variance totale

$$\mathrm{Var}(C) = \mathbb{E}(\mathrm{Var}(C|N)) + \mathrm{Var}(\mathbb{E}(C|N)).$$

Or,

$$\mathrm{Var}(C|N=n) = \mathrm{Var}\left(\sum_{i=1}^{n} D_i\right) = n\mathrm{Var}(D),$$

et donc $\mathrm{Var}(C|N) = N\mathrm{Var}(D)$. De plus, $\mathrm{Var}(\mathbb{E}(C|N)) = \mathrm{Var}(NE(D)) = (\mathbb{E}(D))^2\mathrm{Var}(N)$. Ainsi

$$\mathrm{Var}(C) = \mathbb{E}(N)\mathrm{Var}(D) + (\mathbb{E}(D))^2\mathrm{Var}(N) = 8 \times 10^6,$$

d'où un écart type de chiffre d'affaire égal à $2\sqrt{2}10^3 \approx 3\,000$ euros.

Exercice 8.2 : *Choix d'une stratégie optimale*

D'après l'énoncé, on a que

$$f_S(s) = 2s\mathbb{1}_{[0,1]}(s),\ f_X^{S=s}(x) = \frac{1}{s}\mathbb{1}_{[0,s]}(x).$$

1. La loi conditionnelle de S sachant $X = x$ est donnée par

$$f_S^{X=x}(s) = \frac{f_{(X,S)}(x,s)}{f_X(x)} = \frac{f_X^{S=s}(x)f_S(s)}{f_X(x)}.$$

On a que

$$f_X(x) = \int_{\mathbb{R}} f_{(X,S)}(x,s)ds = \int_{\mathbb{R}} f_X^{S=s}(x)f_S(s)ds = \int_{\mathbb{R}} 2\mathbb{1}_{[0,s]}(x)\mathbb{1}_{[0,1]}(s) = \int_x^1 2ds = 2(1-x)\mathbb{1}_{[0,1]}(x).$$

Donc, pour $x \in [0,1]$,

$$f_S^{X=x}(s) = \frac{2\mathbb{1}_{[0,s]}(x)\mathbb{1}_{[0,1]}(s)}{2(1-x)\mathbb{1}_{[0,1]}(x)} = \frac{1}{1-x}\mathbb{1}_{[x,1]}(s).$$

Ainsi,

$$\mathbb{E}(S|X=x) = \int_{\mathbb{R}} s f_S^{X=x}(s)ds = \frac{1}{1-x}\int_x^1 s ds = \frac{1-x^2}{2(1-x)} = \frac{1+x}{2}.$$

D'où $\mathbb{E}(S|X) = \dfrac{1+X}{2}$.

2. La v.a. G est discrète et prend deux valeurs a (avec probabilité $\mathbb{P}(a+X \leqslant S)$) ou 0 (avec probabilité $\mathbb{P}(a+X > S)$). Ainsi,

$$\mathbb{E}(G|S=s) = a\times\mathbb{P}(a+X \leqslant S|S=s)+0\times\mathbb{P}(a+X > S|S=s) = a\times\mathbb{P}(X \leqslant s-a|S=s) = aF_X^{S=s}(s-a).$$

On a donc que $\mathbb{E}(G|S=s) = a\dfrac{s-a}{s}$ si $s \geqslant a$ et $\mathbb{E}(G|S=s) = 0$ si $s < a$. Ainsi,

$$\mathbb{E}(G|S) = a\frac{S-a}{S}\mathbb{1}_{S\geqslant a}.$$

On obtient alors que

$$\mathbb{E}(G) = \mathbb{E}(\mathbb{E}(G|S)) = a\int_{\mathbb{R}}\left(\frac{s-a}{s}\mathbb{1}_{s\geqslant a}\right)f_S(s)ds = a\int_a^1 2(s-a)ds = a^3 - 2a^2 + a.$$

3. Il s'agit de déterminer le maximum de la fonction $a \mapsto a^3 - 2a^2 + a$ pour $a \in [0,1]$ donc la dérivée est $3a^2 - 4a + 1$ qui s'annule en $a = 1/3$ et $a = 1$. Le maximum de $\mathbb{E}(G)$ est donc $1/27 - 2/9 + 1/3 \approx 0.148$ obtenu pour $a = 1/3$.

4. On sait que $\mathbb{E}(G|S=s) = a\dfrac{s-a}{s}$ si $s \geqslant a$ et 0 sinon. La fonction $a \mapsto a\dfrac{s-a}{s}$ est maximum pour $a(s) = \dfrac{s}{2}$ et dans ce cas $\mathbb{E}(G|S=s) = \dfrac{s}{4}$. La stratégie qui consiste à prendre $a(S) = \dfrac{S}{2}$ conduit alors à un gain moyen $\mathbb{E}(G) = \dfrac{\mathbb{E}(S)}{4} = \dfrac{1}{6} \approx 0.1666$ car $\mathbb{E}(S) = \displaystyle\int_0^1 2s^2 ds = \frac{2}{3}$.

L'information sur S fait gagner en moyenne 0.02 sur un gain de l'ordre de 0.15 soit une amélioration d'environ 12%.

5. On veut s'assurer que $\mathbb{P}(G > 0) \geqslant 0.5$. Remarquons que

$$\begin{aligned}
\mathbb{P}(G > 0) &= \mathbb{P}(X + a \leqslant S) = \mathbb{P}(X \leqslant S - a) = \int_{\mathbb{R}} \mathbb{P}(X \leqslant s - a|S = s)f_S(s)ds \\
&= \int_{\mathbb{R}} \frac{s-a}{s}\mathbb{1}_{s\geqslant a}2s\mathbb{1}_{[0,1]}(s)ds = \int_a^1 2(s-a)ds = (1-a)^2.
\end{aligned}$$

Donc $\mathbb{P}(G > 0) \geqslant 0.5$ est équivalent à la condition $a \leqslant 1 - \sqrt{0.5} \approx 0.293$. Or pour $a = 1 - \sqrt{0.5}$ on peut remarquer que $\mathbb{E}(G) \approx 0.146$ qui est très proche de la valeur maximale de $\mathbb{E}(G)$ égale à 0.148. Si l'on voulait assurer que $\mathbb{P}(G > 0) \geqslant 0.9$ on obtiendrait la condition $a \leqslant 0.0513$ et pour $a = 0.0513$ on aurait que $\mathbb{E}(G) = 0.046$.

Exercice 8.3 : *Estimation bayésienne*

1. Par linéarité de l'espérance

$$Q(T) = \mathbb{E}\left(T^2 - 2Tu + u^2\right) = \mathbb{E}(T^2) - 2u\mathbb{E}(T) + u^2.$$

Or $b^2 = (\mathbb{E}(T) – u)^2 = (\mathbb{E}(T))^2 - 2u\mathbb{E}(T) + u^2$, donc

$$Q(T) = \mathbb{E}(T^2) + b^2 - (\mathbb{E}(T))^2 = b^2 + \mathrm{Var}(T).$$

2. Il est clair que X suit une loi binomiale $\mathcal{B}(n,u)$. Donc $\mathbb{E}(X) = nu$ ce qui implique que $\mathbb{E}(T) = \mathbb{E}\left(\frac{X}{n}\right) = u$ est un estimateur sans biais de u. Sa variance est

$$\mathrm{Var}(T) = \frac{\mathrm{Var}(X)}{n^2} = \frac{nu(1-u)}{n^2} = \frac{u(1-u)}{n}.$$

3. D'après la question précédente, $\mathbb{P}(X = k|U = u) = C_n^k u^k (1-u)^{n-k}$ et on a que $f_U(u) = 2(1-u)\mathbb{1}_{[0,1]}(u)$. D'après la formule de Bayes Par conséquent

$$\begin{aligned} f_U^{X=k}(u) &= \frac{\mathbb{P}(X=k|U=u) f_U(u)}{\mathbb{P}(X=k)} = \frac{2C_n^k u^k (1-u)^{n-k+1}}{\int_0^1 2C_n^k t^k (1-t)^{n-k+1} dt} \\ &= \frac{u^k(1-u)^{n-k+1}}{\int_0^1 t^k (1-t)^{n-k+1} dt} = \frac{(n+2)!}{k!(n-k+1)!} u^k (1-u)^{n-k+1} \mathbb{1}_{[0,1]}(u), \end{aligned}$$

et on obtient alors que

$$\mathbb{E}(U|X=k) = \frac{(n+2)!}{k!(n-k+1)!} \int_0^1 u^{k+1}(1-u)^{n-k+1} = \frac{(n+2)!}{k!(n-k+1)!} \frac{(k+1)!(n-k+1)!}{(n+3)!} = \frac{k+1}{n+3}.$$

Ce qui implique que

$$T_2 = \mathbb{E}(U|X) = \frac{X+1}{n+3}.$$

4. On a alors que

$$\mathbb{E}(T_2) = \mathbb{E}(\mathbb{E}(U|X)) = \mathbb{E}(U) = \int_0^1 2u(1-u)du = \frac{1}{3}.$$

Si on suppose que la vraie valeur de la probabilité à estimer est u, on a en général que $\mathbb{E}(T_2) \neq u$, sauf si $u = \frac{1}{3}$ qui est la valeur de l'espérance de la loi a priori. On a ensuite que (en utilisant de le résultat de la question 2)

$$\mathrm{Var}(T_2) = \mathrm{Var}\left(\frac{X+1}{n+3}\right) = \frac{\mathrm{Var}(X)}{(n+3)^2} < \frac{\mathrm{Var}(X)}{n^2} = \mathrm{Var}(T).$$

5. Pour $u = 0.4$ et $n = 5$ on obtient que

$$Q(T_1) = \mathrm{Var}(T_1) = \frac{0.4 \times 0.6}{5} = 0.048,$$

et

$$Q(T_2) = \mathrm{Var}(T_2) + (\mathbb{E}(T_2) - u)^2 = \frac{5 \times 0.4 \times 0.6}{64} + (1/3 - 0.4)^2 = 0.023.$$

8 Vecteurs gaussiens - PC9

<u>**Énoncés des exercices**</u>

<u>**Exercice 9.1**</u> : *Espérance conditionnelle*
Soit le vecteur gaussien centré $(X_1, X_2, ..., X_n)$ de matrice de covariance Γ. Le but de cet exercice est de déterminer $\mathbb{E}(X_n\ |X_1, X_2, ..., X_{n-1})$.

1. Montrer que
$$\mathbb{E}(X_n X_i) = \mathbb{E}(\mathbb{E}(X_n|X_1, X_2, \cdots, X_{n-1})X_i)$$
pour tout $i = 1,....,n-1$.
2. En déduire un système linéaire de $n-1$ équations dont la solution permet d'exprimer $\mathbb{E}(X_n|X_1, X_2, ..., X_{n-1})$ comme une combinaison linéaire de $(X_1, X_2, ..., X_{n-1})$.
3. Appliquer cette méthode au vecteur (X_1, X_2, X_3) pour lequel
$$\Gamma = \begin{pmatrix} 3 & 6 & 2 \\ 6 & 14 & 3 \\ 2 & 3 & 6 \end{pmatrix}.$$
4. En déduire l'expression de l'espérance conditionnelle de X_3 sachant (X_1, X_2).

<u>**Exercice 9.2**</u> : *Théorème de Cochran*
Le résultat à démontrer dans cet exercice, connu sous le nom de "Théorème de Cochran" est essentiel dans toute la théorie des modèles gaussiens. Il intervient dans la plupart des problèmes d'estimation de paramètres issus d'une loi gaussienne.

Soit X un vecteur gaussien centré de dimension n et de matrice de covariance $\sigma^2 I_n$. Soit $E_1, \ldots, E_q$ une décomposition de $\mathbb{R}^n$ en sous-espaces vectoriels orthogonaux de dimension $p_1, \ldots, p_q$, i.e. $\mathbb{R}^n = \bigoplus_{k=1}^{q} E_k$. Soit $\Pi_k : \mathbb{R}^n \to \mathbb{R}^n$ la matrice de projection orthogonale sur l'espace E_k pour $k = 1, \ldots, q$.

1. Montrer que les vecteurs $W_k = \Pi_k X, k = 1, \ldots, q$ sont des vecteurs gaussiens indépendants.
2. Montrer que $\frac{1}{\sigma^2}\|W_k\|^2$ suit une loi du chi2 à p_k degrés de libertés pour $k = 1, \ldots, q$, où $\|\cdot\|$ est la norme euclidienne standard dans $\mathbb{R}^n$.

<u>Rappel</u> : si $Z_1, \ldots, Z_p$ sont des variables aléatoires iid de loi $N(0,1)$, alors (par définition) la variable aléatoire $W = Z_1^2 + \ldots Z_p^2$ suit une loi du chi2 à p degrés de libertés, notée $\chi^2(p)$.

<u>**Exercice 9.3**</u> : *Coordonnées polaires*

Soient X, Y deux variables aléatoires réelles indépendantes de densité de probabilité f_X et f_Y. On considère le changement de variables en coordonnées polaires $X = R\cos(\Theta)$ et $Y = R\sin(\Theta)$ avec $R \in [0, +\infty[$ et $\Theta \in [0, 2\pi]$.

1. Calculer la densité du rayon R en fonction de f_X et f_Y.
2. Dans le cas où X et Y suivent des lois gaussiennes centrées réduites, déterminer la densité de Θ. Les variables R et Θ sont-elles indépendantes ?

3. Si $X_1, \ldots, X_n$ sont des variables aléatoires iid de loi $N(0,1)$, quelle est la densité de la variable aléatoire $R = \sqrt{X_1^2 + \ldots X_n^2}$?

Corrigés des exercices

Exercice 9.1 : *Espérance conditionnelle*

1. A l'aide des propriétés de l'espérance conditionnelle on peut remarquer que

$$\mathbb{E}(X_n X_i) = \mathbb{E}(\mathbb{E}(X_n X_i | X_1, X_2, \cdots, X_{n-1})) = \mathbb{E}(\mathbb{E}(X_n | X_1, X_2, \cdots, X_{n-1}) X_i), \; i = 1, \ldots, n. \tag{1}$$

2. D'après les théorèmes de conditionnement de vecteurs gaussiens, et étant donné que $X_1, X_2, ..., X_n$ sont des v.a. d'espérance égale à zéro, on a que

$$\mathbb{E}(X_n | X_1, X_2, \cdots, X_{n-1}) = \sum_{j=1}^{n-1} \alpha_j X_j, \tag{2}$$

où les $\alpha_1, \ldots, \alpha_{n-1}$ sont des réels. Il s'agit donc de déterminer les coefficients α_j. Pour cela, on peut multiplier les deux membres de l'égalité (2) par X_i pour $1 \leqslant i \leqslant n-1$

$$\mathbb{E}(X_n | X_1, X_2, \cdots, X_{n-1}) X_i = \sum_{j=1}^{n-1} \alpha_j X_j X_i$$

ce qui conduit à la relation

$$\mathbb{E}(\mathbb{E}(X_n | X_1, X_2, \cdots, X_{n-1}) X_i) = \sum_{j=1}^{n-1} \alpha_j \mathbb{E}(X_j X_i).$$

Ainsi, en utilisant l'égalité (1), si l'on fait varier i dans $\{1, 2, \cdots, n-1\}$, on parvient à un système de $(n-1)$ équations linéaires en les $(\alpha_j)_{1 \leqslant j \leqslant n-1}$:

$$\left(\sum_{j=1}^{n-1} \alpha_j \mathbb{E}(X_i X_j) = \mathbb{E}(X_n X_i) \right)_{1 \leqslant i \leqslant n-1} \tag{3}$$

3. Le vecteur (X_1, X_2, X_3) étant centré, Γ est la matrice $(\mathbb{E}(X_i X_j))_{1 \leqslant i,j \leqslant 3}$. Le système d'équation (3) s'écrit alors

$$\begin{cases} 3\alpha_1 + 6\alpha_2 &= 2 \\ 6\alpha_1 + 14\alpha_2 &= 3 \end{cases}$$

4. La solution de ce système linéaire est $\alpha_1 = \dfrac{5}{3}$ et $\alpha_2 = -\dfrac{1}{2}$, d'où

$$\mathbb{E}(X_3 | X_1, X_2) = \frac{5}{3} X_1 - \frac{1}{2} X_2.$$

Exercice 9.2 : *Théorème de Cochran*

1. Soit $1 \leqslant k \leqslant q$. Le vecteur $W_k = \Pi_k X$ de $\mathbb{R}^n$ est obtenu par une transformation linéaire du vecteur gaussien $X \sim \mathcal{N}(0, \sigma^2 I_n)$. Le loi de W_k est donc celle d'un vecteur gaussien de $\mathbb{R}^n$ d'espérance

$$\mathbb{E}(W_k) = \mathbb{E}(\Pi_k X) = \Pi_k \mathbb{E}(X) = 0,$$

et de matrice de covariance

$$\mathbb{E}((W_k - \mathbb{E}(W_k))(W_k - \mathbb{E}(W_k))^t) = \mathbb{E}(\Pi_k X X^t \Pi_k^t) = \Pi_k \mathbb{E}(XX^t)\Pi_k^t = \sigma^2 \Pi_k,$$

car $\mathbb{E}(XX^t) = \sigma^2 I_n$ et $\Pi_k \Pi_k^t = \Pi_k$.

Soit $1 \leqslant k_1, k2 \leqslant q$ avec $k_1 \neq k_2$. Etant donné que W_{k_1} et W_{k_2} sont des vecteurs gaussiens d'espérance nulle, il suffit de montrer que la matrice de covariances entre leurs coordonnées est nulle. On peut remarquer que

$$\mathbb{E}(W_{k_1} W_{k_2}^t) = \Pi_{k_1} \mathbb{E}(XX^t)\Pi_{k_2}^t = \sigma^2 \Pi_{k_1} \Pi_{k_2}^t = 0,$$

car Π_{k_1} et Π_{k_2} sont des matrices de projection orthogonales sur les sous-espaces vectoriels E_{k_1} et E_{k_2} qui sont orthogonaux. Ainsi, les coordonnées de W_{k_1} et celles de W_{k_2} sont indépendantes deux à deux (car, dans le cas Gaussien, indépendance et covariance nulle sont des propriétés équivalentes) ce qui montre que W_{k_1} et W_{k_2} sont des vecteurs aléatoires indépendants.

2. Soit $1 \leqslant k \leqslant q$ et $e_1, \ldots, e_{p_k}$ une base orthonormée de E_k. En notant $\langle \cdot, \cdot \rangle$ le produit scalaire usuel de $\mathbb{R}^n$ on a que

$$\|W_k\|^2 = \sum_{j=1}^{p_k} Z_j^2,$$

avec $Z_j = \langle X, e_j \rangle = X^t e_j,\ j = 1, \ldots, p_k$. La v.a. réelle Z_j étant obtenue par une transformation linéaire du vecteur X, on a que Z_j est une v.a. de loi normale, d'espérance $\mathbb{E} Z_j = \mathbb{E}(\langle X, e_j \rangle) = \langle \mathbb{E}X, e_j \rangle = 0$ et de variance $\mathrm{Var}(Z_j) = e_j^t \mathbb{E}(XX^t) e_j = \sigma^2 \|e_j\|^2 = \sigma^2$ pour $j = 1, \ldots, p_k$. Etant donné que $\mathbb{E}(Z_j Z_{j'}) = e_j^t \mathbb{E}(XX^t) e_{j'} = \sigma^2 \langle e_j, e_{j'} \rangle = 0$ pour $j \neq j'$, on a donc que les v.a. gaussiennes (et centrées) $Z_1, \ldots, Z_{p_k}$ sont des v.a. indépendantes deux à deux et de même loi $\mathcal{N}(0, \sigma^2)$, ce qui implique (par définition) que $\frac{1}{\sigma^2} = \sum_{j=1}^{p_k} \left(\frac{Z_j}{\sigma} \right)^2$ suit une loi du chi2 à p_k degrés de libertés.

Exercice 9.3 : *Coordonnées polaires*

Soient X, Y deux variables aléatoires réelles indépendantes de densité de probabilité f_X et f_Y. On considère le changement de variables en coordonnées polaires $X = R\cos(\Theta)$ et $Y = R\sin(\Theta)$.

1. Considérons le changement de variable en coordonnées polaires $h : (x, y) \mapsto (r, \theta)$ comme la bijection de $\mathbb{R}^2 \setminus \{(0,0)\}$ sur $]0, +\infty[\times[0, 2\pi[$, de bijection réciproque $h^{-1} : (r, \theta) \mapsto (x, y) = (r\cos\theta, r\sin\theta)$ dont le jacobien est donné par

$$J_{h^{-1}}(r, \theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

Par la formule de changement de variables et en utilisant l'indépendance de X et Y, on a donc que

$$\begin{aligned} f_{R,\Theta}(r, \theta) &= f_{X,Y}(r\cos\theta, r\sin\theta) \times r\, \mathbb{1}_{]0,+\infty[}(r)\, \mathbb{1}_{]0,2\pi[}(\theta) \\ &= f_X(r\cos\theta) f_Y(r\sin\theta) \times r\, \mathbb{1}_{]0,+\infty[}(r)\, \mathbb{1}_{]0,2\pi[}(\theta) \end{aligned}$$

et donc
$$f_R(r) = r\int_0^{2\pi} f_X(r\cos\theta)f_Y(r\sin\theta)d\theta\ \mathbb{1}_{]0,+\infty[}(r).$$

2. Dans le cas où X et Y suivent des lois gaussiennes centrées réduites, on obtient alors que

$$\begin{aligned}
f_{R,\Theta}(r,\theta) &= \frac{1}{2\pi}re^{-\frac{r^2}{2}}\ \mathbb{1}_{]0,+\infty[}(r)\ \mathbb{1}_{]0,2\pi[}(\theta),\\
f_R(r) &= r\left(\int_0^{2\pi}\frac{1}{2\pi}e^{-\frac{r^2}{2}}d\theta\right)\ \mathbb{1}_{]0,+\infty[}(r) = re^{-\frac{r^2}{2}}\ \mathbb{1}_{]0,+\infty[}(r)\\
f_\Theta(\theta) &= \frac{1}{2\pi}\left(\int_0^{+\infty} re^{-\frac{r^2}{2}}\right)\ \mathbb{1}_{]0,2\pi[}(\theta) = \frac{1}{2\pi}\ \mathbb{1}_{]0,2\pi[}(\theta).
\end{aligned}$$

Ainsi, on obtient que $f_{R,\Theta}(r,\theta) = f_R(r)f_\Theta(\theta)$ et donc les v.a. R et Θ sont indépendantes.

3. En dimension supérieure $n \geqslant 3$, on obtient que la densité de $R = \sqrt{X_1^2 + \ldots X_n^2}$, dans le cas où $X_1, \ldots, X_n$ sont des variables aléatoires iid de loi $N(0,1)$, est donnée par

$$f_R(r) = \frac{r^{n-1}e^{-\frac{r^2}{2}}}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}\ \mathbb{1}_{]0,+\infty[}(r)$$

avec $\Gamma(z) = \displaystyle\int_0^{+\infty} t^{z-1}e^{-t}dt$.

9 Fonction caractéristique - PC10

Énoncés des exercices

Exercice 10.1 : *Fonction caractéristique*
Soient X et Y deux variables aléatoires réelles indépendantes de densité de probabilités respective f_X et f_Y. On appelle fonction caractéristique de X, la fonction Φ_X définie par :

$$\forall t \in \mathbb{R},\ \Phi_X(t) = \mathbb{E}(e^{itX}) \text{ (espérance mathématique de la variable aléatoire } e^{itX}).$$

La fonction caractéristique d'une variable aléatoire permet de complètement spécifier sa loi de probabilité.

1. Déterminer la fonction caractéristique de $U = aX + b$ où a et b sont deux réels.
2. Montrer que $\mathbb{E}(X) = -i\Phi'_X(0)$ et $\mathbb{E}(X^2) = -\Phi''_X(0)$.
3. Déterminer la fonction caractéristique de $S = X + Y$.

Exercice 10.2 : *Fonction caractéristique d'une loi uniforme*
Soit X une variable aléatoire de loi uniforme sur $[-1, 1]$.

1. Calculer la fonction caractéristique Φ_X de X.
2. Calculer $\mathbb{E}(X)$ et $\mathbb{E}(X^2)$ à l'aide de Φ_X

Exercice 10.3 : *Fonction caractéristique d'une loi normale*
Soient X et Y deux variables aléatoires indépendantes de loi normale centrée réduite $\mathcal{N}(0, 1)$.

1. Calculer la fonction caractéristique Φ_X de X.
2. Déterminer la loi de la variable aléatoire $S = X + Y$ à l'aide de sa fonction caractéristique.
3. Généraliser aux cas où X et Y suivent des lois normales de moyenne m et d'écart-type σ.

Exercice 10.4 : *Lois gaussiennes*
Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes gaussiennes de loi $\mathcal{N}(m, \sigma^2)$. On pose

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ U_i = X_i - \overline{X},\ \text{et } S^2 = \frac{1}{n}\sum_{i=1}^{n} U_i^2.$$

1. En utilisant la fonction caractéristique déterminer la loi de $\overline{X}$. En déduire son espérance et sa variance.
2. En utilisant la fonction caractéristique déterminer la loi de U_i. En déduire son espérance et sa variance.
3. Calculer $\mathrm{Cov}(U_i, U_j)$, puis $\mathrm{Cov}(U_i, \overline{X})$.
4. Déterminer $\mathbb{E}(S^2)$.
5. Montrer que $nS^2 = \displaystyle\sum_{i=1}^{n} X_i^2 - n\overline{X}^2$ et donc que $n\left(S^2 + (\overline{X} - m)^2\right) = \displaystyle\sum_{i=1}^{n}(X_i - m)^2$.
6. En déduire, en utilisant les fonctions caractéristiques, la loi de S^2.

Corrigés des exercices

Exercice 10.1 : *Fonction caractéristique*

1. En utilisant la définition de la fonction caractéristique et la linéarité de l'espérance, on obtient que

$$\forall t \in \mathbb{R},\ \Phi_U(t) = \mathbb{E}(e^{it(aX+b)}) = e^{itb}\mathbb{E}(e^{itaX}) = e^{itb}\Phi_X(at).$$

2. Par définition,

$$\forall t \in \mathbb{R},\ \Phi_X(t) = \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx} f_X(x)dx,$$

donc, par dérivation sous le signe somme (dans l'hypothèse où $\int_{\mathbb{R}} |x| f_X(x)dx < +\infty$), on a que

$$\forall t \in \mathbb{R},\ \Phi'_X(t) = \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} ixe^{itx} f_X(x)dx,$$

et donc $\Phi'_X(0) = i\int_{\mathbb{R}} x f_X(x)dx = i\mathbb{E}(X)$, d'où la relation : $\mathbb{E}(X) = -i\Phi'_X(0)$. De même (dans l'hypothèse où $\int_{\mathbb{R}} |x|^2 f_X(x)dx < +\infty$), on a que

$$\forall t \in \mathbb{R},\ \Phi''_X(t) = \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} -x^2 e^{itx} f_X(x)dx,$$

et donc $\Phi''_X(0) = -\int_{\mathbb{R}} x^2 f_X(x)dx = -\mathbb{E}(X^2)$, d'où la relation : $\mathbb{E}(X^2) = -\Phi''_X(0)$.

3. En utilisant la définition de la fonction caractéristique et l'indépendance de X et Y on obtient que

$$\forall t \in \mathbb{R},\ \Phi_S(t) = \mathbb{E}(e^{it(X+Y)}) = \mathbb{E}(e^{itX}e^{itY}) = \mathbb{E}(e^{itX})\mathbb{E}(e^{itY}) = \Phi_X(t)\Phi_Y(t).$$

Exercice 10.2 : *Fonction caractéristique d'une loi uniforme*

1. Par définition de la fonction caractéristique et étant donné que $f_X(x) = \frac{1}{2}\mathbb{1}_{[-1,1]}(x)$ (densité de la loi uniforme sur $[-1,1]$), on a que

$$\forall t \in \mathbb{R},\ \Phi_X(t) = \mathbb{E}(e^{itX}) = \frac{1}{2}\int_{-1}^{1} e^{itx}dx = \frac{2i\sin(t)}{2it} = \frac{\sin(t)}{t}.$$

2. Au voisinage de $t = 0$, considérons le développement limité de la fonction $t \mapsto \sin(t)$

$$\sin(t) = t - \frac{t^3}{3!} + \frac{t^5}{5!} + \ldots,$$

qui implique qu'au voisinage de $t = 0$,

$$\frac{\sin(t)}{t} = 1 - \frac{t^2}{3!} + \frac{t^4}{5!} + \ldots$$

Etant donné que $\mathbb{E}(X) = -i\Phi'_X(0) = -i\left(\frac{\sin(t)}{t}\right)'(0)$, on obtient donc que

$$\mathbb{E}(X) = 0.$$

Au voisinage de $t = 0$,

$$\left(\frac{\sin(t)}{t}\right)'(t) = -2\frac{t}{3!} + 4\frac{t^3}{5!} + \dots$$

et donc

$$\left(\frac{\sin(t)}{t}\right)''(t) = -\frac{1}{3} + 4 \times 3\frac{t^2}{5!} + \dots$$

Etant donné que $\mathbb{E}(X^2) = -\Phi_X''(0) = -\left(\frac{\sin(t)}{t}\right)''(0)$, on obtient donc que

$$\mathbb{E}(X^2) = \frac{1}{3}.$$

<u>Exercice 10.3</u> : *Fonction caractéristique d'une loi normale*

1. On a que X est une v.a. de densité $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Donc, par définition de la fonction caractéristique,

$$\forall t \in \mathbb{R},\ \Phi_X(t) = \mathbb{E}(e^{itX}) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{itx}e^{-\frac{x^2}{2}}dx.$$

Calculons $\Phi_X'(t)$ par dérivation sous la signe somme.

$$\begin{aligned}
\Phi_X'(t) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} ixe^{itx-\frac{x^2}{2}}dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} (-i(-x+it)-t)e^{itx-\frac{x^2}{2}}dx, \text{ en remarquant que } ix = -i(-x+it)-t, \\
&= \frac{-i}{\sqrt{2\pi}}\int_{\mathbb{R}} (-x+it)e^{itx-\frac{x^2}{2}}dx - \frac{t}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{itx-\frac{x^2}{2}}dx \\
&= \frac{-i}{\sqrt{2\pi}}\left[e^{itx-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} - t\Phi_X(t) = -t\Phi_X(t),
\end{aligned}$$

car $\left[e^{itx-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} = 0$. On obtient donc l'équation différentielle suivante

$$\Phi_X'(t) + t\Phi_X(t) = 0,$$

dont la solution est de la forme $\Phi_X(t) = ke^{-\frac{t^2}{2}}$ où k est une constante. En utilisant la relation $\Phi_X(0) = \int_{\mathbb{R}} f_X(x)dx = 1$, on obtient donc que $k = 1$ et que

$$\Phi_X(t) = e^{-\frac{t^2}{2}},\ t \in \mathbb{R}.$$

2. En utilisant la définition de la fonction caractéristique et l'indépendance de X et Y on obtient que

$$\forall t \in \mathbb{R},\ \Phi_S(t) = \mathbb{E}(e^{it(X+Y)}) = \mathbb{E}(e^{itX}e^{itY}) = \mathbb{E}(e^{itX})\mathbb{E}(e^{itY}) = \Phi_X(t)\Phi_Y(t) = \left(e^{-\frac{t^2}{2}}\right)^2 = e^{-t^2},$$

ce qui montrer que S suit une loi normale $\mathcal{N}(0,2)$.

3. Soit $X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ deux variables aléatoires indépendantes. En remarquant que $X_1 = m_1 + \sigma_1 Z_1$ et $X_2 = m_2 + \sigma_2 Z_2$ où Z_1 et Z_2 sont deux v.a. indépendantes de loi $\mathcal{N}(0,1)$, on a que (d'après l'exercice 10.1)

$$\forall t \in \mathbb{R},\ \Phi_{X_1}(t) = e^{itm_1} e^{-\frac{\sigma_1^2}{2}t^2},$$

et

$$\forall t \in \mathbb{R},\ \Phi_S(t) = \Phi_{X_1}(t)\Phi_{X_2}(t) = e^{itm_1}\Phi_{Z_1}(t\sigma_1)e^{itm_2}\Phi_{Z_2}(t\sigma_2) = e^{it(m_1+m_2)}e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}.$$

On reconnait donc la fonction caractéristique d'une loi $\mathcal{N}(m_1+m_2, \sigma_1^2+\sigma_2^2)$, et on retrouve donc bien le fait que

$$S \sim \mathcal{N}(m_1+m_2, \sigma_1^2+\sigma_2^2).$$

Exercice 10.4 : *Lois gaussiennes*

1. D'après l'exercice 10.3, on a que pour tout $k = 1, \ldots, n$

$$\forall t \in \mathbb{R},\ \Phi_{X_k}(t) = e^{itm} e^{-\frac{\sigma^2}{2}t^2}.$$

Donc, par indépendance de $X_1, \ldots, X_n$ on obtient que

$$\forall t \in \mathbb{R},\ \Phi_{\overline{X}}(t) = \mathbb{E}\left(e^{it\frac{1}{n}\sum_{k=1}^n X_k}\right) = \prod_{k=1}^n \mathbb{E}\left(e^{i\frac{t}{n}X_k}\right) = \prod_{k=1}^n \Phi_{X_k}\left(\frac{t}{n}\right) = \left(\Phi_{X_1}\left(\frac{t}{n}\right)\right)^n = e^{itm}e^{-\frac{\sigma^2}{2n}t^2}.$$

On a donc que $\overline{X} \sim \mathcal{N}\left(m, \dfrac{\sigma^2}{n}\right)$, ce qui implique que $\mathbb{E}(\overline{X}) = m$ et $\mathrm{Var}(\overline{X}) = \dfrac{\sigma^2}{n}$.

2. On va chercher à écrire $U_k = X_k - \overline{X}$ sous la forme d'une somme de v.a. indépendantes (attention X_i et $\overline{X}$ ne sont pas indépendantes) :

$$U_k = X_k - \frac{1}{n}\sum_{j=1}^n X_j = -\frac{1}{n}\sum_{j=1, j\neq k}^n X_j + \frac{n-1}{n}X_k.$$

Par indépendance de $X_1, \ldots, X_n$, on obtient donc que

$$\begin{aligned}
\forall t \in \mathbb{R},\ \Phi_{U_k}(t) &= \Phi_{X_k}\left(\frac{(n-1)t}{n}\right)\prod_{j=1, j\neq k}^n \Phi_{X_j}\left(-\frac{t}{n}\right) \\
&= e^{i\frac{(n-1)t}{n}m}e^{-\frac{\sigma^2}{2}\frac{(n-1)^2t^2}{n^2}}\left(e^{-i\frac{t}{n}m}e^{-\frac{\sigma^2}{2}\frac{t^2}{n^2}}\right)^{n-1} \\
&= e^{-\frac{\sigma^2}{2}\frac{n-1}{n}t^2}.
\end{aligned}$$

On donc que $U_k \sim \mathcal{N}\left(0, \dfrac{n-1}{n}\sigma^2\right)$, ce qui implique que $\mathbb{E}(U_k) = 0$ et $\mathrm{Var}(U_k) = \dfrac{n-1}{n}\sigma^2$.

3. Par définition de la covariance, en utilisant l'indépendance de $X_1, \dots, X_n$ et étant donné que $\mathbb{E}(U_i) = \mathbb{E}(U_j) = 0$ on obtient que

$$\begin{aligned}
\mathrm{Cov}(U_i, U_j) &= \mathbb{E}(U_i U_j) - \mathbb{E}(U_i)\mathbb{E}(U_j) = \mathbb{E}\left((X_i - \overline{X})(X_j - \overline{X})\right) \\
&= \mathbb{E}(X_i X_j) - \mathbb{E}(X_i \overline{X}) - \mathbb{E}(X_j \overline{X}) + \mathbb{E}(\overline{X}^2) \\
&= \mathbb{E}(X_i)\mathbb{E}(X_j) - \frac{2}{n} \sum_{k=1, k \neq i}^{n} \mathbb{E}(X_i)\mathbb{E}(X_k) - \frac{2}{n}\mathbb{E}(X_i^2) + \mathbb{E}(\overline{X}^2) \\
&= m^2 - \frac{2(n-1)}{n} m^2 - \frac{2}{n}\left(\sigma^2 + m^2\right) + \left(\frac{\sigma^2}{n} + m^2\right) = -\frac{\sigma^2}{n}.
\end{aligned}$$

De même, on a que

$$\begin{aligned}
\mathrm{Cov}(U_i, \overline{X}) &= \mathbb{E}(U_i \overline{X}) - \mathbb{E}(U_i)\mathbb{E}(\overline{X}) = \mathbb{E}(U_i \overline{X}) = \mathbb{E}(X_i \overline{X}) - \mathbb{E}(\overline{X}^2) \\
&= -\frac{n-1}{n} m^2 - \frac{1}{n}\left(\sigma^2 + m^2\right) + \left(\frac{\sigma^2}{n} + m^2\right) = 0.
\end{aligned}$$

Etant donné que U_i et $\overline{X}$ sont des v.a. gaussiennes, on en déduit donc qu'elle sont indépendantes.

4. Par linéarité de l'espérance et utilisant le fait que $U_i \sim \mathcal{N}\left(0, \frac{n-1}{n}\sigma^2\right)$, on obtient que

$$\mathbb{E}(S^2) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(U_i^2) = \mathrm{Var}(U_1) = \frac{n-1}{n}\sigma^2,$$

car $\mathbb{E}(U_i^2) = \mathrm{Var}(U_1)$ pour tout $i = 1, \dots, n$. On ne trouve donc pas $\mathbb{E}(S^2) = \sigma^2$ comme on aurait pu le penser a priori !

5. Par définition, on a que

$$\begin{aligned}
S^2 &= \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 - 2\left(\frac{1}{n} \sum_{i=1}^{n} X_i \overline{X}\right) + \overline{X}^2 \\
&= \frac{1}{n} \sum_{i=1}^{n} X_i^2 - 2\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) \overline{X} + \overline{X}^2 \\
&= \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \overline{X}^2,
\end{aligned}$$

et donc

$$\begin{aligned}
n\left(S^2 + (\overline{X} - m)^2\right) &= \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 + n\left(\overline{X}^2 - 2\overline{X}m + m^2\right) \\
&= \sum_{i=1}^{n} X_i^2 - 2\left(\sum_{i=1}^{n} X_i\right) m + nm^2 = \sum_{i=1}^{n} \left(X_i^2 - 2X_i m + m^2\right) \\
&= \sum_{i=1}^{n} (X_i - m)^2.
\end{aligned}$$

6. Déterminons la fonction caractéristique de la v.a. $Z_k = (X_k - m)^2$. En utilisant le fait que $X_k \sim \mathcal{N}(m, \sigma^2)$ on obtient que

$$\begin{aligned}
\forall t \in \mathbb{R},\ \Phi_{Z_k}(t) &= \int_{\mathbb{R}} e^{it(x-m)^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \\
&= \int_{\mathbb{R}} e^{itu^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}} du,\ (\text{en posant } u = x - m) \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} e^{\left(it - \frac{1}{2\sigma^2}\right)u^2} du.
\end{aligned}$$

Donc, en faisant le changement de variables $y = \left(\frac{1}{2\sigma^2} - it\right)^{1/2} x$, et utilisant le fait que

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2} dy = \frac{1}{\sqrt{2}},$$

on obtient que

$$\begin{aligned}
\forall t \in \mathbb{R},\ \Phi_{Z_k}(t) &= \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{\left(\frac{1}{2\sigma^2} - it\right)^{1/2}} \int_{\mathbb{R}} e^{-y^2} dy = \frac{1}{\sqrt{2}\sigma} \frac{1}{\left(\frac{1}{2\sigma^2} - it\right)^{1/2}} \\
&= \frac{1}{(1 - 2i\sigma^2 t)^{1/2}},
\end{aligned}$$

On en déduit donc que

$$\forall t \in \mathbb{R},\ \prod_{k=1}^{n} \Phi_{(X_k - m)^2}\left(\frac{t}{n}\right) = \frac{1}{\left(1 - 2i\sigma^2 \frac{t}{n}\right)^{n/2}} \quad \text{et} \quad \Phi_{(\overline{X} - m)^2}(t) = \frac{1}{\left(1 - 2i\frac{\sigma^2}{n}t\right)^{1/2}}$$

car $\overline{X} \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$. Donc, en utilisant la relation

$$S^2 + (\overline{X} - m)^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2,$$

et l'indépendance de $U_i = X_i - m$ avec $\overline{X} - m$, on obtient finalement que $S^2 = \frac{1}{n} \sum_{i=1}^{n} U_i^2$ et $(\overline{X} - m)^2$ sont indépendantes ce qui implique que

$$\forall t \in \mathbb{R},\ \Phi_{S^2}(t) \Phi_{(\overline{X} - m)^2}(t) = \prod_{k=1}^{n} \Phi_{(X_k - m)^2}\left(\frac{t}{n}\right)$$

et donc que

$$\begin{aligned}
\forall t \in \mathbb{R},\ \Phi_{S^2}(t) &= \frac{1}{\left(1 - 2i\sigma^2 \frac{t}{n}\right)^{n/2}} \left(1 - 2i\frac{\sigma^2}{n}t\right)^{1/2} \\
&= \frac{1}{\left(1 - 2i\frac{\sigma^2}{n}t\right)^{(n-1)/2}}.
\end{aligned}$$

On reconnait ici la fonction caractéristique d'une loi Gamma de paramètres $\left(\frac{n-1}{2}, \frac{n}{2\sigma^2}\right)$ qui est plus connue sous le nom de loi du χ^2 à $(n-1)$ degrés de liberté de paramètre $\frac{\sigma}{\sqrt{n}}$. On sait que l'espérance d'une v.a. Z de loi Gamma de paramètres (α, λ) est (cf. Exercice 2.5) $\mathbb{E}(Z) = \frac{\alpha}{\lambda}$ et $\mathrm{Var}(Z) = \frac{\alpha}{\lambda^2}$, ce qui donne que

$$\mathbb{E}(S^2) = \frac{n-1}{n}\sigma^2 \quad \text{et} \quad \mathrm{Var}(S^2) = \frac{2(n-1)}{n^2}\sigma^4.$$

10 Convergence de variables aléatoires - PC11

Énoncés des exercices

Exercice 11.1 : *Loi binomiale et loi de Poisson*
Soit Z une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda > 0$.

1. Calculer la fonction caractéristique $\Phi_Z(t)$.
2. Pour tout entier n, on considère une variable aléatoire X_n qui suit une loi binomiale $B(n, p_n)$ pour laquelle lorsque n tend vers l'infini, p_n tend vers 0 et np_n tend vers λ. Calculer la fonction caractéristique de X_n.
3. En déduire que la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers la variable aléatoire qui suit une loi de Poisson de paramètre $\lambda > 0$.
4. Application : une entreprise fabrique des produits dont 1% sont défectueux. Les produits sont vendus par paquets de 100 et garantis à 98%. Modéliser le problème. Donner une interprétation de la garantie. Calculer la probabilité que cette garantie soit en défaut.

Exercice 11.2 : *Loi uniforme discrète et loi uniforme continue*

1. Soit X une variable aléatoire continue uniforme sur $[0, 1]$. Calculer sa fonction caractéristique $\Phi_X(t)$.
2. On considère pour tout $n > 0$ une variable aléatoire X_n discrète et uniforme sur $\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}$. Déterminer sa fonction caractéristique $\Phi_{X_n}(t)$.
3. Démontrer que la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers la variable aléatoire X de loi uniforme sur $[0, 1]$.

Exercice 11.3 : *Erreur d'arrondi d'un calculateur*

1. Un calculateur arrondit les nombres à $0,5.10^{-k}$ (k chiffres après la virgule). Ce calculateur effectue $n = 10^6$ opérations indépendantes. On modélise l'erreur commise à chaque pas par une variable aléatoire continue uniforme sur $[-a, a]$ avec $a = 0,5.10^{-k}$. Soit S l'erreur totale après n opérations. En utilisant le théorème de la limite centrale déterminer la probabilité pour que l'erreur totale S soit inférieure à 1000 fois l'erreur élémentaire d'arrondi, c'est-à-dire $\mathbb{P}(|S| \leqslant 10^3 a)$.
2. Application : dans le cas d'un calculateur qui effectue un arrondi à $k = 6$ chiffres après la virgule, quelle est la probabilité d'avoir une précision meilleure que le millième après un million d'opérations ?

Exercice 11.4 : *Calcul d'intégrale par Monte-Carlo*
Soit f une fonction définie sur $[a, b]$ et à valeurs sur $[m, M]$ avec $(m > 0)$. On note $\mathcal{D} = [a, b] \times [m, M]$ et $\mathcal{A} = \{(x, y) \in \mathcal{D} : f(x) \geqslant y\}$. Soit p le rapport de l'aire de $\mathcal{A}$ sur l'aire de $\mathcal{D}$. Pour tout $k = 1, 2, \ldots, n$, on considère un couple de variables aléatoires (X_k, Y_k) de loi uniforme sur $\mathcal{D}$. On définit

$$Z_k = \begin{cases} 1 & \text{si} \quad Y_k \leqslant f(X_k) \\ 0 & \text{si} \quad Y_k > f(X_k), \end{cases}$$

et on pose $\overline{Z}_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} Z_k$. On suppose en outre que les couples (X_k, Y_k) sont indépendants entre eux.

1. Calculer $\mathbb{E}(Z_k)$, $\mathrm{Var}(Z_k)$, $\mathbb{E}(\overline{Z}_n)$ et $\mathrm{Var}(\overline{Z}_n)$
2. En utilisant la loi des grands nombres montrer que $\overline{Z}_n \to p$ presque sûrement quand n tend vers l'infini.
3. En utilisant l'inégalité de Bienaymé-Tchebichev, déterminer n pour que $\overline{Z}_n$ approche p à 10^{-2} près dans 95% des cas.
4. En utilisant le théorème de la limite centrale, déterminer la loi suivie par $\overline{Z}_n$ quand n est grand. Fournir une nouvelle valeur de n permettant d'atteindre la précision fixée dans la question précédente. Expliquer l'amélioration du résultat.

Corrigés des exercices

Exercice 11.1 : *Loi binomiale et loi de Poisson*

1. Etant donné que Z suit une loi de Poisson, il s'agit donc d'une v.a. discrète à valeurs dans $\mathbb{N}$ telle que

$$\forall k \in \mathbb{N},\ \mathbb{P}(Z = k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Par définition, sa fonction caractéristique est donc donnée par

$$\forall t \in \mathbb{R},\ \phi_Z(t) = \mathbb{E}(e^{itZ}) = \sum_{k=0}^{+\infty} e^{itk-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\frac{(\lambda e^{it})^k}{k!} = e^{-\lambda}e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

2. Etant donné que X_n suit une loi binomiale $B(n, p_n)$, on obtient que

$$\forall t \in \mathbb{R},\ \phi_{X_n}(t) = \mathbb{E}(e^{itX_n}) = \sum_{k=0}^{n} e^{itk}\mathbb{P}(X_n = k) = \sum_{k=0}^{n} C_n^k(p_n e^{it})^k(1-p_n)^{n-k} = \left(1 + p_n(e^{it} - 1)\right)^n.$$

3. On va montrer que sous les hypothèses $p_n \to 0$ et $np_n \to \lambda$ (quand $n \to +\infty$) alors

$$\forall t \in \mathbb{R},\ \lim_{n\to+\infty} \phi_{X_n}(t) = e^{\lambda(e^{it}-1)}.$$

Pour cela, on pose $c = e^{it} - 1$ (nombre complexe dépendant de t mais indépendant de n). Ainsi, on se ramène donc au problème de montrer que

$$\lim_{n\to+\infty} (1 + p_n c)^n = e^{\lambda c},$$

qui est équivalent à montrer que

$$\lim_{n\to+\infty} \left(1 + \frac{z_n}{n}\right)^n = 1.$$

où $z_n = n(e^{-\frac{\lambda c}{n}} - 1) + np_n e^{-\frac{\lambda c}{n}}c$. On peut tout d'abord remarquer que sous l'hypothèse que $np_n \to \lambda > 0$, alors $\lim_{n\to+\infty} z_n = 0$. On a ensuite la majoration suivante :

$$\begin{aligned}
\left|\left(1 + \frac{z_n}{n}\right)^n - 1\right| &= \left|\sum_{k=1}^{n} C_n^k \left(\frac{z_n}{n}\right)^k\right| \\
&\leqslant \sum_{k=1}^{n} C_n^k \left(\frac{|z_n|}{n}\right)^k \\
&= \sum_{k=0}^{n-1} C_n^{k+1} \left(\frac{|z_n|}{n}\right)^{k+1} \\
&= \sum_{k=0}^{n-1} C_{n-1}^k \left(\frac{|z_n|}{n}\right)^k \frac{|z_n|}{k+1} \leqslant |z_n| \left(1 + \frac{|z_n|}{n}\right)^{n-1}.
\end{aligned}$$

Etant donné que $\ln\left(\left(1 + \frac{|z_n|}{n}\right)^{n-1}\right) = (n-1)\ln\left(1 + \frac{|z_n|}{n}\right) = \frac{n-1}{n}|z_n| + o(\frac{n-1}{n}|z_n|) \to 0$, on a que $\lim_{n\to+\infty} \left(1 + \frac{|z_n|}{n}\right)^{n-1} = 1$, et donc

$$\lim_{n\to+\infty} \left|\left(1 + \frac{z_n}{n}\right)^n - 1\right| = 0,$$

car $\lim\limits_{n\to+\infty} |z_n| = 0$. Ainsi on obtient que

$$\forall t \in \mathbb{R},\ \lim_{n\to+\infty} \phi_{X_n}(t) = e^{\lambda(e^{it}-1)} = \phi_Z(t),$$

ce qui montre la convergence en loi de X_n vers une v.a. Z qui suit une loi de Poisson de paramètre $\lambda > 0$.

4. On garantit que dans un paquet de 100, au moins 98 produits sont dans un bon état de fonctionnement. Notons X_{100}, la variable aléatoire "nombre de produits défectueux dans un paquet de 100". D'après l'énoncé, X_{100} suit une loi binomiale $B\left(100, \dfrac{1}{100}\right)$. On est donc bien dans les hypothèses des questions précédentes avec $n = 100$, $p_n = \dfrac{1}{100}$ et $np_n = \lambda = 1$. Ainsi, on peut considérer que X_{100} suit approximativement une loi de Poisson de paramètre $\lambda = 1$ et donc

$$\mathbb{P}(X_{100} = k) \approx \frac{e^{-1}}{k!} = \frac{1}{ek!}.$$

La probabilité que la garantie soit en défaut est (en utilisant la loi binomiale)

$$\begin{aligned}
\mathbb{P}(X_{100} > 2) &= 1 - \mathbb{P}(X_{100} \leqslant 2) = 1 - \sum_{k=0}^{2} \mathbb{P}(X_{100} = k) \\
&= 1 - \left[\left(\frac{99}{100}\right)^{100} + 100\left(\frac{99}{100}\right)^{99} + \frac{100\times 99}{2}\left(\frac{99}{100}\right)^{98}\right] \approx 7,94.10^{-2}.
\end{aligned}$$

et en utilisant l'approximation par une loi de Poisson

$$\mathbb{P}(X_{100} > 2) = 1 - \mathbb{P}(X_{100} \leqslant 2) = 1\left[\frac{1}{e} + \frac{1}{e} + \frac{1}{2e}\right] = 1 - \frac{5}{2e} \approx 8,03.10^{-2}.$$

<u>Exercice 11.2</u> : *Loi uniforme discrète et loi uniforme continue*

1. La densité de X est $f_X(x) = \mathbb{1}_{[0,1]}(x),\ x \in \mathbb{R}$. Donc, par définition de la fonction caractéristique,

$$\forall t \in \mathbb{R},\ \phi_X(t) = \mathbb{E}(e^{itX}) = \int_0^1 e^{itx}dx = \frac{e^{it}-1}{it}.$$

2. SI X_n est de loi uniforme sur $\left\{0, \dfrac{1}{n}, \dfrac{2}{n}, \dots, \dfrac{n-1}{n}, 1\right\}$, alors X_n est une v.a. discrète telle que $\mathbb{P}\left(X_n = \dfrac{k}{n}\right) = \dfrac{1}{n+1}$, $k = 0, \dots, n$. Ceci implique que

$$\forall t \in \mathbb{R},\ \phi_{X_n}(t) = \mathbb{E}(e^{itX_n}) = \sum_{k=0}^{n} e^{it\frac{k}{n}}\mathbb{P}\left(X_n = \frac{k}{n}\right) = \frac{1}{n+1}\sum_{k=0}^{n}\left(e^{i\frac{t}{n}}\right)^k = \frac{1}{n+1}\left(\frac{1-e^{i\frac{n+1}{n}t}}{1-e^{i\frac{t}{n}}}\right).$$

3. Quand $n \to +\infty$, on a que $1 - e^{i\frac{t}{n}} \sim -i\dfrac{t}{n}$ et donc

$$\frac{1}{n+1}\left(\frac{1-e^{i\frac{n+1}{n}t}}{1-e^{i\frac{t}{n}}}\right) \sim \frac{n}{n+1}\frac{e^{i\frac{n+1}{n}t}-1}{it}$$

quand $n \to +\infty$. Ce qui prouve que

$$\forall t \in \mathbb{R}, \lim_{n\to+\infty} \phi_{X_n}(t) = \frac{e^{it}-1}{it},$$

et donc X_n converge en loi vers une v.a. X continue et de loi uniforme sur $[0,1]$ (d'après la question 1).

Exercice 11.3 : *Erreur d'arrondi d'un calculateur*

1. Soit X_i l'erreur à la i-ème opération. On a que $S = \sum_{i=1}^{10^6} X_i$. Les X_i étant indépendantes, de même loi (continue uniforme sur $[-a,a]$) qui admet une espérance et une variance finies, d'après le théorème central limite (TCL), la v.a. centrée réduite $\dfrac{S-\mathbb{E}(S)}{\sqrt{\mathrm{var}(S)}}$ suit approximativement une loi $\mathcal{N}(0,1)$ quand n est grand.

Or $\mathbb{E}(S) = 10^6\mathbb{E}(X_i)$, avec $\mathbb{E}(X_i) = 0$ car X_i, erreur due à l'arrondi à $0,5.10^{-k}$ près, suit la loi uniforme sur $[-a,a]$, et $\mathrm{var}(S) = 10^6\mathrm{var}(X_i)$ (variables indépendantes), avec $\mathrm{var}(X_i) = \mathbb{E}(X_i^2) = \dfrac{1}{2a}\displaystyle\int_{-a}^{a} x^2\,dx = \dfrac{a^2}{3}$, et donc $\sqrt{\mathrm{var}(S)} = 10^3\dfrac{a}{\sqrt{3}}$.

Ainsi, par le TCL, la v.a. et $\dfrac{\sqrt{3}}{10^3 a}S$ suit approximativement une loi $\mathcal{N}(0,1)$ quand n est grand. On a alors l'approximation suivante :

$$\begin{aligned}\mathbb{P}(|S| \leqslant 10^3 a) &= \mathbb{P}\left(\frac{|S|}{10^3a} \leqslant 1\right) = \mathbb{P}\left(\frac{\sqrt{3}}{10^3a}|S| \leqslant \sqrt{3}\right)\\ &\approx \Phi(\sqrt{3}) - \Phi(-\sqrt{3}) = 2\Phi(\sqrt{3}) - 1,\end{aligned}$$

où Φ est la fonction de répartition de la loi $\mathcal{N}(0,1)$. Etant donné que $\Phi(\sqrt{3}) \approx \Phi(1,732) \approx 0,958$, on obtient donc que

$$\mathbb{P}(|S| \leqslant 10^3 a) \approx 0,916.$$

On peut donc conclure qu'après 1 million d'opérations, la probabilité d'avoir une erreur supérieure à 1000 fois l'erreur de troncature est inférieure à 10 %.

2. On prend $a = 0,5.10^{-6}$ (arrondi à 5 chiffres après la virgule). Par le TCL, on en déduit donc que $\dfrac{\sqrt{3}}{10^3a}S = 2\sqrt{3}10^3 S$ peut être approximée par la loi gaussienne $\mathcal{N}(0,1)$. Ainsi,

$$\mathbb{P}(|S| \leqslant 10^{-3}) = \mathbb{P}(2\sqrt{3}10^3|S| \leqslant 2\sqrt{3}) \approx \phi(2\sqrt{3}) - \phi(-2\sqrt{3}) = 2\phi(2\sqrt{3}) - 1 \approx 99,95\%.$$

Ainsi, après 1 million d'opérations, on est presque sûr d'avoir une précision meilleure que le millième.

Exercice 11.4 : *Calcul d'intégrale par Monte-Carlo*

1. On a que $Z_k = 1$ si $f(X_k) \geqslant Y_k$ c'est à dire si (X_k, Y_k) est un couple de points au dessous de la courbe $x \mapsto f(x)$ i.e. qui appartient à l'ensemble $\mathcal{A}$. On a donc que

$$\mathbb{P}(Z_k = 1) = \frac{\text{Aire de } \mathcal{A}}{\text{Aire de } \mathcal{D}} = p$$

avec Aire de $\mathcal{D} = (M - m) \times (b - a)$ et

$$\text{Aire de } \mathcal{A} = \int_a^b (f(x) - m)\, dx = \int_a^b f(x)\, dx - m(b - a).$$

Etant donné que Z_k suit une loi de Bernouilli de paramètre p, on obtient immédiatement que $\mathbb{E}(Z_k) = p$ et $\operatorname{var}(Z_k) = p(1-p)$, ainsi que $\mathbb{E}(\overline{Z}_n) = \frac{\sum_{i=1}^{n} \mathbb{E}(Z_i)}{n} = p$ et $\operatorname{var}(\overline{Z}_n) = \frac{1}{n^2}\operatorname{var}\left(\sum_{i=1}^{n} Z_i\right) = \frac{1}{n}\operatorname{var}(Z_i) = \frac{p(1-p)}{n}$ car les Z_i sont indépendantes. D'où, finalement

$$\mathbb{E}(\overline{Z}_n) = p \text{ et } \operatorname{var}(\overline{Z}_n) = \frac{1}{n} p(1-p).$$

2. Par la loi forte des grands nombres, on obtient immédiatement que $\overline{Z}_n \stackrel{p.s.}{\to} \mathbb{E}(Z_i) = p$.

3. On cherche à calculer la probabilité $\mathbb{P}\left(|\overline{Z}_n - p| \leqslant 10^{-2}\right) \geqslant 0,95$. Pour cela on va calculer la probabilité $\mathbb{P}\left(|\overline{Z}_n - p| > 10^{-2}\right) < 0,05 = 5.10^{-2}$. Par l'inégalité de Bienaymé-Tchebichev on a que

$$\mathbb{P}\left(|\overline{Z}_n - p| > 10^{-2}\right) \leqslant \frac{1}{10^{-4}} \mathbb{E}((\overline{Z}_n - p)^2) = \frac{\operatorname{Var}(\overline{Z}_n)}{10^{-4}} = \frac{p(1-p)}{10^{-4} n}.$$

On cherche donc un entier n tel que $\frac{p(1-p)}{10^{-4} n} < 5.10^{-2}$, soit $n > \frac{p(1-p)}{5.10^{-6}}$.

Soit $\varphi(p) = p(1-p) = p - p^2$. On a que $\varphi'(p) = 1 - 2p$, et donc φ admet un maximum en $p = 1/2$ avec $\varphi(1/2) = 1/4$. Ainsi, pour avoir $n > \frac{p(1-p)}{5.10^{-6}}$ avec p quelconque, il suffit de prendre $n > \frac{1}{4}\frac{1}{5.10^{-6}}$, soit $n > \frac{10^6}{20} = 5.10^4$ i.e. $n > 50\,000$.

4. On cherche de nouveau un entier tel n que $\mathbb{P}\left(|\overline{Z}_n - p| > \epsilon\right) < \alpha$ avec $\epsilon = 10^{-2}$ et $\alpha = 0,05$. D'après le TCL, la v.a. $\frac{\overline{Z}_n - \mathbb{E}(\overline{Z}_n)}{\sqrt{\operatorname{Var}(\overline{Z}_n)}} = \sqrt{n}\frac{\overline{Z}_n - p}{\sqrt{p(1-p)}}$ suit approximativement une loi $\mathcal{N}(0,1)$ quand n est grand. On a donc l'approximation suivante :

$$\begin{aligned}
\mathbb{P}\left(|\overline{Z}_n - p| > \epsilon\right) &= \mathbb{P}\left(\sqrt{n}\frac{|\overline{Z}_n - p|}{\sqrt{p(1-p)}} > \frac{\sqrt{n}}{\sqrt{p(1-p)}}\epsilon\right) \approx 1 - \Phi\left(\frac{\sqrt{n}}{\sqrt{p(1-p)}}\epsilon\right) - \Phi\left(-\frac{\sqrt{n}}{\sqrt{p(1-p)}}\epsilon\right) \\
&\approx 2\left(1 - \Phi\left(\frac{\sqrt{n}}{\sqrt{p(1-p)}}\epsilon\right)\right)
\end{aligned}$$

où Φ est la fonction de répartition de la loi $\mathcal{N}(0,1)$. On cherche donc un entier n tel que

$$2\left(1 - \Phi\left(\frac{\sqrt{n}}{\sqrt{p(1-p)}}\epsilon\right)\right) < \alpha,$$

ce qui est équivalent à la condition $n > \frac{p(1-p)}{\epsilon^2}\left(\Phi^{-1}(1-\alpha/2)\right)^2$. Comme on l'a vu à la question précédente, pour avoir $n > \frac{p(1-p)}{\epsilon^2}\left(\Phi^{-1}(1-\alpha/2)\right)^2$ avec p quelconque, il suffit de prendre $n > \frac{1}{4\epsilon^2}\left(\Phi^{-1}(1-\alpha/2)\right)^2$. Or pour $\alpha = 0,05$ on a que $\Phi^{-1}(1-\alpha/2) \approx 1,96$, ce qui conduit à choisir n tel que

$$n > \frac{1}{4.10^{-4}}(1.96)^2 \approx 9604.$$

On trouve donc une valeur de n qui est $50000/9604 \approx 5,2$ fois plus petite qu'à la question 3. Ainsi, l'information relative à la convergence de $\overline{Z}_n$ vers une loi gaussienne est plus riche. L'inégalité de Bienaymé-Tchebichev ne fait aucune hypothèse sur la variable aléatoire concernée, car elle s'applique dans tous les cas (indépendamment de la loi) et fournit donc une majoration plus grossière de la probabilité de déviation $\mathbb{P}\left(|\overline{Z}_n - p| > \epsilon\right)$, basée uniquement sur l'information donnée par l'espérance et la variance de $\overline{Z}_n$.

11 Chaînes de Markov - PC12 & PC13

Énoncés des exercices

Exercice 12.1 : *Etude d'une chaîne de Markov à 4 états*
On définit une chaîne de Markov sur les états $\{1,2,3,4\}$ par sa matrice de transition P définie par

$$P = \begin{bmatrix} 1/3 & a & 0 & 0 \\ b & 1/2 & 0 & 0 \\ 0 & 0 & 1/4 & c \\ 0 & 2/5 & d & 2/5 \end{bmatrix}.$$

1. Déterminer les réels a, b, c et d et tracer le graphe de transition de la chaîne.
2. Quelle est la nature des états 1, 2, 3 et 4 ?
3. Déterminer les différentes classes d'équivalence de la chaîne et leur nature. La chaîne est-elle irréductible ?
4. Déterminer la distribution stationnaire de la chaîne restreinte aux états $\{1,2\}$ (caractérisée par sa matrice de transition notée A).
5. Déterminer la loi de probabilité du temps T_1 de premier retour à l'état 1. Calculer $\mathbb{E}(T_1)$.
6. Montrer que $p_{1,1}(n) = \frac{1}{7}\left(3 + 4\left(-\frac{1}{6}\right)^n\right)$ (utiliser la matrice A^n).
7. Calculer la limite de $p_{1,1}(n)$ quand $n \to +\infty$. Vérifier que $\mathbb{E}(T_1) = \frac{1}{\lim_{n\to+\infty} p_{1,1}(n)}$.

Exercice 12.2 : *Etude d'une chaîne de Markov à 6 états*
On définit une chaîne de Markov sur les états $\{1,2,3,4,5,5\}$ par sa matrice de transition P définie par

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 3/4 & 0 & 0 & 0 \\ 0 & 0 & 1/5 & 0 & 4/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

1. Tracer le graphe de transition de la chaîne.
2. Déterminer les différentes classes d'équivalence de la chaîne et leur nature.
3. Soit la distribution initiale $\pi(0) = (\mathbb{P}(X_0 = 1), \ldots, \mathbb{P}(X_0 = 6)) = (0, 0, 1/2, 0, 1/2, 0)$. Calculer $\pi(1) = (\mathbb{P}(X_1 = 1), \ldots, \mathbb{P}(X_1 = 6))$ et $\pi(2) = (\mathbb{P}(X_2 = 1), \ldots, \mathbb{P}(X_2 = 6))$.
4. On considère la sous-chaîne sur $\{1,2\}$. Quelle est sa nature ? Déterminer sa matrice de transition et sa distribution stationnaire π.
5. On considère à nouveau la chaîne initiale. Soit $\pi(0) = (0,0,1,0,0,0)$ et $n \in \mathbb{N}^*$. Calculer $\mathbb{P}(X_n > 3)$ et $\mathbb{P}(X_n < 3)$. Que peut-on dire de la limite de $\pi(n)$ quand $n \to +\infty$ où $\pi(n)$ est le vecteur $(\mathbb{P}(X_n = 1), \ldots, \mathbb{P}(X_n = 6))$?

Exercice 12.3 : *Marche aléatoire.*
Etude des marches aléatoires sur :

1. $\mathbb{Z}$

2. $\mathbb{Z}^2$
3. $\mathbb{Z}^3$

Exercice 13.1 : *Chaîne de victoires.*
Lorsqu'une équipe de foot a joué n matches, on note X_n la série de matches sans défaites en cours. Si $X_n = k$, on note p_k la probabilité de ne pas perdre le match suivant. Ainsi, $\mathbb{P}([X_{n+1} = k+1]/[X_n = k]) = p_k$ et $\mathbb{P}([X_{n+1} = 0]/[X_n = k]) = 1 - p_k$.

1. Montrer rapidement que (X_n) définit une chaîne de Markov irréductible et faire son graphe. Si T_0 est l'instant de premier retour en 0, exprimer $\mathbb{P}([T_0 = n]/[X_0 = 0])$ et donner des conditions pour que la haîne soit transitoire, récurrente nulle ou récurrente positive (à l'aide des $r_n = p_0 p_1 \cdots p_{n-1}$).
2. On suppose que $p_k = p$ pour tout k. Déterminer la distribution stationnaire de la chaîne et la longueur moyenne d'une série sans défaite. (*Application numérique :* $p = \dfrac{1}{2}$)

Exercice 13.2 : *Gestion d'un stock.*
Un stock contient au maximum s pièces identiques. La variable aléatoire X_n est le nombre de pièces au début de la $n^{ième}$ semaine. La variable aléatoire Y_n est le nombre de pièces sorties du stock au cours de la $n^{ième}$ semaine. On supposera que $X_0 = s$ et que pour tout $j = 0, 1, ..., i$

$$\mathbb{P}(Y_n = j | X_n = i) = \frac{1}{1+i},$$

quel que soit $n \in \mathbb{N}^*$. Le stock est alimenté de telle manière que : $X_{n+1} = s - Y_n$ (c'est-à-dire qu'en cours de semaine on le complète à la quantité s sans se préoccuper des pièces qui sortent dans la semaine en cours).

1. Montrer que $(X_n)_n$ est une chaîne de Markov, et qu'elle admet une distribution stationnaire définie par $\pi_i = \dfrac{2(i+1)}{(s+1)(s+2)}$ pour $i = 1, \ldots, s$.
2. Quel est le nombre moyen de pièces disponibles au début de chaque semaine ? Application numérique : $s = 30$.

Exercice 13.3 : *Les urnes d'Ehrenfest.*
On dispose de N boules numérotées de 1 à N, $(N \geqslant 2)$, qui sont réparties dans deux urnes : k dans l'urne U_1 et $N - k$ dans l'urne U_2 (avec $0 \leqslant k \leqslant N$). On tire au sort un numéro compris entre 1 et N de façon equiprobable, puis on change d'urne la boule dont le numéro a été tiré. On note alors X_1 le nombre de boules dans l'urne U_1. On réitère l'opération et on note X_n le nombre de boules dans l'urne U_1 après n opérations.

1. Déterminer la loi de X_1 (dans un premier temps, traiter à part les cas $k = 0$ et $k = N$).
2. Montrer que $\mathbb{E}(X_1) = k\left(1 - \dfrac{2}{N}\right) + 1$.
3. Déterminer la loi conditionnelle de X_{n+1} sachant $X_n = q$, c'est à dire calculer $\mathbb{P}(X_{n+1} = p | X_n = q)$ pour p et q deux entiers entre 0 et N.
4. Déduire des questions précédentes l'espérance conditionnelle de X_{n+1} sachant X_n puis établir une relation de récurrence entre $\mathbb{E}(X_{n+1})$ et $\mathbb{E}(X_n)$.

5. En déduire la valeur de $\mathbb{E}(X_n)$ et déterminer la limite de $\mathbb{E}(X_n)$ quand $n \to +\infty$.

   On peut également considérer que $(X_n)_{n\geqslant 0}$ est une chaîne de Markov à valeurs dans $E = \{0, 1, \dots, N\}$ de loi initiale $\pi(0) = (\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1), \dots, \mathbb{P}(X_0 = N))$ où X_0 est un nombre aléatoire de boules dans l'urne U_1.

6. Donner la matrice de transition P de la chaîne.
7. Tracer le graphe de transition de la chaîne. Est-elle irréductible ?

   On pose $\pi_q = \dfrac{1}{2^N}\dbinom{N}{q}$ et $\pi = (\pi_q)_{0\leqslant q\leqslant N}$

8. Montrer que π est une distribution stationnaire de la chaîne.
9. Montrer que π est l'unique distribution stationnaire et que $(X_n)_{n\geqslant 0}$ est récurrente positive.
10. Soit $0 \leqslant k \leqslant N$. Si l'on démarre avec k boules dans l'urne U_1, au bout de combien de temps en moyenne a-t-on de nouveau k boules dans l'urne U_1 ? Comparer les cas $k = 0$ et $k = 5$ pour $N = 10$. Si N tend vers l'infini, que peut-on dire du temps moyen de retour à $k = 0$ boules dans l'urne U_1 ?

Exercice 13.4 : *Lancés d'un dé.*

Soit X_n la v.a. égale à la valeur maximale obtenue après n jets d'un dé.

1. Montrer que $(X_n)_n$ est une chaîne de Markov.
2. Déterminer la matrice de transition.
3. Déterminer la limite de $\pi(n)$, loi de (X_n), quand n tend vers l'infini.

Exercice 13.5 : *Disponibilité de machines.*

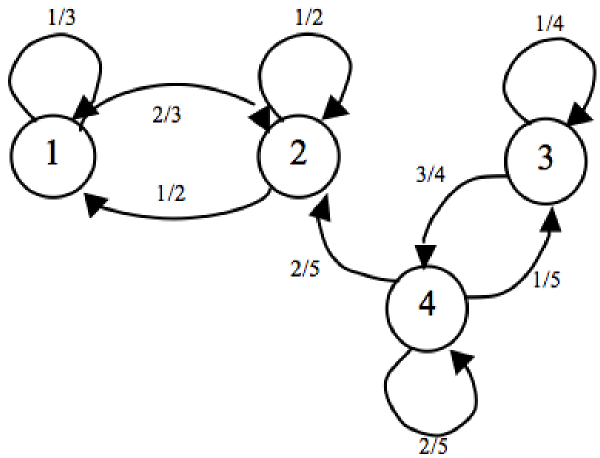
Deux machines identiques fonctionnent indépendamment ; chacune pouvant tomber en panne avec la probabilité q. Soit X_n le nombre de machines en panne au début de la $n^{ième}$ journée.

1. Si une machine tombe en panne elle est réparée la nuit suivante et on ne peut réparer qu'une machine par nuit. Caractériser la chaîne de Markov $(X_n)_n$. Déterminer la distribution stationnaire.
2. Même question en supposant qu'une machine en panne n'est réparée que le lendemain et qu'on ne peut réparer qu'une machine dans la journée.

Corrigés des exercices

Exercice 12.1 : *Etude d'une chaîne de Markov à 4 états*

1. La matrice de transition d'une chaîne de Markov est une matrice stochastique : la somme des coefficients de chacune de ses lignes est égale à 1. On a donc que $a = 2/3$, $b = 1/2$, $c = 3/4$ et $d = 1/5$. Le graphe de transition de cette chaîne de Markov est :



2. L'état 1 est récurrent positif (partant de cet état on est sûr de pouvoir y revenir) et apériodique (par exemple : il existe une boucle sur l'état 1 qui permet un retour en 1 coup) Pour les mêmes raisons l'état 2 est récurrent positif et apériodique. L'état 3 est transitoire (on finit toujours par le quitter vers l'état 4 sans chemin de retour possible) et apériodique. L'état 4 est transitoire (on finit toujours par le quitter vers l'état 2 sans chemin de retour possible) et apériodique.

3. Les états 1 et 2 communiquent entre eux, de même pour les états 3 et 4. Il existe donc 2 classes d'équivalence $\{1, 2\}$ et $\{3, 4\}$. La chaîne n'est donc pas irréductible. La classe $\{1, 2\}$ est récurrente positive et apériodique. La classe $\{3, 4\}$ est transitoire et apériodique. Ainsi, quelque soit le point départ, on finit toujours par quitter la classe $\{3, 4\}$ pour rester dans la classe $\{1, 2\}$.

4. On considère désormais la chaîne réduite à sa classe récurrente $\{1, 2\}$. Sa matrice de transition est :

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1/3 & 2/3 \\ 1/2 & 1/2 \end{bmatrix}.$$

Cette chaîne est récurrente positive et apériodique d'après le théorème (dit d'ergodicité) des chaînes de Markov elle admet une distribution stationnaire unique, $\Pi = (\pi_1, \pi_2)$, qui vérifie : $\Pi = \Pi A$ avec $\pi_1 + \pi_2 = 1$ et $\pi_1, \pi_2 \geqslant 0$, d'où le système d'équations :

$$\begin{aligned} \pi_1 &= \frac{1}{3}\pi_1 + \frac{1}{2}\pi_2 \\ \pi_2 &= \frac{2}{3}\pi_1 + \frac{1}{2}\pi_2 \\ \pi_1 + \pi_2 &= 1, \end{aligned}$$

dont la résolution conduit à la solution : $\pi_1 = \dfrac{3}{7}$ et $\pi_2 = \dfrac{4}{7}$.

5. On appelle T_1 le temps de premier retour à l'état 1. On a alors que

$$\mathbb{P}(T_1 = 1) = a_{11} = \frac{1}{3},\ \mathbb{P}(T_1 = 2) = a_{12}a_{21} = \frac{2}{3}\frac{1}{2} = \frac{1}{3},\ \mathbb{P}(T_1 = 3) = a_{12}a_{22}a_{21} = \frac{2}{3}\frac{1}{2}\frac{1}{2} = \frac{2}{3}\left(\frac{1}{2}\right)^2.$$

On a alors la généralisation : $\forall k \geqslant 2$, $\mathbb{P}(T_1 = k) = \frac{2}{3}\left(\frac{1}{2}\right)^{k-2}\frac{1}{2} = \frac{2}{3}\left(\frac{1}{2}\right)^{k-1}$. On peut alors vérifier que

$$\sum_{k=1}^{+\infty} \mathbb{P}(T_1 = k) = \frac{1}{3} + \frac{2}{3}\sum_{k=2}^{+\infty}\left(\frac{1}{2}\right)^{k-1} = \frac{1}{3} + \frac{2}{3}\sum_{k=1}^{+\infty}\left(\frac{1}{2}\right)^{k} = 1,$$

ce qui confirme que l'état 1 est récurrent positif. On a également que

$$\mathbb{E}(T_1) = \sum_{k=1}^{+\infty} k\mathbb{P}(T_1 = k) = \frac{1}{3} + \frac{2}{3}\sum_{k=2}^{+\infty} k\left(\frac{1}{2}\right)^{k-1} = \frac{1}{3} + \frac{2}{3}\left(\frac{1}{(1-\frac{1}{2})^2} - 1\right) = \frac{1}{3} + 2 = \frac{7}{3}.$$

On retrouve bien, sur ce cas simple, la formule générale vue en cours : $\mathbb{E}(T_1) = \frac{1}{\pi_1}$.

6. Pour calculer A^n on va diagonaliser la matrice A. Les valeurs propres de A sont 1 (toujours valeur propre d'une matrice stochastique) et $tr(A) - 1 = \frac{5}{6} - 1 = -\frac{1}{6}$ On peut prendre comme vecteurs propres respectivement associés à ces 2 valeurs propres : $(1,1)^t$ et $(4,-3)^t$. La matrice de passage de la diagonalisation est donc

$$Q = \begin{bmatrix} 1 & 4 \\ 1 & -3 \end{bmatrix}.$$

et son inverse

$$Q^{-1} = \frac{1}{7}\begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix}.$$

D'où

$$A^n = \frac{1}{7}\begin{bmatrix} 1 & 4 \\ 1 & -3 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{6} \end{bmatrix}^n\begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 3 + 4\left(-\frac{1}{6}\right)^n & 4 - 4\left(-\frac{1}{6}\right)^n \\ 3 - 3\left(-\frac{1}{6}\right)^n & 4 + 3\left(-\frac{1}{6}\right)^n \end{bmatrix}$$

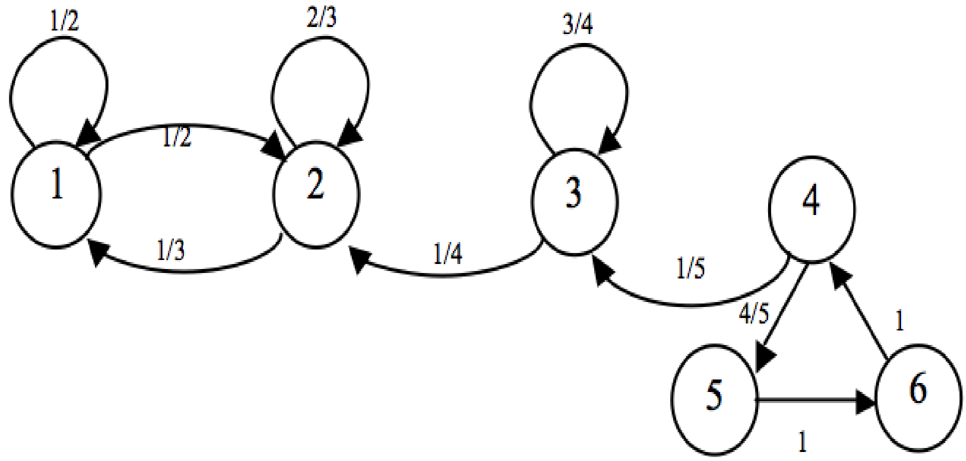
et donc on a bien que $p_{1,1}(n) = \frac{1}{7}\left(3 + 4\left(-\frac{1}{6}\right)^n\right)$.

7. Quand $n \to +\infty$, $p_{1,1}(n) \to \frac{3}{7} = \pi_1$, et donc on vérifie bien que $\mathbb{E}(T_1) = \frac{7}{3} = \frac{1}{\lim_{n\to+\infty} p_{1,1}(n)}$.

De même on peut remarquer que $\lim_{n\to+\infty} p_{2,2}(n) = \frac{4}{7} = \pi_2$.

Exercice 12.2 : *Etude d'une chaîne de Markov à 6 états*

1. Le graphe de transition de cette chaîne de Markov est :



2. Il y a 3 classes d'équivalence $\{1,2\}$, $\{3\}$ et $\{4,5,6\}$. La chaîne n'est donc pas irréductible. La classe $\{1,2\}$ est récurrente positive et apériodique. La classe $\{3\}$ est transitoire et apériodique. La classe $\{4,5,6\}$ est transitoire et périodique de période 3.

3. Soit la distribution initiale $\pi(0) = (0,0,1/2,0,1/2,0)$. On a que

$$\pi(1) = \pi(0)P = (0,1/8,3/8,0,0,1/2)$$

que l'on peut déterminer à partir du graphe, et

$$\pi(2) = \pi(1)P = (1/24, 2/24+3/32, 9/32, 1/2, 0, 0) = (1/24, 17/96, 9/32, 1/2, 0, 0)$$

que l'on peut également déterminer à partir du graphe.

4. On considère désormais la chaîne réduite à sa classe récurrente positive et apériodique $\{1,2\}$. Sa matrice de transition est

$$A = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}.$$

Cette chaîne est récurrente positive et apériodique d'après le théorème (dit d'ergodicité) des chaînes de Markov elle admet une distribution stationnaire unique, $\Pi = (\pi_1, \pi_2)$, qui vérifie : $\Pi = \Pi A$ avec $\pi_1 + \pi_2 = 1$ et $\pi_1, \pi_2 \geqslant 0$, d'où le système d'équations :

$$\begin{aligned} \pi_1 &= \frac{1}{2}\pi_1 + \frac{1}{3}\pi_2 \\ \pi_2 &= \frac{1}{2}\pi_1 + \frac{2}{3}\pi_2 \\ \pi_1 + \pi_2 &= 1, \end{aligned}$$

dont la résolution conduit à la solution : $\pi_1 = \dfrac{2}{5}$ et $\pi_2 = \dfrac{3}{5}$.

5. On considère à nouveau la chaîne complète et on part maintenant de l'état 3 i.e. de la distribution initiale $\pi(0) = (0,0,1,0,0,0)$. D'après le graphe de la chaîne, il est clair que $\mathbb{P}(X_n > 3) = 0$ pour tout $n \geqslant 1$. On a donc que $\mathbb{P}(X_n < 3) = \mathbb{P}(X_n = 1) + \mathbb{P}(X_n = 2) = 1 - \mathbb{P}(X_n = 3)$. Etant donné le graphe de la chaîne, on a que $\mathbb{P}(X_n = 3) = \left(\dfrac{3}{4}\right)^n$, et donc

$$\mathbb{P}(X_n < 3) = 1 - \left(\frac{3}{4}\right)^n.$$

Ainsi, à l'aide de la question 5, on obtient donc que

$$\lim_{n\to+\infty} \pi(n) = (\pi_1, \pi_2, 0, 0, 0, 0) = \left(\frac{2}{5}, \frac{3}{5}, 0, 0, 0, 0\right).$$

Exercice 12.3 : *Marche aléatoire.*

1. L'espace des états est $E = \mathbb{Z}$. Pour être de nouveau au point de départ après n étapes, il faut avoir fait autant de pas vers la droite que vers la gauche : ainsi, n doit être pair ($n = 2m$). Il y a autant de trajets possibles que de $2m$ u-plets avec m "d" et m "g", soit C_{2m}^m, et ils sont tous de probabilité $\left(\frac{1}{2}\right)^{2m}$. Ainsi, $p_{x,x}^{(2m)} = C_{2m}^m \left(\frac{1}{2}\right)^{2m}$ (et $p_{x,x}^{(2m+1)} = 0$).

Pour la nature de la série, on utilise un équivalent grâce à Stirling :

$$p_{xx}^{(2m)} = \frac{(2m)!}{(m!)^2}\left(\frac{1}{2}\right)^{2m} \sim \frac{(2m)^{2m}}{e^{2m}}\sqrt{2\pi\times 2m}\left(\frac{e^m}{m^m\sqrt{2\pi m}}\right)^2 \times \frac{1}{2^{2m}} = \frac{1}{\sqrt{\pi m}}.$$

Ainsi, $p_{xx}^{(2m)} \sim \frac{1}{\sqrt{\pi m}}$ terme général tendant vers 0 d'une série divergente entraîne que la chaîne est récurrente nulle. On démontre sans difficulté qu'une marche aléatoire disymétrique dans $\mathbb{Z}$, donc telle que $p \neq q$, est transiente.

2. L'espace des états est $E = \mathbb{Z}^2$. Pour être de nouveau au point de départ après n étapes, il faut avoir fait dans chacune des 2 directions autant de pas dans un sens que dans l'autre : ainsi, n doit être pair ($n = 2m$). Si il y a $2k$ pas verticaux (et donc $2m-2k$ pas horizontaux), alors il doit y avoir k pas vers le haut, k pas vers le bas, $m-k$ pas vers la gauche et $m-k$ pas vers la droite. Ainsi, dans ces cas-là, il y a $C_{2m}^{2k}\times C_{2k}^k \times C_{2m-2k}^{m-k}$ trajets possibles (on a C_{2m}^{2k} façons de choisir les pas verticaux, puis, parmi ceux-ci, on en choisit C_{2k}^k vers le haut et, parmi les $2m-2k$ horizontaux on choisit les C_{2m-2k}^{m-k} vers la droite par exemple), et ils sont tous de probabilité $\left(\frac{1}{4}\right)^{2m}$. Mais, comme k peut prendre toutes les valeurs de 0 à m, il vient

$$\begin{aligned} p_{xx}^{(2m)} &= \sum_{k=0}^{m} \frac{(2m)!}{(2k)!(2m-2k)!}\frac{(2k)!}{(k!)^2}\frac{(2m-2k)!}{((m-k)!)^2}\left(\frac{1}{4}\right)^{2m} \\ &= C_{2m}^m\left(\frac{1}{4}\right)^{2m}\left(\sum_{k=0}^{m}(C_m^k)^2\right) = \left[C_{2m}^m\left(\frac{1}{2}\right)^{2m}\right]^2 \sim \frac{1}{\pi m} \end{aligned}$$

car $\sum_{k=0}^{m}(C_m^k)^2 = \sum_{k=0}^{m} C_m^k C_m^{m-k} = C_{2m}^m$ (on peut par exemple développer $(1+x)^{2m} = (1+x)^m(1+x)^m$ des 2 façons et identifier le coefficient de x^m dans chacune des expressions). Même conclusion que précédemment.

3. L'espace des états est $E = \mathbb{Z}^3$. Pour être de nouveau au point de départ après n étapes, il faut avoir fait dans chacune des 3 directions autant de pas dans un sens que dans l'autre : ainsi, n doit être pair ($n = 2m$). Si il y a $2k_1$ pas dans la direction 1, $2k_2$ pas dans la direction 2 (et donc $2m-2k_1-2k_2$ pas dans la direction 3), alors il doit y avoir k_1 pas dans chaque sens de la direction 1, k_2 dans ceux de la direction 2 et $m-k_1-k_2$ dans ceux de la direction 3. Ainsi, dans ces cas-là, il y a $C_{2m}^{2k_1}\times C_{2m-2k_1}^{2k_2}\times C_{2k_1}^{k_1}\times C_{2k_2}^{k_2}\times C_{2(m-k_1-k_2)}^{m-k_1-k_2}$ trajets possibles et ils sont

tous de probabilité $\left(\dfrac{1}{6}\right)^{2m}$ (3 directions et 2 sens dans chacune, donc 6 possibilités à chaque pas). Il vient alors

$$\begin{aligned} p_{xx}^{(2m)} &= \sum_{k_1,k_2} \frac{(2m)!}{(2k_1)!(2m-2k_1)!} \frac{(2m-2k_1)!}{(2k_2)!(2m-2k_1-2k_2)!} \frac{(2k_1)!}{(k_1!)^2} \frac{(2k_2)!}{(k_2!)^2} \frac{(2m-2k_1-2k_2)!}{((m-k_1-k_2)!)^2} \left(\frac{1}{6}\right)^{2m} \\ &= \sum_{k_1,k_2} \frac{(2m)!}{(k_1!)^2(k_2!)^2((m-k_1-k_2)!)^2} \left(\frac{1}{6}\right)^{2m} \end{aligned}$$

On pourrait montrer que l'on peut majorer cette expression par un terme équivalent à $\dfrac{K_3}{m^{3/2}}$. Ainsi $p_{xx}^{(2m)}$ est le terme général d'une série convergente donc la chaîne est <u>transitoire</u>.

<u>Exercice 13.1</u> : *Chaîne de victoires.*

1. La valeur de X_n n'est fonction que de X_{n-1} et du résultat du n-ième match (c'est $X_{n-1}+1$ si le n-ième match n'est pas perdu et c'est 0 si le n-ième match est perdu). Ainsi, l'axiome de Markov est vérifié. De plus, $\mathbb{P}(X_n = k+1 \mid X_{n-1} = k) = p_k$ et $\mathbb{P}(X_n = 0 \mid X_{n-1} = k) = 1 - p_k$ indépendants de n donc l'axiome d'homogénéité est aussi vérifié. On a donc bien une chaîne de Markov de matrice de transition $(p_{k,j})$ avec

$$p_{k,j} = \begin{cases} p_k \text{ si } j = k+1 \\ 1 - p_k \text{ si } j = 0 \\ 0 \text{ sinon} \end{cases} .$$

On a $P^{[X_0=0]}([T_0 = 1]) = 1 - p_0$, $P^{[X_0=0]}([T_0 = 2]) = p_0(1-p_1)$ et, pour $n \geqslant 3$, il faut aller jusqu'à $n-1$ matches sans défaites, puis perdre, d'où $P^{[X_0=0]}([T_0 = n]) = p_0 p_1 \cdots p_{n-2}(1-p_{n-1})$.
0 est récurrent si et seulement si $P^{[X_0=0]}([T_0 < +\infty]) = 1$.
Or $P^{[X_0=0]}([T_0 < +\infty]) = P^{[X_0=0]}\left(\bigcup_{k=1}^{+\infty} [T_0 = k]\right) = \lim_{n \to +\infty} P^{[X_0=0]}\left(\bigcup_{k=1}^{n} [T_0 = k]\right)$ avec

$$P^{[X_0=0]}\left(\bigcup_{k=1}^{n} [T_0 = k]\right) = (1-p_0) + (p_0 - p_0 p_1) + \cdots - p_0 p_1 \cdots p_{n-1} = 1 - r_n.$$

$r_n = p_0 p_1 \cdots p_{n-1}$ donc (r_n) est une suite décroissante de réels positifs, qui converge nécessairement vers une limite $r \geqslant 0$.
On a donc 0 récurrent si et seulement si $r = \lim_n p_0 p_1 \cdots p_{n-1} = 0$.

La chaîne étant irréductible, tous les états sont de même nature. Elle admet une distribution stationnaire si et seulement si elle est récurrente positive.

On cherche donc à résoudre $\pi = \pi P$ avec $P = \begin{pmatrix} 1-p_0 & p_0 & 0 & \cdots & \cdots \\ 1-p_1 & 0 & p_1 & \ddots & \\ 1-p_2 & 0 & 0 & p_2 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$ et $\displaystyle\sum_{n=0}^{+\infty} \pi_n = 1$.

On obtient, à partir de la deuxième équation, $\pi_1 = p_0 \pi_0$, $\pi_2 = p_1 \pi_1$,...,$\pi_n = p_{n-1}\pi_{n-1}$,..., puis $\pi_2 = p_0 p_1 \pi_0$,...,$\pi_n = p_0 p_1 \cdots p_{n-1} \pi_0 = r_n \pi_0$ et on doit avoir $\pi_0 \left(1 + \displaystyle\sum_{n=1}^{+\infty} r_n\right) = 1$. On a donc

une solution si et seulement si $(\sum r_n)$ converge et cette solution est alors unique.
La chaîne est donc :

- transitoire si $\lim_n p_0p_1\cdots p_{n-1} \neq 0$;
- récurrente nulle si $\lim_n p_0p_1\cdots p_{n-1} = 0$ et $(\sum_n p_0p_1\cdots p_{n-1})$ diverge ;
- récurrente positive si $(\sum_n p_0p_1\cdots p_{n-1})$ converge.

2. Dans le cas où $p_n = p$ pour tout n, on a $r_n = p^n$, avec $p \in]0,1[$, donc la chaîne est récurrente positive, de distribution stationnaire donnée par $\pi_n = p^n\pi_0$ et donc $\pi_0\left(1+\sum_{n=1}^{+\infty}p^n\right) = \dfrac{\pi_0}{1-p} = 1$, soit $\pi_0 = 1-p$ et finalement $\pi_n = (1-p)p^n$ pour tout n.
Le temps moyen de retour à l'état 0 est $\mu_0 = \dfrac{1}{\pi_0} = 2$ si $p = \dfrac{1}{2}$ mais, lorsque le temps de retour à 0 est n, la longueur de la série est $n-1$. Ainsi, si $p = \dfrac{1}{2}$, il y a en moyenne 1 match sans défaite.

Exercice 13.2 : *Gestion d'un stock.*

1. On a $\mathbb{P}([X_{n+1} = k]/[X_n = i]) = \mathbb{P}([s - Y_n = k]/[X_n = i]) = \mathbb{P}([Y_n = s-k]/[X_n = i]) = \dfrac{1}{i+1}$ si $0 \leqslant s-k \leqslant i$, *i.e.* si $s-i \leqslant k \leqslant s$. $\mathbb{P}([X_{n+1} = k]/[X_n = i])$ est donc indépendant de n et l'axiome d'homogénéité est bien vérifié. De plus, la connaissance de $X_0, \cdots, X_{n-1}$ n'apporterait rien de plus pour déterminer la loi de X_{n+1} donc l'axiome de Markov est aussi vérifié et (X_n) est bien une chaîne de Markov de matrice de transition $P = (p_{ij})$ où $p_{ij} = \dfrac{1}{i+1}$ si $s-i \leqslant j \leqslant s$ et 0 sinon.
Pour $\pi_i = \dfrac{2(i+1)}{(s+1)(s+2)}$, on a $\sum_{i=0}^{s}\pi_i = \dfrac{2}{(s+1)(s+2)}\sum_{i=0}^{s}(i+1) = \dfrac{2}{(s+1)(s+2)}\sum_{i=1}^{s+1} i = 1$ et on a aussi :
$(\pi P)_j = \sum_i \pi_i p_{ij} = \sum_{i=s-j}^{s}\dfrac{2(i+1)}{(s+1)(s+2)}\dfrac{1}{i+1} = \dfrac{2}{(s+1)(s+2)}(j+1) = \pi_j$. La chaîne étant irréductible finie, elle admet une unique distribution stationnaire : c'est donc (π_i) où

$$\pi_i = \frac{2(i+1)}{(s+1)(s+2)}.$$

2. $N = \sum_{i=0}^{s} i\pi_i = \dfrac{2}{(s+1)(s+2)}\left[\sum_{i=1}^{s} i(i+1)\right]$ avec $\sum_{i=1}^{s} i^2 = \dfrac{s(s+1)(2s+1)}{6}$ et $\sum_{i=1}^{s} i = \dfrac{s(s+1)}{2}$, donc $\sum_{i=1}^{s} i(i+1) = \dfrac{1}{6}\left[s(s+1)(2s+4)\right] = \dfrac{s(s+1)(s+2)}{3}$, donc $N = \dfrac{2s}{3}$ et pour $s = 30$, $N = 20$.

Exercice 13.3 : *Les urnes d'Ehrenfest.*

1. Si $k = 0$, $X_1 = 1$ et si $k = N$, $X_1 = N-1$. Si $1 \leqslant k \leqslant N-1$, $X_1 = k-1$ si la boule est tirée dans l'urne U_1, ce qui se fait avec la probabilité $\dfrac{k}{N}$ et $X_1 = k+1$ si la boule est tirée

dans l'urne U_2, ce qui se fait avec la probabilité $\dfrac{N-k}{N}$. On a donc $X_1(\Omega)=\{k-1,k+1\}$ avec $\boxed{\mathbb{P}(X_1=k-1)=k/N \text{ et } \mathbb{P}(X_1=k+1)=(N-k)/N}$.

2. Ainsi, si $1\leqslant k\leqslant N-1$, $\mathbb{E}(X_1)=(k-1)\times\dfrac{k}{N}+(k+1)\times\dfrac{N-k}{N}=\dfrac{k^2-k+Nk-k^2+N-k}{N}$

soit $\boxed{\mathbb{E}(X_1)=k\left(1-\dfrac{2}{N}\right)+1}$ et cette relation est encore vraie pour $k=0$ et pour $k=N$.

3. Le passage de X_n à X_{n+1} se fait de la même façon que le passage de X_0 à X_1. On a donc $\mathbb{P}(X_{n+1}=1|X_n=0)=\mathbb{P}(X_{n+1}=N-1|X_n=N)=1$, $\mathbb{P}(X_{n+1}=q-1|X_n=q)=q/N$ et $\mathbb{P}(X_{n+1}=q+1|X_n=q)=(N-q)/N$ si $1\leqslant q\leqslant N-1$, les autres probabilités étant nulles.

4. D'après la question 2, on a $\mathbb{E}^{X_n=k}(X_{n+1})=k\left(1-\dfrac{2}{N}\right)+1$ donc $\boxed{\mathbb{E}^{X_n}(X_{n+1})=X_n\left(1-\dfrac{2}{N}\right)+1}$.
D'après le théorème de l'espérance totale, puis la linéarité de l'espérance, il vient, en prenant l'espérance des deux membres, $\boxed{\mathbb{E}(X_{n+1})=\mathbb{E}(X_n)\left(1-\dfrac{2}{N}\right)+1}$.

5. La suite $(\mathbb{E}(X_n))$ est donc une suite arithmético-géométrique. On cherche un point fixe éventuel ℓ : ℓ vérifie donc $\ell=\ell\left(1-\dfrac{2}{N}\right)+1$, soit $2\ell/N=1$ donc $\ell=N/2$. On a alors $\mathbb{E}(X_{n+1})-\ell=(\mathbb{E}(X_n)-\ell)\left(1-\dfrac{2}{N}\right)$ donc la suite $(\mathbb{E}(X_n)-\ell)$ est géométrique, et $\mathbb{E}(X_n)-\ell=(\mathbb{E}(X_0)-\ell)\left(1-\dfrac{2}{N}\right)^n$, soit $\boxed{\mathbb{E}(X_n)=\dfrac{N}{2}+\left(\mathbb{E}(X_0)-\dfrac{N}{2}\right)\left(1-\dfrac{2}{N}\right)^n}$.

Pour $N\geqslant 2$, $0<\dfrac{2}{N}\leqslant 1$ donc $0\leqslant 1-\dfrac{2}{N}<1$ et $\lim\limits_{n\to+\infty}\left(1-\dfrac{2}{N}\right)^n=0$. On a donc $\boxed{\lim\limits_{n\to+\infty}\mathbb{E}(X_n)=\dfrac{N}{2}}$.

6. On a $P=\begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ q_1 & 0 & p_1 & \ddots & & \vdots \\ 0 & q_2 & 0 & p_2 & & \vdots \\ \vdots & \ddots & & & & 0 \\ \vdots & & & q_{N-1} & 0 & p_{N-1} \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{bmatrix}$ avec $p_k=k/N$ et $q_k=(N-k)/N$.

7. On a un graphe "en boudins" en alignant les états de 0 à N. Tous les états communiquent : $\boxed{\text{la chaîne est donc irréductible}}$.

8. Pour $1\leqslant j\leqslant N-1$,

$$\begin{aligned}\sum_{k=0}^{N}\pi_k P_{k,j} &= \pi_{j-1}P_{j-1,j}+\pi_{j+1}P_{j+1,j}=\frac{1}{2^N}\left[\binom{N}{j-1}\frac{N-j+1}{N}+\binom{N}{j+1}\frac{j+1}{N}\right]\\ &= \frac{1}{2^N}\left[\frac{(N-1)!}{(j-1)!(N-j)!}+\frac{(N-1)!}{j!(N-j-1)!}\right]=\frac{1}{2^N}\binom{N}{j}\left[\frac{j}{N}+\frac{N-j}{N}\right]=\pi_j\end{aligned}$$

On a aussi $\sum_{k=0}^{N} \pi_k P_{k,0} = \pi_1 \frac{1}{N} = \pi_0$ et $\sum_{k=0}^{N} \pi_k P_{k,N} = \pi_{N-1} \frac{1}{N} = \pi_N$. De plus $\sum_{k=0}^{N} \pi_k = 1$ (formule du binôme de Newton) donc $\boxed{\pi \text{ est bien une distribution stationnaire de la chaîne}}$.

9. La chaîne étant irréductible et finie, on sait alors qu'elle est récurrente. Etant donnée que l'on a montré qu'elle admet une distribution stationnaire, on a donc que la chaîne est récurrente positive. De plus, d'après les résultats du cours, la distribution stationnaire π, trouvée à la question précédente, est l'unique distribution stationnaire de la chaîne.

10. Le temps moyen de retour à l'état k est $\boxed{\mu_k = \frac{1}{\pi_k} = \frac{2^N}{\binom{N}{k}}}$. Pour $k = 0$, $\mu_0 = 2^N$ qui tend vers $+\infty$ quand $N \to +\infty$. Pour $N = 10$, on a $\mu_0 = 2^{10}$ et $\mu_5 = \frac{2^{10}}{\binom{10}{5}} = \frac{1024}{252}$, soit $\boxed{\mu_0 = 1024}$ et $\boxed{\mu_5 \approx 4}$.

Exercice 13.4 : *Lancés d'un dé.*

1. Soit Y_i le résultat du i-ème lancer. On a que $X_{n+1} = \max(X_n, Y_{n+1})$ ce qui implique que $(X_n)_n$ est une chaîne de Markov.

2. L'espace des états est $E = \{1, 2, 3, 4, 5, 6\}$ et $P = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 1/3 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 1/2 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 2/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

3. Soit $\pi(0) = (p_1, p_2, p_3, p_4, p_5, p_6)$. On a alors que

$$\pi(0) \lim_{n \to +\infty} P^n = (0, 0, 0, 0, 0, 1)$$

car $\lim_{n \to +\infty} X_n = 6$. D'où :

$$\lim_{n \to +\infty} P^n = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Exercice 13.5 : *Disponibilité de machines.*

1. La connaissance de l'état au n-ième jour suffit pour déterminer les probabilités d'occupation des états au $(n+1)$-ième jour et celles-ci sont indépendantes de n. Ainsi, on a bien une chaîne de Markov telle que :

− si $X_n = 0$, $X_{n+1} = 1$ si les 2 machines tombent en panne le n-ième jour, et $X_{n+1} = 0$ sinon,

− si $X_n = 1$, $X_{n+1} = 1$ si la machine en service tombe en panne le n-ième jour, et $X_{n+1} = 0$ sinon.

En notant $p_{ij} = \mathbb{P}([X_{n+1} = j]/[X_n = i])$, on a $p_{0,1} = q^2$ (les machines fonctionnent indépendamment) ; $p_{0,0} = 1 - q^2$; $p_{1,1} = q$; $p_{1,0} = 1 - q$.

$$P = \begin{pmatrix} 1-q^2 & q^2 \\ 1-q & q \end{pmatrix}.$$

On cherche $\pi = (\pi_0, \pi_1)$ tel que $\pi = \pi P$ et $\pi_0 + \pi_1 = 1$.
Ceci donne $\pi_1 = q^2\pi_0 + q\pi_1$, soit $\pi_1 = \dfrac{q^2}{1-q}\pi_0$ et $\pi_0\left(1 + \dfrac{q^2}{1-q}\right) = 1$. Ainsi, la distribution stationnaire est $\pi = \left(\dfrac{1-q}{1-q+q^2}, \dfrac{q^2}{1-q+q^2}\right)$.

2. Le réparateur ne travaillant plus la nuit, il peut y avoir le matin 0, 1 ou 2 machines en pannes.
− si $X_n = 0$, on peut avoir $X_{n+1} = 0$ (0 panne), $X_{n+1} = 1$ (1 panne) ou $X_{n+1} = 2$ (2 pannes) ;
− si $X_n = 1$, (1 seule machine en service), alors $X_{n+1} = 0$ si elle ne tombe pas en panne le n-ième jour, et $X_{n+1} = 1$ sinon (l'autre remarchera le lendemain) ;
− si $X_n = 2$, alors $X_{n+1} = 1$ car la machine réparée remarchera.
On a alors $P = \begin{pmatrix} (1-q)^2 & 2q(1-q) & q^2 \\ 1-q & q & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

$\pi = \pi P$ et $\pi_0 + \pi_1 + \pi_2 = 1$ donne ici : $\begin{cases} \pi_0 = (1-q)^2\pi_0 + (1-q)\pi_1 \\ \pi_2 = q^2\pi_0 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases}$,
soit $\pi_1 = \dfrac{2q-q^2}{1-q}\pi_0$, $\pi_2 = q^2\pi_0$ et $\pi_0\left(1 + \dfrac{2q-q^2}{1-q} + q^2\right) = 1$, soit $\pi_0 = \dfrac{1-q}{1+q-q^3}$.

Ainsi, la distribution stationnaire est $\pi = \left(\dfrac{1-q}{1+q-q^3}, \dfrac{2q-q^2}{1+q-q^3}, \dfrac{q^2-q^3}{1+q-q^3}\right)$.

12 Processus de Poisson - PC14

Énoncés des exercices

Exercice 14.1 : *Paradoxe des stations de bus*
On suppose que les arrivées de bus à une station forment un processus de Poisson d'intensité λ, ce qui n'est pas réaliste, compte tenu de la régulation imposée à un réseau citadin de bus. Un passager arrive à la station au temps θ : soient U le temps qui sépare θ du temps de la prochaine arrivée et V le temps qui sépare θ de la dernière arrivée si un bus est passé avant θ, θ si aucun bus n'est passé.

1. Déterminer les lois de U et V.
2. Démontrer l'indépendance des v.a. U et V, en calculant $\mathbb{P}((U > y) \cap (V > x))$ dans chacun des deux cas : $0 < x < \theta$ et $x \geqslant \theta$. De l'égalité des lois de U et T, temps aléatoire qui sépare deux arrivées consécutives, on déduit que quelque soit l'instant d'arrivée à une station de bus, le temps d'attente moyen est constant. On constate que $\mathbb{E}(U + V)$ est distinct de $\mathbb{E}(T)$.

Exercice 14.2 : *Traversée d'une route*
Sur une route à sens unique, l'écoulement des voitures est poissonnien d'intensité $\lambda = \dfrac{1}{6}$. Un piéton arrive à l'instant t sur le bord de la route.

1. Quelle est la probabilité pour qu'il attende, sachant qu'il lui faut 4 secondes pour traverser ?
2. Quelle est la durée moyenne des intervalles lui permettant de traverser ?
3. Quel est le nombre moyen de voitures qu'il voit passer avant de traverser ?

Exercice 14.3 : *Nombre d'occurrences sur un intervalle de temps aléatoire*

1. Calculer la loi du nombre aléatoire X d'occurrences d'un processus de Poisson d'intensité λ sur une durée U aléatoire de densité f_U. Application au cas où U est de loi exponentielle.
2. Déterminer la fonction génératrice g_X de X, et en déduire $\mathbb{E}(X)$ et $\mathrm{Var}(X)$.

Exercice 14.4 : *Radar au bord d'une route*
Un radar est placé sur une route où il passe en moyenne 5 véhicules en excès de vitesse par heure. On admet que ces véhicules forment un Processus de Poisson (N_t).

1. Déterminer la probabilité qu'1 voiture ait été prise dans le premier 1/4 d'heure sachant que 2 ont été prises en 1 heure.
2. On suppose ici que la durée de fonctionnement du radar T suit une loi exponentielle de moyenne 100 heures. Déterminer la loi du nombre N de véhicules détectés par le radar et le nombre moyen de ces véhicules. Pour cela :
   - Expliquer la relation $\mathbb{P}^{(T=t)}([N = k]) = \mathbb{P}([N_t = k])$.
   - En utilisant $\mathbb{P}([N = k]) = \displaystyle\int \mathbb{P}^{(T=t)}([N = k]) f_T(t)\, dt$, établir que N suit une loi géométrique sur $\mathbb{N}$.

Corrigés des exercices

Exercice 14.1 : *Paradoxe des stations de bus*

1. Déterminons tout d'abord la loi de V. Il est clair que $\forall x > \theta$, $P([V > x]) = 0$. De plus, $\forall x \in [0, \theta[$, l'événement $[V > x]$ signifie que pendant la durée x précédant θ, il n'y a aucune arrivée :

$$P([V > x]) = P([N_\theta - N_{\theta - x} = 0]) = P([N_x = 0]) = e^{-\lambda x}$$

pour $x \in [0, \theta[$, 0 si $x \geqslant \theta$, où N_x est un processus de Poisson d'intensité λ. D'où une loi de V, de forme mixte, définie par :

$$F_V(x) = (1 - e^{-\lambda x})\mathbf{1}_{[0,\theta[}(x) + \mathbf{1}_{[\theta,+\infty[}(x).$$

Déterminons maintenant la loi de U. D'après l'énoncé, on a que :

$$P([U > x]) = P([N_{\theta + x} - N_\theta = 0]) = P([N_x = 0]) = e^{-\lambda x}$$

car $[U > x]$ signifie que pendant la durée x qui suit θ, il n'y a aucune arrivée.

2. Montrons l'indépendance de U et V en distinguant deux cas :
- si $0 < x < \theta$ et $y > 0$: $[V > x] \cap [U > y]$ signifie qu'il n'y a aucune arrivée entre $\theta - x$ et $\theta + y$ donc

$$P([V > x] \cap [U > y]) = e^{-\lambda(x+y)} = e^{-\lambda x} e^{-\lambda y} = P([V > x])P([U > y]).$$

- si $x \geqslant \theta$ et $y > 0$: $P([V > x] \cap [U > y]) = 0 = P([V > x])P([U > y])$.

$$\mathbb{E}(U) = \frac{1}{\lambda}\,;\ \mathbb{E}(V) = \int_0^\theta P([V > s])\,ds = \frac{1}{\lambda}(1 - e^{-\lambda\theta}) \text{ donc :}$$

$$\mathbb{E}(U + V) = \frac{2}{\lambda} - \frac{1}{\lambda}\,e^{-\lambda\theta} > \frac{1}{\lambda} = \mathbb{E}(T).$$

Exercice 14.2 : *Traversée d'une route*

Le piéton arrive au bord de la route à un instant θ. On note U la v.a.r. qui sépare θ de la prochaine arrivée de voiture.

1. Le piéton ayant besoin d'au moins 4s pour traverser, il devra attendre si $U \leqslant 4$.

Or $P([U \leqslant 4]) = \displaystyle\int_0^4 \lambda e^{-\lambda u}\,du = \Big[-e^{-\lambda u}\Big]_0^4 = 1 - e^{-4\lambda}$ car on a vu à l'exercice 14.1 que U suit la loi exponentielle $\mathcal{E}(\lambda)$. Comme $\lambda = 1/6$, $P([U \leqslant 4]) = 1 - e^{-4/6}$.
La probabilité pour qu'il doive attendre est donc $1 - e^{-2/3} \approx 0,487$.

2. On cherche $\mathbb{E}^{[U>4]}(U)$ c'est-à-dire $\displaystyle\int u f_U^{[U>4]}(u)\,du$. On a

$$\begin{aligned} F_U^{[U>4]}(u) &= P^{[U>4]}([U \leqslant u]) = \frac{P([U \leqslant u] \cap [U > 4])}{P([U > 4])} \\ &= \begin{cases} 0 & \text{si } u \leqslant 4 \\ \dfrac{P([4 < U \leqslant u])}{P([U > 4])} = \dfrac{F_U(u) - F_U(4)}{P([U > 4])} & \text{si } u > 4 \end{cases} \end{aligned}$$

et donc $f_U^{[U>4]}(u) = \dfrac{f_U(u)}{P([U>4])} \mathbf{1}_{]4,+\infty[}(u) = \lambda\, e^{-\lambda(u-4)}\, \mathbf{1}_{]4,+\infty[}(u)$ puis

$$\begin{aligned}
\mathbb{E}^{[U>4]}(U) &= \int_4^{+\infty} \lambda u\, e^{-\lambda(u-4)}\, du \underset{v=u-4}{=} \int_0^{+\infty} \lambda(v+4)\, e^{-\lambda v}\, dv \\
&= \int_0^{+\infty} v\lambda\, e^{-\lambda v}\, dv + 4\int_0^{+\infty} \lambda\, e^{-\lambda v}\, dv = \frac{1}{\lambda} + 4
\end{aligned}$$

car $\displaystyle\int_0^{+\infty} v\lambda\, e^{-\lambda v}\, dv$ est l'espérance d'une v.a.r. de loi exponentielle $\mathcal{E}(\lambda)$ et $\displaystyle\int_0^{+\infty} \lambda\, e^{-\lambda v}\, dv$ l'intégrale de sa densité.

Ainsi, il faut en moyenne $6+4=10$s pour que le piéton puisse traverser.

3. Soit N le nombre moyen de voitures que voit passer le piéton avant de pouvoir traverser la route. On a

– $[N=0]=[U>4]$, où U est le temps entre l'arrivée du piéton et la prochaine voiture ;

– $[N=1]=[U\leqslant 4]\cap[U_1>4]$ où U_1 est le temps entre la première et la deuxième voiture qui se présentent après l'arrivée du piéton.

– Plus généralement, $[N=k]=[U\leqslant 4]\cap[U_1\leqslant 4]\cdots\cap[U_{k-1}\leqslant 4]\cap[U_k>4]$ où U_k est le temps entre la k-ième et la $(k+1)$-ième voiture qui se présentent après l'arrivée du piéton.

Les v.a.r. U, $U_1, \cdots, U_k$ étant toutes indépendantes et de même loi exponentielle $\mathcal{E}(\lambda)$ on a $P([N=0]) = P([U>4]) = e^{-4\lambda}$ et

$$\begin{aligned}
P([N=k]) &= P([U\leqslant 4])P([U_1\leqslant 4])\cdots P([U_{k-1}\leqslant 4])P([U_k>4]) \\
&= e^{-4\lambda}(1-e^{-4\lambda})^k = e^{-2/3}(1-e^{-2/3})^k
\end{aligned}$$

(valable aussi pour $k=0$). On reconnait la loi géométrique $\mathcal{G}_0(e^{-2/3})$ sur $\mathbb{N}$ définie par : $P(X=k)=p(1-p)^k$.

La fonction génératrice de N est :

$$G_N(s) = \sum_{k=0}^{+\infty} s^k\, P([N=k]) = e^{-2/3}\sum_{k=0}^{+\infty}(s(1-e^{-2/3}))^k$$

soit $G_N(s) = \dfrac{e^{-2/3}}{1-s(1-e^{-2/3})}$ et $G_N'(s) = \dfrac{e^{-2/3}(1-e^{-2/3})}{(1-s(1-e^{-2/3}))^2}$.

On a alors $\mathbb{E}(N) = G_N'(1) = \dfrac{1-e^{-2/3}}{e^{-2/3}} = e^{2/3}-1 \approx 0,948$.

Le piéton voit donc passer en moyenne 1 voiture avant de traverser.

<u>Exercice 14.3</u> : *Nombre d'occurrences sur un intervalle de temps aléatoire*

1. $P(X=k) = \displaystyle\int_0^{+\infty} P(X=k \mid U=u) f_U(u)\, du = \int_0^{+\infty} \frac{e^{-\lambda u}(\lambda u)^k}{k!} f_U(u)\, du.$

Si U est de loi exponentielle de paramètre α,

$$\begin{aligned}
P([X=k]) &= \int_0^{+\infty} \frac{e^{-\lambda u}(\lambda u)^k}{k!}\alpha e^{-\alpha u}\, du = \frac{\lambda^k}{k!}\alpha \int_0^{+\infty} e^{-(\alpha+\lambda)u} u^k\, du \\
&= \frac{\lambda^k}{k!}\alpha \frac{1}{(\alpha+\lambda)^{k+1}}\,\Gamma(k+1) = \left(\frac{\lambda}{\alpha+\lambda}\right)^k \left(\frac{\alpha}{\alpha+\lambda}\right)
\end{aligned}$$

Conclusion : X est une variable aléatoire de loi géométrique $\mathcal{G}_0\left(\dfrac{\alpha}{\alpha+\lambda}\right)$.

2. Par définition, la la fonction génératrice de X est

$$g_X(s) = \sum_{k=0}^{+\infty} s^k P([X=k]) = \sum_{k=0}^{+\infty} \frac{e^{-\lambda u}(\lambda us)^k}{k!} f_U(u)\,du = \int_0^{+\infty} e^{\lambda u(s-1)}\,f_U(u)\,du.$$

Ainsi, $g'_X(s) = \displaystyle\int_0^{+\infty} \lambda u e^{\lambda u(s-1)}\,f_U(u)\,du$ et $g''_X(s) = \displaystyle\int_0^{+\infty} \lambda^2 u^2 e^{\lambda u(s-1)}\,f_U(u)\,du$ et, pour $s=1$, on obtient $\mathbb{E}(X) = g'_X(1) = \lambda\mathbb{E}(U)$ et $g''_X(1) = \lambda^2\mathbb{E}(U^2)$ donc $\mathrm{var}(X) = g''_X(1) + g'_X(1) - g'_X(1)^2$, soit $\mathrm{var}(X) = \lambda^2\mathrm{var}(U) + \lambda\mathbb{E}(U)$.

Si U est de loi exponentielle de paramètre α, on a alors $\mathbb{E}(U) = \dfrac{1}{\alpha}$ et $\mathrm{var}(U) = \dfrac{1}{\alpha^2}$, donc $\mathbb{E}(X) = \dfrac{\lambda}{\alpha}$ et $\mathrm{var}(X) = \dfrac{\lambda^2}{\alpha^2} + \dfrac{\lambda}{\alpha}$.

<u>Exercice 14.4</u> : *Radar au bord d'une route*

1. En utilisant successivement les propriétés d'un processus de Poisson (indépendance des accroissements, stationnarité, et N_t de loi $\mathcal{P}(\lambda t)$), il vient :

$$\begin{aligned}
P(N_{1/4}=1 \mid N_1=2) &= \frac{P(N_{1/4}=1,\, N_1=2)}{P(N_1=2)} = \frac{P(N_{1/4}=1,\, N_1-N_{1/4}=1)}{P(N_1=2)} \\
&= \frac{P(N_{1/4}=1)P(N_1-N_{1/4}=1)}{P(N_1=2)} = \frac{P(N_{1/4}=1)P(N_{3/4}=1)}{P(N_1=2)} \\
&= \frac{e^{-\frac{\lambda}{4}}\ \frac{\lambda}{4} \times e^{-\frac{3\lambda}{4}}\ \frac{3\lambda}{4}}{e^{-\lambda}\ \frac{\lambda^2}{2!}} = \frac{3\times 2}{4\times 4} = \frac{3}{8}.
\end{aligned}$$

2. Si $T=t$ le radar aura enregistré les véhicules de 0 à t et leur nombre est exactement N_t. Ainsi, $P_N^{T=t} = P_{N_t}$. Avec $f_T(t) = \dfrac{1}{100}e^{-\frac{t}{100}}\mathbf{1}_{]0,+\infty[}(t)$, on a alors

$$\begin{aligned}
P([N=n]) &= \int P^{T=t}([N=n])f_T(t)dt = \int P([N_t=n])f_T(t)dt \\
&= \int_0^{+\infty} e^{-\lambda t}\ \frac{(\lambda t)^n}{n!}\ \frac{1}{100}e^{-\frac{t}{100}}\ dt = \frac{\lambda^n}{100n!}\int_0^{+\infty} t^n\ e^{-(\lambda+\frac{1}{100})t}\ dt
\end{aligned}$$

et en posant $u = \left(\lambda + \dfrac{1}{100}\right)t$, on a $\displaystyle\int_0^{+\infty} t^n\ e^{-(\lambda+\frac{1}{100})t}\ dt = \frac{\Gamma(n+1)}{(\lambda+\frac{1}{100})^{n+1}} = \frac{100^{n+1}\ n!}{(100\lambda+1)^{n+1}}$ donc $P([N=n]) = \dfrac{(100\lambda)^n}{(100\lambda+1)^{n+1}} = \dfrac{1}{100\lambda+1}\left(\dfrac{100\lambda}{100\lambda+1}\right)^n$ de la forme pq^n. Ainsi, avec $\lambda=5$, N suit la loi géométrique $\mathcal{G}_0\left(\dfrac{1}{501}\right)$, et $\mathbb{E}(N) = \dfrac{q}{p} = 500$.