PC 1 : Probabilités discrètes

Les exercices 1 et 6 sont corrigés pour vous donner un exemple de rédaction. Les exercices 4, 5 et 7 seront particulièrement discutés en PC.

1 Événements, probabilités et indépendance

Exercice 1 (Événements indépendants)**.** Soit $\Omega := \{\omega_1, \omega_2, \omega_3, \omega_4\}$ un ensemble à quatre éléments, muni de la probabilité uniforme $\mathbb{P}$. On définit les événements $A := \{\omega_1, \omega_2\}$, $B := \{\omega_1, \omega_3\}$ et $C := \{\omega_2, \omega_3\}$. Montrer que A, B et C sont indépendants deux à deux. Comparer $\mathbb{P}(A \cap B \cap C)$ et $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

Solution. Pour chaque $i \in \{1, \dots, 4\}$, on a par définition $\mathbb{P}(\{w_i\}) = 1/|\Omega| = 1/4$. D'une part, $\mathbb{P}(A) = \mathbb{P}(\{\omega_1\} \cup \{\omega_2\}) = \mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) = 1/2$ et, de même, $\mathbb{P}(B) = \mathbb{P}(C) = 1/2$. D'autre part, $A \cap B = \{\omega_1\}$ et donc $\mathbb{P}(A \cap B) = 1/4$. Donc, on a montré que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, d'où l'indépendance de A et B. Les événements A et C sont indépendants pour la même raison, de même que B et C sont indépendants.
Comme $A \cap B \cap C = \emptyset$, on a $\mathbb{P}(A \cap B \cap C) = 0$. En revanche, $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = 1/8$. Cela implique que A, B et C ne sont pas mutuellement indépendants.

Exercice 2. Soit Ω un ensemble fini ou dénombrable et soit $\mathbb{P}$ une mesure de probabilité quelconque sur Ω. On fixe deux événements A et B.

1. Supposons que $\mathbb{P}(A) = \frac{3}{4}$ et $\mathbb{P}(B) = \frac{1}{3}$, montrer que $\frac{1}{12} \leq \mathbb{P}(A \cap B) \leq \frac{1}{3}$. Exhiber des exemples qui montrent que les deux bornes peuvent être atteintes.
2. Montrer que si $A \cup B = \Omega$, alors

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A^c)\mathbb{P}(B^c).$$

Solution.

1. On sait que $\mathbb{P}(A \cup B) = \mathbb{P}(A) + P(B) - \mathbb{P}(A \cap B)$ et $\mathbb{P}(A \cup B) \leq 1$, d'où, comme par ailleurs $A \cap B \subset B$,

$$\frac{1}{3} = \mathbb{P}(B) \geq \mathbb{P}(A \cap B) = \mathbb{P}(A) + P(B) - \mathbb{P}(A \cup B) \geq \mathbb{P}(A) + P(B) - 1 = \frac{3}{4} + \frac{1}{3} - 1 = \frac{1}{12}.$$

   Prenons $\Omega = \{1, \dots, 12\}$ et $\mathbb{P}$ la loi uniforme ; pour $A = \{1, \dots, 9\}$ et $B = \{9, \dots, 12\}$ on a $A \cap B = \{9\}$, d'où $\mathbb{P}(A) = \frac{3}{4}$, $\mathbb{P}(B) = \frac{1}{3}$ et $\mathbb{P}(A \cap B) = \frac{1}{12}$, tandis que pour le même $A = \{1, \dots, 9\}$ mais $B = \{1, \dots, 4\}$ on a $A \cap B = B$, d'où $\mathbb{P}(A) = \frac{3}{4}$, $\mathbb{P}(B) = \frac{1}{3} = \mathbb{P}(A \cap B)$.

2. Si $\mathbb{P}(A \cup B) = 1$, alors

$$\begin{aligned}
\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(A^c)\mathbb{P}(B^c) &= \mathbb{P}(A)\mathbb{P}(B) - (1 - \mathbb{P}(A))(1 - \mathbb{P}(B)) \\
&= \mathbb{P}(A)\mathbb{P}(B) - (1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B)) \\
&= -1 + \mathbb{P}(A) + \mathbb{P}(B) \\
&= -\mathbb{P}(A \cup B) + \mathbb{P}(A) + \mathbb{P}(B) \\
&= \mathbb{P}(A \cap B).
\end{aligned}$$

Exercice 3 (Conditionnement)**.** L'exercice suivant est très classique et a de nombreuses variantes. Il illustre l'utilité d'une formulation mathématique rigoureuse pour éviter des pièges et des paradoxes dus à des raisonnements spécieux.
Une famille a deux enfants. On suppose les 4 configurations (ω_1, ω_2) avec ω_i le sexe du $i$ème enfant équiprobables.

1. Montrer que la probabilité pour que les deux enfants soient des filles sachant que le plus jeune enfant est une fille vaut $\frac{1}{2}$.
2. Montrer que la probabilité pour que les deux enfants soient des filles sachant que l'enfant plus âgé est une fille vaut $\frac{1}{2}$.
3. Montrer que la probabilité pour que les deux enfants soient des filles sachant que l'un des enfants est une fille vaut $\frac{1}{3}$.

Solution.

1. Ici $\Omega = \{(F,F),(G,G),(F,G),(G,F)\}$. L'événement que le plus jeune enfant est une fille correspond à l'ensemble $A \coloneqq \{(F,F),(G,F)\}$ qui se produit avec probabilité $\mathbb{P}(A) = |A|/|\Omega| = 1/2$. Pour la probabilité que les deux enfants soient des filles sachant que le plus jeune enfant est une fille on obtient alors
$$\mathbb{P}(\{(F,F)\} \mid A) = \frac{\mathbb{P}(\{(F,F)\} \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\{(F,F)\})}{\mathbb{P}(A)} = \frac{1/4}{1/2} = \frac{1}{2}.$$
2. Par le même raisonnement, en notant $B \coloneqq \{(F,F),(F,G)\}$ l'événement que l'enfant plus âgé est une fille avec $\mathbb{P}(B) = |B|/|\Omega| = 1/2$, on obtient
$$\mathbb{P}(\{(F,F)\} \mid B) = \frac{\mathbb{P}(\{(F,F)\} \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{(F,F)\})}{\mathbb{P}(B)} = \frac{1/4}{1/2} = \frac{1}{2}.$$
3. Notons $C \coloneqq \{(F,F),(F,G),(G,F)\}$ l'événement que l'un des enfants est une fille. On a $\mathbb{P}(C) = |C|/|\Omega| = 3/4$ et donc
$$\mathbb{P}(\{(F,F)\} \mid C) = \frac{\mathbb{P}(\{(F,F)\})}{\mathbb{P}(C)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

2 Borel-Cantelli

Exercice 4 (Limite supérieure d'ensembles)**.** Soit $(A_n)_{n\geq 1}$ une suite d'événements.

1. Que représentent les événements
$$\bigcap_{k\geq 1}\bigcup_{n\geq k} A_n \qquad \text{et} \qquad \bigcup_{k\geq 1}\bigcap_{n\geq k} A_n$$
respectivement ? Le premier est noté $\boldsymbol{\limsup_{n\to\infty} A_n}$.
2. Si l'espace Ω est $\mathbb{R}$, donner $\limsup_{n\to\infty} A_n$ dans les trois cas suivants :
   (a) $A_n = [-1/n, 3+1/n]$,
   (b) $A_n = [-2-(-1)^n, 2+(-1)^n[$,
   (c) $A_n = p_n\mathbb{N}$, où $(p_n)_{n\geq 1}$ est la suite ordonnée des nombres premiers et $p_n\mathbb{N} = \{0, p_n, 2p_n, \dots\}$ est l'ensemble des multiples de p_n.
3. Comparer les événements $\limsup_{n\to\infty}(A_n \cup B_n)$ et $\limsup_{n\to\infty}(A_n \cap B_n)$ respectivement avec $\limsup_{n\to\infty} A_n$ et $\limsup_{n\to\infty} B_n$.

Solution. 1. On a

$$\omega \in \bigcap_{k\geq 1}\bigcup_{n\geq k} A_n \iff \text{il existe une infinité d'indices } n \text{ tels que } \omega \in A_n,$$

tandis que

$$\omega \in \bigcup_{k\geq 1}\bigcap_{n\geq k} A_n \iff \text{pour tout indice } n \text{ à partir d'un certain rang } \omega \in A_n.$$

2. (a) La suite $(A_n)_{n\geq 1}$ étant monotone décroissante ($A_n \supset A_{n+1}$ pour tout $n \geq 1$), on a pour tout $k \geq 1$, $\bigcup_{n\geq k} A_n = A_k = [-1/k, 3+1/k]$. D'une part, on voit que $[0,3] \subset A_k \subset \bigcup_{n\geq k} A_n$ pour tout k ; d'autre part, pour tout $s < 0$ et pour tout $t > 3$ il existe k tel que pour tout $n \geq k$, on a $s < -1/n$ et $t > 3 + 1/n$ donc ni s ni t n'appartiennent à $\bigcup_{n\geq k} A_n$. Donc, $\limsup_{n\to\infty} A_n = [0,3]$.

   (b) Pour tout n pair, on a $A_n = [-3,3[$. Pour tout n impair, on a $A_n = [-1,1[$. Donc, pour tout $k \geq 1$, $\bigcup_{n\geq k} A_n = [-3,3[$, et ainsi $\limsup_{n\to\infty} A_n = [-3,3[$.

   (c) On a $A_1 = 2\mathbb{N} = \{0,2,4,6,...\}$, $A_2 = 3\mathbb{N} = \{0,3,6,9,...\}$, et plus généralement, pour tout $n \geq 1$, on a $A_n = p_n\mathbb{N} = \{0, p_n, 2p_n, 3p_n, ...\}$. Clairement, $A_n \subset \mathbb{N}$ pour tout n, donc $\limsup_{n\to\infty} A_n \subset \mathbb{N}$. On voit que $0 \in A_n$ pour tout n, donc $0 \in \limsup_{n\to\infty} A_n$. Enfin, pour tout $j \in \mathbb{N}^*$ et pour tout k tel que $j < p_k$, on a $j \notin A_n$ pour tout $n \geq k$. Donc $\limsup_{n\to\infty} A_n = \{0\}$.

3. D'une part,

$$\begin{aligned}
\omega \in \limsup_{n\to\infty}(A_n \cup B_n) &\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \cup B_n \\
&\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \textbf{ ou } \omega \in B_n \\
&\iff \omega \in \limsup_{n\to\infty} A_n \textbf{ ou } \omega \in \limsup_{n\to\infty} B_n \\
&\iff \omega \in \left(\limsup_{n\to\infty} A_n\right) \cup \left(\limsup_{n\to\infty} B_n\right).
\end{aligned}$$

D'autre part,

$$\begin{aligned}
\omega \in \limsup_{n\to\infty}(A_n \cap B_n) &\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \cap B_n \\
&\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \textbf{ et } \omega \in B_n \\
&\implies \omega \in \left(\limsup_{n\to\infty} A_n\right) \cap \left(\limsup_{n\to\infty} B_n\right).
\end{aligned}$$

On a montré que

$$\begin{aligned}
\limsup_{n\to\infty}(A_n \cup B_n) &= \left(\limsup_{n\to\infty} A_n\right) \cup \left(\limsup_{n\to\infty} B_n\right) \\
\limsup_{n\to\infty}(A_n \cap B_n) &\subset \left(\limsup_{n\to\infty} A_n\right) \cap \left(\limsup_{n\to\infty} B_n\right).
\end{aligned}$$

La dernière inclusion peut être stricte. Par exemple, pour $A_{2\ell} = B_{2\ell+1} = \{0\}$ et $A_{2\ell+1} = B_{2\ell} = \{1\}$ pour tout ℓ, on a $\limsup_{n\to\infty} A_n = \limsup_{n\to\infty} B_n = \{0,1\}$, alors que $\limsup_{n\to\infty}(A_n \cap B_n) = \emptyset$.

Exercice 5 (Retours en zéro d'une marche aléatoire)**.** Soit $p \in]0,1[$ et soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi donnée par :

$$\mathbb{P}(X_1 = 1) = p \qquad \text{et} \qquad \mathbb{P}(X_1 = -1) = 1 - p.$$

On note $S_n = X_1 + \cdots + X_n$, $n \geq 1$. Montrer que si $p \neq \frac{1}{2}$, alors avec probabilité 1 la suite $(S_n)_{n\geq 1}$ ne prend la valeur 0 qu'un nombre fini de fois. Peut-on conclure aussi facilement lorsque $p = \frac{1}{2}$?
On pourra faire appel à la formule de Stirling : $n! \sim \sqrt{2\pi n}\,(n/\mathrm{e})^n$ lorsque $n \to \infty$.

Solution. Comme la marche est issue de 0, l'événement $\{S_n = 0\}$ est impossible si n est impair. Pour $n = 2k$, on a $S_{2k} = 0$ si et seulement si exactement k variables parmi $X_1, \ldots, X_{2k}$ prennent la valeur 1, d'où

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

En appliquant la formule de Stirling, on voit que lorsque $k \to +\infty$,

$$\binom{2k}{k} = \frac{(2k)!}{(k!)^2} \sim \frac{(2k)^{2k}\sqrt{4\pi k}\,\mathrm{e}^{-2k}}{k^{2k} 2\pi k\,\mathrm{e}^{-2k}} = \frac{1}{\sqrt{\pi k}} 4^k.$$

Par conséquent,

$$\mathbb{P}(S_{2k} = 0) \sim \frac{1}{\sqrt{\pi k}} \left(4p(1-p)\right)^k.$$

Comme $p \neq 1/2$, $4p(1-p) < 1$ et donc $\sum_{k\geq 1} \mathbb{P}(S_{2k} = 0) < +\infty$. Le lemme de Borel-Cantelli entraîne donc que $\mathbb{P}(\limsup\{S_{2k} = 0\}) = 0$. Autrement dit la marche ne repasse par 0 qu'un nombre fini de fois avec probabilité 1.
Si $p = 1/2$, on a $\mathbb{P}(S_{2k} = 0) \sim 1/\sqrt{\pi k}$ et donc $\sum_{k=1}^{+\infty} \mathbb{P}(S_{2k} = 0) = +\infty$. En revanche les événements $\{S_{2k} = 0\}$ n'étant pas indépendants, on ne peut pas utiliser le deuxième lemme de Borel-Cantelli qui nous conduirait à la conclusion que $\mathbb{P}(\limsup\{S_{2k} = 0\}) = 1$, c'est-à-dire au fait que la marche aléatoire repasse par 0 une infinité de fois avec probabilité 1.

Remarque : cette conclusion est cependant vraie ! Et ce dès lors que la série $\sum_{k=1}^{+\infty} \mathbb{P}(S_k = 0)$ n'est pas sommable ; ce n'est pas si difficile à montrer mais nous préférons nous concentrer sur autre chose.

3 Variables aléatoires

Exercice 6 (Une autre formule pour l'espérance). Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

1. Montrer que l'espérance de X est finie si et seulement si la série $\sum_n \mathbb{P}(X > n)$ est convergente, et dans ce cas on a

$$\mathbb{E}[X] = \sum_{n\geq 0} \mathbb{P}(X > n).$$

Application : on considère une urne avec N boules numérotées de 1 à N, on tire successivement n boules avec remise et on note $X_N^{(n)}$ le plus grand numéro sorti.

2. Calculer $\mathbb{P}(X_N^{(n)} \leq k)$ pour tout $k \geq 1$ et en déduire la valeur de $\mathbb{E}[X_N^{(n)}]$.
3. L'entier $n \geq 1$ étant fixé, montrer que la suite $(N^{-1}\mathbb{E}[X_N^{(n)}])_{N\geq 1}$ converge et calculer sa limite.

Solution. 1. C'est une simple application du théorème de Fubini pour les séries à termes positifs : comme $X \geq 0$, on peut toujours définir $\mathbb{E}[X] \in [0, +\infty]$, et alors en remarquant que $k = \sum_{n=0}^{k-1} 1$ pour tout $k \geq 0$ (une somme vide étant nulle), on a en échangeant deux séries à termes positifs :

$$\mathbb{E}[X] = \sum_{k\geq 0} k\,\mathbb{P}(X = k) = \sum_{k\geq 0}\sum_{n=0}^{k-1} \mathbb{P}(X = k) = \sum_{n\geq 0}\sum_{k>n} \mathbb{P}(X = k) = \sum_{n\geq 0} \mathbb{P}(X > n),$$

où la dernière égalité provient du fait que l'événement $\{X > n\}$ est la réunion disjointe des $\{X = k\}$ pour $k > n$.

2. On a $\{X_N^{(n)} \leq k\}$ si et seulement si les n boules, qui sont tirées indépendamment et uniformément au hasard, sont toutes plus petites que k, de sorte que

$$\mathbb{P}(X_N^{(n)} \leq k) = \left(\frac{k}{N}\right)^n \quad \text{pour tout } 1 \leq k \leq N.$$

On en déduit que

$$\mathbb{E}[X_N^{(n)}] = \sum_{k=1}^{N} 1 - \left(\frac{k}{N}\right)^n = N - \sum_{k=1}^{N} \left(\frac{k}{N}\right)^n.$$

3. Pour tout $N \geq 1$, on a

$$\frac{\mathbb{E}[X_N^{(n)}]}{N} = 1 - \frac{1}{N}\sum_{k=1}^{N} \left(\frac{k}{N}\right)^n \underset{n\to\infty}{\longrightarrow} 1 - \int_0^1 x^n \mathrm{d}x = 1 - \frac{1}{n+1} = \frac{n}{n+1}.$$

Exercice 7 (Loi géométrique)**.** On modélise le jeu de pile ou face par une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p \in]0,1[$, en codant 1 pour succès (pile) et 0 pour échec (face) : $\mathbb{P}(X_1 = 1) = 1 - \mathbb{P}(X_1 = 0) = p$.

1. On pose $T_1 = \inf\{n \geq 1 : X_n = 1\}$. Que représente la variable aléatoire T_1, qu'elle est sa loi, sa moyenne, sa variance ?
2. Soit $k \geq 2$; on s'intéresse à la variable T_k définie par $T_k = \inf\{n \geq 1 : \sum_{i=1}^n X_i = k\}$ représentant l'instant où le joueur réalise son $k$ème succès. Déterminer la loi de T_k.
3. Posons $T_0 = 0$ et $\Delta_k = T_k - T_{k-1}$ pour $k \geq 1$. Montrer que les variables aléatoires Δ_k sont indépendantes et de même loi.

Solution. 1. La variable aléatoire T_1 correspond à la date du premier succès. Déterminer la loi de T_1 consiste à donner $\mathbb{P}(T_1 = k)$ pour tout $k \geq 1$. On remarque que

$$\{T_1 = k\} = \{X_1 = 0, \ldots, X_{k-1} = 0, X_k = 1\} = \left(\bigcap_{i=1}^{k-1}\{X_i = 0\}\right) \cap \{X_k = 1\}$$

et donc par indépendance et équidistribution des variables X_i on trouve

$$\mathbb{P}(T_1 = k) = (1-p)^{k-1}p.$$

Autrement dit, la variable T_1 suit la loi géométrique de paramètre p. Pour l'espérance de T_1, on trouve

$$\mathbb{E}[T_1] = \sum_{k=1}^{+\infty} k\mathbb{P}(T_1 = k) = \sum_{k=1}^{+\infty} k(1-p)^{k-1}p$$

donc en notant $f(x) = \sum_{k=0}^{+\infty}(1-x)^k = x^{-1}$, $x \in]0,1]$, on voit que

$$\mathbb{E}[T_1] = -pf'(p) = \frac{1}{p}.$$

De même, on a

$$\mathbb{E}[T_1(T_1-1)] = \sum_{k=1}^{+\infty} k(k-1)\mathbb{P}(T_1 = k) = \sum_{k=1}^{+\infty} k(k-1)(1-p)^{k-1}p,$$

autrement dit,

$$\mathbb{E}[T_1(T_1-1)] = p(1-p)f''(p) = 2\frac{1-p}{p^2},$$

et ainsi, par linéarité de l'espérance,

$$\mathbb{E}[T_1^2] = \mathbb{E}[T_1(T_1-1)] + \mathbb{E}[T_1] = 2\frac{1-p}{p^2} + \frac{1}{p} = \frac{2-p}{p^2};$$

finalement on a

$$\mathrm{Var}(T_1) = \mathbb{E}[T_1^2] - \mathbb{E}[T_1]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

2. Déterminons la loi de T_k. Remarquons tout d'abord que $\mathbb{P}(T_k = n) = 0$ si $n < k$. Pour $n \geq k$, on a

$$\{T_k = n\} = \left\{ \sum_{i=1}^{n-1} X_i = k-1, X_n = 1 \right\}.$$

Par indépendance, on a donc

$$\mathbb{P}(T_k = n) = \mathbb{P}\left(\sum_{i=1}^{n-1} X_i = k-1 \right) \mathbb{P}(X_n = 1) = p\,\mathbb{P}\left(\sum_{i=1}^{n-1} X_i = k-1 \right).$$

Par ailleurs, $\sum_{i=1}^{n-1} X_i$ suit une loi binomiale de paramètres $n-1$ et p :

$$\mathbb{P}\left(\sum_{i=1}^{n-1} X_i = j \right) = \binom{n-1}{j} p^j (1-p)^{n-1-j}, \qquad \forall j \in \{0, 1, \ldots, n-1\}.$$

D'où la formule

$$\mathbb{P}(T_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}, \qquad \forall n \geq k.$$

3. Afin de montrer l'indépendance des variables Δ_k, remarquons que pour toute suite d'entiers $n_1, \ldots, n_{k+1} \geq 1$, on a

$$\begin{aligned}
&\{\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_{k+1} = n_{k+1}\} \\
&= \{T_1 - T_0 = n_1, T_2 - T_1 = n_2, \ldots, T_{k+1} - T_k = n_{k+1}\} \\
&= \{T_1 - T_0 = n_1, T_2 - T_1 = n_2, \ldots, T_k - T_{k-1} = n_k\} \cap \bigcap_{j=1}^{n_{k+1}-1} \{X_{n_1+\cdots+n_k+j} = 0\} \\
&\qquad \cap \{X_{n_1+\cdots+n_{k+1}} = 1\}.
\end{aligned}$$

Comme l'événement $\{T_1 - T_0 = n_1, T_2 - T_1 = n_2, \ldots, T_k - T_{k-1} = n_k\}$ s'exprime en fonction des variables $X_1, \ldots, X_{n_1+\cdots+n_k}$ uniquement, on en déduit qu'il est indépendant du reste, et ainsi

$$\begin{aligned}
&\mathbb{P}(\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_{k+1} = n_{k+1}) \\
&\quad = \mathbb{P}(\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_k = n_k)(1-p)^{n_{k+1}-1} p
\end{aligned}$$

et par une récurrence immédiate,

$$\mathbb{P}(\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_{k+1} = n_{k+1}) = \prod_{j=1}^{k+1} (1-p)^{n_j - 1} p.$$

Comme $(1-p)^{n_j-1}p$ est la probabilité qu'une variable aléatoire de loi géométrique de paramètre p prenne la valeur n_j, on déduit immédiatement de la forme produit que les variables Δ_k sont indépendantes et suivent toutes cette loi géométrique de paramètre p.

Exercice 8 (Espérance conditionnelle)**.** Soit X_1, X_2 des variables aléatoires indépendantes de loi de Poisson de paramètre respectif $\theta_1 > 0$ et $\theta_2 > 0$.

1. Calculer la loi $\mathbb{P}(X_1 + X_2 = k)$ pour tout $k \geq 0$; quelle est la loi de $X_1 + X_2$?
2. Calculer $\mathbb{E}(z^{X_1} \mid X_1 + X_2)$ pour $z > 0$; quelle est la loi conditionnelle de X_1 sachant $X_1 + X_2$?
3. Calculer $\mathbb{E}(X_1 \mid X_1 + X_2)$.

Solution. 1. Pour $k \in \mathbb{N}$ on a par la formule des probabilités totales

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = k) &= \sum_{\ell=0}^{\infty} \mathbb{P}(X_1 + X_2 = k \mid X_2 = \ell)\mathbb{P}(X_2 = \ell) \\
&= \sum_{\ell=0}^{k} \mathbb{P}(X_1 = k - \ell \mid X_2 = \ell)\mathbb{P}(X_2 = \ell) \\
&= \sum_{\ell=0}^{k} \mathbb{P}(X_1 = k - \ell)\mathbb{P}(X_2 = \ell) \qquad \text{(par indépendance)} \\
&= \sum_{\ell=0}^{k} \frac{\theta_1^{k-\ell}\,\mathrm{e}^{-\theta_1}}{(k-\ell)!} \frac{\theta_2^{\ell}\,\mathrm{e}^{-\theta_2}}{\ell!} \\
&= \frac{\mathrm{e}^{-\theta_1-\theta_2}}{k!} \sum_{\ell=0}^{k} \binom{k}{\ell} \theta_1^{k-\ell}\theta_2^{\ell} \\
&= \frac{(\theta_1 + \theta_2)^k\,\mathrm{e}^{-(\theta_1+\theta_2)}}{k!}.
\end{aligned}$$

Donc, $X_1 + X_2$ suit la loi de Poisson de paramètre $\theta_1 + \theta_2$.

2. D'après la définition de l'espérance conditionnelle on a

$$\mathbb{E}[z^{X_1} \mid X_1 + X_2 = m] = \sum_{k=0}^{\infty} z^k \mathbb{P}(X_1 = k \mid X_1 + X_2 = m)$$

Bien évidemment, on a $\mathbb{P}(X_1 = k \mid X_1 + X_2 = m) = 0$ pour tout $k > m$. Pour tout $0 \leq k \leq m$ on trouve

$$\begin{aligned}
\mathbb{P}(X_1 = k \mid X_1 + X_2 = m) &= \frac{\mathbb{P}(X_1 = k, X_1 + X_2 = m)}{\mathbb{P}(X_1 + X_2 = m)} = \frac{\mathbb{P}(X_1 = k, X_2 = m - k)}{\mathbb{P}(X_1 + X_2 = m)} \\
&= \frac{\mathbb{P}(X_1 = k)\mathbb{P}(X_2 = m - k)}{\mathbb{P}(X_1 + X_2 = m)} \quad \text{(par indépendance)} \\
&= \frac{\frac{\theta_1^k\,\mathrm{e}^{-\theta_1}}{k!} \frac{\theta_2^{m-k}\,\mathrm{e}^{-\theta_2}}{(m-k)!}}{\frac{(\theta_1+\theta_2)^m\,\mathrm{e}^{-(\theta_1+\theta_2)}}{m!}} \\
&= \binom{m}{k} \frac{\theta_1^k \theta_2^{m-k}}{(\theta_1 + \theta_2)^m}.
\end{aligned}$$

En fait, $(X_1 \mid X_1 + X_2 = m)$ suit une loi binomiale de paramètres m et $\theta_1/(\theta_1 + \theta_2)$. On en déduit que

$$\mathbb{E}[z^{X_1} \mid X_1 + X_2 = m] = \frac{1}{(\theta_1 + \theta_2)^m} \sum_{k=0}^{m} \binom{m}{k} (z\theta_1)^k \theta_2^{m-k} = \left(\frac{z\theta_1 + \theta_2}{\theta_1 + \theta_2}\right)^m.$$

En conclusion : pour tout $z > 0$ on a

$$\mathbb{E}[z^{X_1} \mid X_1 + X_2] = \left(\frac{z\theta_1 + \theta_2}{\theta_1 + \theta_2}\right)^{X_1 + X_2}.$$

3. Comme, conditionnellement à $X_1 + X_2$, la variable X_1 suit une loi binomiale de paramètres $X_1 + X_2$ et $\theta_1/(\theta_1 + \theta_2)$, on a

$$\mathbb{E}[X_1 \mid X_1 + X_2] = \frac{(X_1 + X_2)\theta_1}{\theta_1 + \theta_2}.$$

On peut aussi utiliser la proposition 3.19 dans le cadre de l'espérance conditionnelle :

$$\mathbb{E}[X_1 \mid X_1 + X_2 = m] = \frac{\mathrm{d}}{\mathrm{d}z}\mathbb{E}[z^{X_1} \mid X_1 + X_2]\bigg|_{z=1} = \frac{m\theta_1 \, (z\theta_1 + \theta_2)^{m-1}}{(\theta_1 + \theta_2)^m}\bigg|_{z=1} = \frac{m\theta_1}{\theta_1 + \theta_2}.$$