Variables aléatoires réelles discrètes

Corrigé des exercices

Correction de l'exercice 20. La réponse est non : si X étant une telle variable, la série des $\mathbb{P}(X = k) = \frac{\alpha}{k}$ convergerait, or ce n'est pas le cas puisque la série harmonique diverge.

Correction de l'exercice 21. On remarque tout d'abord que $0 < \frac{a}{a+b} < 1$. Soit à présent $X \sim \mathcal{B}\left(n, \frac{a}{a+b}\right)$. On a $X(\Omega) = [\![0, n]\!]$ donc

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

soit

$$1 = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{a}{a+b}\right)^k \left(1 - \frac{a}{a+b}\right)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \frac{a^k b^{n-k}}{(a+b)^n}.$$

En multipliant cette dernière égalité par $(a+b)^n$, on obtient alors

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

ce qu'il fallait démontrer.

Correction de l'exercice 22. La proposition à démontrer est évidemment vraie au rang 1 puisque $X_1 = 1$ (donc X_1 suit une loi uniforme sur $[\![1, 1]\!]$).
Soit à présent $n \in \mathbb{N}^*$; supposons que $X_n \sim \mathcal{U}([\![1, n]\!])$. Le nombre de boules blanches dans l'urne juste avant le n-ième tirage peut donc prendre toutes les valeurs de $[\![1, n]\!]$; comme ce tirage conduit à rajouter une boule blanche ou aucune, le nombre de boules blanches juste avant le tirage suivant, qui vaut X_{n+1}, peut prendre toutes les valeurs de $[\![1, n+1]\!]$. Soit à présent $k \in [\![1, n+1]\!]$. En utilisant la formule des probabilités totales sur le système complet d'événements de probabilités non nulles $((X_n = i))_{i \in [\![1,n]\!]}$, on peut écrire

$$\mathbb{P}(X_{n+1} = k) = \sum_{i=1}^{n} \mathbb{P}(X_n = i)\, \mathbb{P}(X_{n+1} = k \,|\, X_n = i).$$

Si $k = 1$, l'unique cas de figure permettant la réalisation de l'événement $(X_{n+1} = 1)$ est que seule une boule blanche ait été présente juste avant le n-ième tirage (ce qui se produit avec probabilité $\frac{1}{n}$ d'après l'hypothèse de récurrence) et que cette boule n'ait pas été tirée (ce qui se produit avec probabilité $\frac{n}{n+1}$ puisque l'urne contient n boules noires et une seule boule blanche), donc

$$\mathbb{P}(X_{n+1}=1)=\mathbb{P}(X_n=1)\mathbb{P}(X_{n+1}=1\,|\,X_n=1)=\frac{1}{n}\cdot\frac{n}{n+1}=\frac{1}{n+1}.$$

Si $k\geqslant 2$, alors k boules blanches sont présentes dans l'urne juste avant le $(n+1)$-ième tirage si et seulement si l'on est dans l'un deux cas suivants :

- $k-1$ boules blanches sont présentes dans l'urne lors du n-ième tirage (ce qui se produit avec probabilité $\frac{1}{n}$ d'après l'hypothèse de récurrence), et une boule blanche est tirée (ce qui, dans ce cas, se produit avec probabilité $\frac{k-1}{n+1}$) ;
- ou bien k boules blanches sont présentes dans l'urne lors du n-ième tirage (ce qui se produit avec probabilité $\frac{1}{n}$ d'après l'hypothèse de récurrence) et une boule noire est tirée (ce qui, dans ce cas, se produit avec probabilité $\frac{n+1-k}{n+1}$).

Ainsi, on a $\mathbb{P}(X_{n+1}=k\mid X_n=i)=0$ si $i\notin\{k-1,k\}$, donc

$$\begin{aligned}\mathbb{P}(X_{n+1}=k)&=\mathbb{P}(X_n=k-1)\mathbb{P}(X_{n+1}=k\,|\,X_n=k-1)+\mathbb{P}(X_n=k)\mathbb{P}(X_{n+1}=k\,|\,X_n=k)\\&=\frac{1}{n}\cdot\frac{k-1}{n+1}+\frac{1}{n}\cdot\frac{n+1-k}{n+1}\\&=\frac{1}{n}\cdot\frac{k-1+n+1-k}{n+1}\\&=\frac{1}{n+1}.\end{aligned}$$

On a donc $\mathbb{P}(X_{n+1}=k)=\frac{1}{n+1}$ quel que soit $k\in[\![1,n+1]\!]$, donc $X_{n+1}\sim\mathcal{U}([\![1,n+1]\!])$.

La propriété à démontrer est donc vraie pour tout $n\in\mathbb{N}^*$ d'après le principe de récurrence.

*Le système d'*urnes à renforcement *étudié dans cet exercice a été introduit par le mathématicien hongrois Georg Pólya (1887 – 1985). Il possède des applications en science économique, où il permet par exemple de modéliser les comportements mimétiques de consommateurs qui se tournent vers une marque dominante.*

Correction de l'exercice 23.

1. Supposons que l'on ait $ab=cd=\frac{1}{11}$ et $ac+bd\leqslant\frac{1}{11}$. Alors les nombres positifs a,b,c et d sont tous non nuls et on peut écrire $b=\frac{1}{11a}$ ainsi que $d=\frac{1}{11c}$. Alors

   $$ac+bd=ac+\frac{1}{11a}\cdot\frac{1}{11c}=ac+\frac{1}{121ac}=\frac{121(ac)^2+1}{121ac}.$$

   Or $ac+bd\leqslant\frac{1}{11}$, donc $\frac{121(ac)^2+1}{121ac}\leqslant\frac{1}{11}$, soit $121(ac)^2+1\leqslant 11ac$ (puisque $ac>0$), soit encore $121(ac)^2-11ac+1\leqslant 0$. Or la fonction polynomiale $f:x\mapsto 121x^2-11x+1$ a pour discriminant $11^2-4\cdot121=-3\cdot121<0$, donc elle ne change pas de signe sur $\mathbb{R}$; son coefficient dominant étant strictement positif, elle ne prend que des valeurs strictement positives sur $\mathbb{R}$, ce qui induit une contradiction avec le fait que $f(ac)\leqslant 0$.

   Ainsi, il n'est pas possible d'avoir à la fois $ab=cd=\frac{1}{11}$ et $ac+bd\leqslant\frac{1}{11}$.

2. Comme X et Y sont à valeurs dans $[\![1,6]\!]$, on a

$$\begin{aligned}\mathbb{P}(X+Y=2) &= \mathbb{P}\Big((X=1)\cap(Y=1)\Big)\\ &= \mathbb{P}(X=1)\,\mathbb{P}(Y=1) \quad \text{par indépendance de } X \text{ et } Y\\ &= ab,\end{aligned}$$

et par le même argument

$$\mathbb{P}(X+Y=12) = \mathbb{P}(X=6)\,\mathbb{P}(Y=6) = dc.$$

Comme l'événement $(X+Y=7)$ est réalisé notamment lorsque X et Y prennent les valeurs respectives 1 et 6 ou 6 et 1, on a

$$\begin{aligned}\mathbb{P}(X+Y=7) &\geqslant \mathbb{P}\Big(\big((X=1)\cap(Y=6)\big)\sqcup\big((X=6)\cap(Y=1)\big)\Big)\\ &= \mathbb{P}\Big((X=1)\cap(Y=6)\Big)+\mathbb{P}\Big((X=6)\cap(Y=1)\Big) \quad \text{par incompatibilité}\\ &= \mathbb{P}(X=1)\,\mathbb{P}(Y=6)+\mathbb{P}(X=6)\,\mathbb{P}(Y=1) \quad \text{par indépendance de } X \text{ et } Y\\ &= ac+db.\end{aligned}$$

Or les événements $(X+Y=2), (X+Y=12)$ et $(X+Y=7)$ sont tous trois de probabilité $\frac{1}{11}$ puisque $X+Y \sim \mathcal{U}([\![2,12]\!])$; on a donc $ab = cd = \frac{1}{11}$ et $ac+bd \leqslant \frac{1}{11}$, ce qui est impossible d'après la question précédente.

Il est donc impossible de définir des VARD X et Y indépendantes et à valeurs dans $[\![1,6]\!]$ telles que $X+Y$ suive la loi $\mathcal{U}([\![2,12]\!])$. On obtient ainsi la proposition que l'on voulait démontrer : on ne peut pas piper deux dés à six faces de sorte que leurs résultats X et Y vérifient $X+Y \sim \mathcal{U}([\![2,12]\!])$.

Correction de l'exercice 24. Posons $u_k := \mathbb{P}(X=k) = \binom{n}{k}p^k(1-p)^{n-k}$ pour tout $k \in [\![0,n]\!]$. Les u_k sont tous strictement positifs, et si $k \in [\![0,n-1]\!]$ on a

$$\frac{u_{k+1}}{u_k} = \frac{\binom{n}{k+1}p^{k+1}(1-p)^{n-k-1}}{\binom{n}{k}p^k(1-p)^{n-k}} = \frac{\frac{n!}{(k+1)!(n-k-1)!}}{\frac{n!}{k!(n-k)!}}\,\frac{p}{1-p} = \frac{(n-k)p}{(k+1)(1-p)},$$

donc

$$\begin{aligned}\frac{u_{k+1}}{u_k} \geqslant 1 \quad &\iff \quad (n-k)p \geqslant (k+1)(1-p)\\ &\iff \quad np-\cancel{kp} \geqslant k+1-p-\cancel{kp}\\ &\iff \quad k \leqslant (n+1)p-1.\end{aligned}$$

Ainsi, tant que $k \leqslant (n+1)p-1$, on a $u_k \leqslant u_{k+1}$, et lorsque $k > (n+1)p-1$ on a $u_k > u_{k+1}$. La valeur maximale de u_k est donc atteinte pour le plus petit entier k tel que $k > (n+1)p-1$, c'est-à-dire pour $k = \lfloor (n+1)p \rfloor$.

Correction de l'exercice 25. Les VARD X et Y ne peuvent prendre que les valeurs 0 et 1, donc elles suivent des lois de Bernoulli.

Or l'événement $(X=1)$ est réalisé si et seulement si tous les événements $(X_i=1)$

sont réalisés, pour $i \in [\![1, n]\!]$, si bien que

$$\begin{aligned}\mathbb{P}(X = 1) &= \mathbb{P}\Big((X_1 = 1) \cap (X_2 = 1) \cap \ldots \cap (X_n = 1)\Big) \\ &= \mathbb{P}(X_1 = 1)\, \mathbb{P}(X_2 = 1) \cdots \mathbb{P}(X_n = 1) = p^n\end{aligned}$$

par indépendance des X_i : ainsi, $X \sim \mathcal{B}(p^n)$.
De même, l'événement $(Y = 0)$ est réalisé si et seulement si tous les événements $(X_i = 0)$, pour $i \in [\![1, n]\!]$, sont réalisés, si bien que

$$\begin{aligned}\mathbb{P}(Y = 0) &= \mathbb{P}\Big((X_1 = 0) \cap (X_2 = 0) \cap \ldots \cap (X_n = 0)\Big) \\ &= \mathbb{P}(X_1 = 0)\, \mathbb{P}(X_2 = 0) \cdots \mathbb{P}(X_n = 0) = (1 - p)^n.\end{aligned}$$

On a donc $\mathbb{P}(Y = 1) = 1 - (1 - p)^n$, d'où $Y \sim \mathcal{B}(1 - (1 - p)^n)$.

Correction de l'exercice 26.

1. X compte le nombre de succès obtenus lors de la répétition de 100 expériences de Bernoulli indépendantes de probabilité de succès p, donc $X \sim \mathcal{B}(100, p)$.
2. La variable aléatoire $100 - X$ compte le nombre d'échecs obtenus dans le schéma décrit dans la question précédente ; la probabilité d'échec d'une expérience individuelle étant égale à $1 - p$, on a $100 - X \sim \mathcal{B}(100, 1 - p)$.
3. Le profit réalisé par la compagnie grâce à la vente des 100 tickets est égal à $100 \cdot 200 = 20\,000$ €. Le nombre de passagers surnuméraires vaut $\max(0, X - 97)$ (il est nul si moins de 97 passagers se présentent à l'embarquement et prend la valeur $k - 97$ si $k > 97$ passagers se présentent), et chacun de ces passagers occasionne un coût de $200 + 400 = 600$ € pour la compagnie. Ainsi, le profit final réalisé par la compagnie est bien égal à $20\,000 - 600 \max(0, X - 97)$ €.
4. Comme 100 tickets ont été vendus pour 97 places, la variable aléatoire S représentant le nombre de passagers surnuméraires prend les valeurs $0, 1, 2$ ou 3. Par ailleurs, on a

$$\mathbb{P}(S = 1) = \mathbb{P}(X = 98) = \binom{100}{98} p^{98}(1 - p)^2 = 4950p^{98}(1 - p)^2,$$

$$\mathbb{P}(S = 2) = \mathbb{P}(X = 99) = \binom{100}{99} p^{99}(1 - p) = 99p^{99}(1 - p)$$

et

$$\mathbb{P}(S = 3) = \mathbb{P}(X = 100) = \binom{100}{100} p^{100} = p^{100}.$$

5. Comme $\max(0, X - 97)$ est à valeurs dans $\{0, 1, 2, 3\}$, elle admet une espérance donnée par

$$\begin{aligned}\mathbb{E}(S) &= \mathbb{P}(S = 1) + 2\,\mathbb{P}(S = 2) + 3\,\mathbb{P}(S = 3) \\ &= 4950p^{98}(1 - p)^2 + 198p^{99}(1 - p) + 3p^{100} \\ &= 4755p^{100} - 9702p^{99} + 4950p^{98}.\end{aligned}$$

La variable aléatoire $Y = 20\,000 - 600\,S$ admet une espérance égale par linéarité à
$$\mathbb{E}(Y) = 20\,000 - 600\,\mathbb{E}(S).$$
En l'absence de surbooking, le profit réalisé serait de $97 \cdot 200 = 19\,400$ € ; ainsi, le profit moyen en cas de surbooking est supérieur au profit sans surbooking si et seulement si $20\,000 - 600\,\mathbb{E}(S) \geqslant 19\,400$, donc si et seulement $600\,\mathbb{E}(S) \leqslant 600$, soit encore $\mathbb{E}(S) \leqslant 1$. Une condition nécessaire et suffisante sur p pour que le profit moyen en cas de surbooking soit supérieur est donc que
$$4755p^{100} - 9702p^{99} + 4950p^{98} \leqslant 1.$$
Une analyse numérique suggère que cette condition est réalisée si et seulement si p est inférieur à un seuil situé autour de 0,9767, c'est-à-dire si la probabilité qu'un passager soit effectivement présent à l'embarquement ne dépasse pas 97,67 % ; l'existence d'un tel seuil est conforme à l'intuition puisque la pratique du surbooking n'a d'intérêt que si p est significativement inférieur à 1.

Correction de l'exercice 27. On remarque tout d'abord que si $\theta \equiv 0[\pi]$ alors les variables aléatoires $\sin(X\theta)$ sont nulles (puisque les VARD X considérées sont à valeurs entières), donc d'espérance nulle. On considère donc dans la suite que $\theta \not\equiv 0[\pi]$.

(i) Comme X prend un nombre fini de valeurs, c'est aussi le cas de $\sin(X\theta)$ qui admet donc une espérance. D'après le théorème de transfert, cette dernière vaut
$$\begin{aligned}\mathbb{E}(\sin(X\theta)) &= \sum_{k=1}^{n} \sin(k\theta)\,\mathbb{P}(X = k)\\ &= \sum_{k=1}^{n} \frac{1}{n}\sin(k\theta) = \frac{1}{n}\sum_{k=1}^{n}\sin(k\theta).\end{aligned}$$
Or
$$\sum_{k=1}^{n}\sin(k\theta) = \operatorname{Im}\left(\sum_{k=1}^{n} e^{ik\theta}\right) = \operatorname{Im}\left(\sum_{k=1}^{n}\left(e^{i\theta}\right)^k\right) = \operatorname{Im}\left(\frac{e^{i\theta} - e^{i\theta(n+1)}}{1 - e^{i\theta}}\right)$$
car $e^{i\theta} \neq 1$ (puisque $\theta \not\equiv 0\,[2\pi]$). En faisant apparaître la demi-somme des arguments, on peut écrire
$$\begin{aligned}\frac{e^{i\theta} - e^{i\theta(n+1)}}{1 - e^{i\theta}} &= \frac{e^{i\frac{(n+2)\theta}{2}}}{e^{i\frac{\theta}{2}}} \cdot \frac{e^{-i\frac{n\theta}{2}} - e^{i\frac{n\theta}{2}}}{e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}}\\ &= e^{i\frac{(n+1)\theta}{2}} \cdot \frac{2i\sin\left(-\frac{n\theta}{2}\right)}{2i\sin\left(-\frac{\theta}{2}\right)} \quad \text{par la deuxième formule d'Euler}\\ &= e^{i\frac{(n+1)\theta}{2}}\frac{\sin\left(\frac{n\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)},\end{aligned}$$

si bien que

$$\sum_{k=1}^{n}\sin(k\theta) = \operatorname{Im}\left(e^{i\frac{(n+1)\theta}{2}}\frac{\sin\left(\frac{n\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}\right) = \sin\left(\frac{(n+1)\theta}{2}\right)\frac{\sin\left(\frac{n\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}.$$

On a donc

$$\mathbb{E}(\sin(X\theta)) = \frac{\sin\left(\frac{(n+1)\theta}{2}\right)\sin\left(\frac{n\theta}{2}\right)}{n\sin\left(\frac{\theta}{2}\right)}.$$

(ii) La VARD X prenant là encore un nombre fini de valeurs, c'est aussi le cas de $\sin(X\theta)$ qui admet donc une espérance. D'après le théorème de transfert, cette dernière vaut

$$\begin{aligned}
\mathbb{E}(\sin(X\theta)) &= \sum_{k=0}^{n}\sin(k\theta)\,\mathbb{P}(X=k)\\
&= \sum_{k=0}^{n}\sin(k\theta)\binom{n}{k}\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{n-k}\\
&= \operatorname{Im}\left(\sum_{k=0}^{n}e^{ik\theta}\binom{n}{k}\frac{1}{2^n}\right)\\
&= \frac{1}{2^n}\operatorname{Im}\left(\sum_{k=0}^{n}\binom{n}{k}(e^{i\theta})^k\right)\\
&= \frac{1}{2^n}\operatorname{Im}\Big((e^{i\theta}+1)^n\Big) \qquad \text{par la formule du binôme.}
\end{aligned}$$

Or une factorisation par la demi-somme des arguments amène

$$(e^{i\theta}+1)^n = \left(e^{i\frac{\theta}{2}}(e^{-i\frac{\theta}{2}}+e^{i\frac{\theta}{2}})\right)^n = \left(e^{i\frac{\theta}{2}}\cdot 2\cos\left(\frac{\theta}{2}\right)\right)^n = 2^n e^{i\frac{n\theta}{2}}\left(\cos\left(\frac{\theta}{2}\right)\right)^n,$$

donc

$$\mathbb{E}(\sin(X\theta)) = \frac{1}{2^n}\operatorname{Im}\left(2^n e^{i\frac{n\theta}{2}}\cos^n\left(\frac{\theta}{2}\right)\right) = \sin\left(\frac{n\theta}{2}\right)\left(\cos\left(\frac{\theta}{2}\right)\right)^n.$$

(iii) Pour tout $N \geqslant 1$, on a [1]

$$\begin{aligned}
\sum_{k=1}^{N}\sin(k\theta)\,\mathbb{P}(X=k) &= p\sum_{k=1}^{N}\sin(k\theta)(1-p)^{k-1}\\
&= p\operatorname{Im}\left(\sum_{n=0}^{N-1}e^{i(n+1)\theta}(1-p)^n\right) \qquad \text{en posant } n=k+1
\end{aligned}$$

1. Le passage par les sommes partielles a pour unique but d'éviter d'utiliser un résultat sur les séries géométriques complexes qui stipule que pour tout $z \in \mathbb{C}$ tel que $|z| < 1$, la série de terme général z^k converge et $\sum_{k=0}^{+\infty} z^k = \frac{1}{1-z}$ (au sens d'une limite de suite complexe). Les lecteurs qu'une telle égalité ne dérange pas pourront utiliser ce résultat pour écrire directement l'égalité $\mathbb{E}(\sin(X\theta)) = p\operatorname{Im}\left(\frac{e^{i\theta}}{1-(1-p)e^{i\theta}}\right)$ et simplifier considérablement les calculs développés ici...

$$= p\,\mathrm{Im}\left(e^{i\theta}\sum_{n=0}^{N-1}(e^{i\theta}(1-p))^n\right)$$
$$= p\,\mathrm{Im}\left(e^{i\theta}\frac{1-(1-p)^N e^{iN\theta}}{1-(1-p)e^{i\theta}}\right) \quad \text{car } (1-p)e^{i\theta}\neq 1.$$

Or on peut accéder à la forme algébrique de $\frac{1-(1-p)^N e^{iN\theta}}{1-(1-p)e^{i\theta}}$ en multipliant les deux termes du quotient par le conjugué du dénominateur :

$$\begin{aligned}
\frac{1-(1-p)^N e^{iN\theta}}{1-(1-p)e^{i\theta}} &= \frac{(1-(1-p)^N e^{iN\theta})(1-(1-p)e^{-i\theta})}{|1-(1-p)e^{i\theta}|^2}\\
&= \frac{1-(1-p)^N e^{iN\theta}-(1-p)e^{-i\theta}+(1-p)^{N+1}e^{i(N-1)\theta}}{(1-(1-p)\cos(\theta))^2+((1-p)\sin(\theta))^2}\\
&= \frac{1-(1-p)^N e^{iN\theta}-(1-p)e^{-i\theta}+(1-p)^{N+1}e^{i(N-1)\theta}}{1+(1-p)^2-2(1-p)\cos(\theta)}
\end{aligned}$$

soit

$$\frac{1-(1-p)^N e^{iN\theta}}{1-(1-p)e^{i\theta}} = \frac{1-(1-p)^N e^{iN\theta}-(1-p)e^{-i\theta}+(1-p)^{N+1}e^{i(N-1)\theta}}{(2-2p)(1-\cos(\theta))+p^2},$$

d'où

$$e^{i\theta}\frac{1-(1-p)^N e^{iN\theta}}{1-(1-p)e^{i\theta}} = \frac{e^{i\theta}-(1-p)^N e^{i(N+1)\theta}-(1-p)+(1-p)^{N+1}e^{iN\theta}}{(2-2p)(1-\cos(\theta))+p^2},$$

et donc

$$\begin{aligned}
\sum_{k=1}^{N}\sin(k\theta)\,\mathbb{P}(X=k) &= p\,\mathrm{Im}\left(e^{i\theta}\frac{1-(1-p)^N e^{iN\theta}}{1-(1-p)e^{i\theta}}\right)\\
&= p\,\frac{\sin(\theta)-(1-p)^N\sin((N+1)\theta)-(1-p)+(1-p)^{N+1}\sin(N\theta)}{(2-2p)(1-\cos(\theta))+p^2}.
\end{aligned}$$

Or sin est bornée et $|1-p|<1$, donc on a $(1-p)^N\sin((N+1)\theta)\xrightarrow[N\to+\infty]{}0$ et $(1-p)^{N+1}\sin(N\theta)\xrightarrow[N\to+\infty]{}0$, ce qui montre que la série de terme général $\sin(k\theta)\,\mathbb{P}(X=k)$ converge vers

$$\sum_{k=1}^{+\infty}\sin(k\theta)\,\mathbb{P}(X=k) = \frac{p(\sin(\theta)-(1-p))}{(2-2p)(1-\cos(\theta))+p^2}.$$

Ainsi, $\mathbb{E}(\sin(X\theta))$ existe et on a

$$\mathbb{E}(\sin(X\theta)) = \frac{p(\sin(\theta)-(1-p))}{(2-2p)(1-\cos(\theta))+p^2}.$$

Correction de l'exercice 28.

1. La VARD X prend un nombre fini de valeurs, donc elle admet une espérance. On a par ailleurs

$$\begin{aligned}\sum_{k=0}^{N} \mathbb{P}(X > k) &= \sum_{k=0}^{N} \sum_{j=k+1}^{N} \mathbb{P}(X = j) \\ &= \sum_{j=1}^{N} \sum_{k=0}^{j-1} \mathbb{P}(X = j) = \sum_{j=1}^{N} j\, \mathbb{P}(X = j) = \mathbb{E}(X),\end{aligned}$$

ce qu'il fallait démontrer.

2. Soient comme dans l'énoncé X_1, X_2, X_3 indépendantes et de même loi uniforme sur $[\![1, N]\!]$ et $X = \min(X_1, X_2, X_3)$. On a alors $X(\Omega) = [\![1, N]\!]$, et pour tout $k \in [\![0, N]\!]$:

$$\begin{aligned}\mathbb{P}(X > k) &= \mathbb{P}((X_1 > k) \cap (X_2 > k) \cap (X_3 > k)) \\ &= \mathbb{P}(X_1 > k)\, \mathbb{P}(X_2 > k)\, \mathbb{P}(X_3 > k) \quad \text{par indépendance des } X_k \\ &= \left(\frac{N-k}{N}\right)^3.\end{aligned}$$

Ainsi, d'après la question précédente :

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=0}^{N} \mathbb{P}(X > k) = \sum_{k=0}^{N} \left(\frac{N-k}{N}\right)^3 \\ &= \sum_{i=0}^{N} \left(\frac{i}{N}\right)^3 \quad \text{en posant } i = N - k \\ &= \frac{1}{N^3} \sum_{i=0}^{N} i^3 = \frac{1}{N^3} \left(\frac{N(N+1)}{2}\right)^2 \\ &= \frac{(N+1)^2}{4N}.\end{aligned}$$

Correction de l'exercice 29.

1. On a $(X+Y)(\Omega) = [\![2, 12]\!]$ puisque $X + Y$ peut prendre chaque valeur entre 2 et 12 et uniquement ces valeurs. Pour tout $k \in [\![2, 7]\!]$, on a

$$\begin{aligned}\mathbb{P}(X + Y = k) &= \mathbb{P}\left(\bigsqcup_{i=1}^{k-1} (X = i) \cap (Y = k - i)\right) \\ &= \sum_{i=1}^{k-1} \mathbb{P}((X = i) \cap (Y = k - i)) \quad \text{par incompatibilité} \\ &= \sum_{i=1}^{k-1} \mathbb{P}(X = i)\, \mathbb{P}(Y = k - i) \quad \text{par indépendance de } X \text{ et } Y \\ &= \sum_{i=1}^{k-1} \frac{1}{6} \cdot \frac{1}{6} \\ &= \frac{k-1}{36}.\end{aligned}$$

Si $k \in [\![8, 12]\!]$, on ne peut pas effectuer le même calcul puisque le produit $\mathbb{P}(X = i)\,\mathbb{P}(Y = k - i)$ est nul dès lors que $i > 6$ ou $k - i > 6$ (c'est-à-dire $i < k - 6$) : on écrit alors

$$\begin{aligned}
\mathbb{P}(X + Y = k) &= \mathbb{P}\left(\bigsqcup_{i=k-6}^{6} (X = i) \cap (Y = k - i)\right) \\
&= \sum_{i=k-6}^{6} \mathbb{P}((X = i) \cap (Y = k - i)) \quad \text{par incompatibilité} \\
&= \sum_{i=k-6}^{6} \mathbb{P}(X = i)\,\mathbb{P}(Y = k - i) \quad \text{par indépendance de } X \text{ et } Y \\
&= \sum_{i=k-6}^{6} \frac{1}{6} \cdot \frac{1}{6} \\
&= \frac{13 - k}{36}.
\end{aligned}$$

On peut présenter la loi de $X + Y$ dans le tableau suivant [2]. Les fractions y sont données sous une forme non simplifiée pour faire apparaître la progression des probabilités associées :

x	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

2. Par le même raisonnement [3] que dans le point précédent, on obtient que

$$(X + Y)(\Omega) = \{k \in \mathbb{N} \,:\, k \geqslant 2\} = \mathbb{N} \setminus \{0, 1\}$$

et que pour tout $k \geqslant 2$ on a :

$$\begin{aligned}
\mathbb{P}(X + Y = k) &= \mathbb{P}\left(\bigsqcup_{i=1}^{k-1} (X = i) \cap (Y = k - i)\right) \\
&= \sum_{i=1}^{k-1} \mathbb{P}((X = i) \cap (Y = k - i)) \quad \text{par incompatibilité} \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X = i)\,\mathbb{P}(Y = k - i) \quad \text{par indépendance de } X \text{ et } Y \\
&= \sum_{i=1}^{k-1} p(1 - p)^{i-1} p(1 - p)^{k-i-1}
\end{aligned}$$

2. Cette loi est importante dans certains jeux de société puisqu'elle est celle de la somme des résultats de deux dés à six faces indépendants.

3. La loi de $X+Y$ est la *loi de Pascal de paramètres* 2 *et* p *(voir l'exercice 55). La VARD* $X+Y$ *représente le rang d'obtention du* deuxième *succès lors de la répétition d'épreuves de Bernoulli indépendantes de probabilité de succès* p.

$$
\begin{aligned}
&= p^2 \sum_{i=1}^{k-1} (1-p)^{k-2} \\
&= (k-1)p^2(1-p)^{k-2}.
\end{aligned}
$$

3. On a une fois encore $(X+Y)(\Omega) = \{k \in \mathbb{N} \,:\, k \geqslant 2\} = \mathbb{N} \setminus \{0,1\}$: en effet, les valeurs prises par $X+Y$ sont des entiers supérieurs ou égaux à 2, et pour tout $k \geqslant 2$ la VARD $X+Y$ peut prendre la valeur k, par exemple dès que X prend la valeur $k-1$ et Y prend la valeur 1.

   Le cas le plus facile est celui où $k \geqslant 7$: dans ce cas, toutes les valeurs de Y sont compatibles avec la réalisation de l'événement $(X+Y=k)$ et on a

$$
\begin{aligned}
\mathbb{P}(X+Y=k) &= \mathbb{P}\left(\bigsqcup_{i=1}^{6} (X=k-i) \cap (Y=i)\right) \\
&= \sum_{i=1}^{6} \mathbb{P}((X=k-i) \cap (Y=i)) \quad \text{par incompatibilité} \\
&= \sum_{i=1}^{6} \mathbb{P}(X=k-i)\, \mathbb{P}(Y=i) \quad \text{par indépendance de } X \text{ et } Y \\
&= \sum_{i=1}^{6} p(1-p)^{k-i-1} \cdot \frac{1}{6} \\
&= \frac{p}{6}(1-p)^{k-1} \sum_{i=1}^{6} \left(\frac{1}{1-p}\right)^i \\
&= \frac{p}{6}(1-p)^{k-1} \cdot \frac{\frac{1}{1-p} - \left(\frac{1}{1-p}\right)^7}{1 - \frac{1}{1-p}} \\
&= \frac{1}{6}\left((1-p)^{k-7} - (1-p)^{k-1}\right).
\end{aligned}
$$

   Si $k \in [\![2,6]\!]$, toutes les valeurs de Y ne sont pas à même de réaliser l'événement $(X+Y=k)$, mais seulement celles comprises dans $[\![1,k-1]\!]$: on a dans ce cas

$$
\begin{aligned}
\mathbb{P}(X+Y=k) &= \mathbb{P}\left(\bigsqcup_{i=1}^{k-1} (X=k-i) \cap (Y=i)\right) \\
&= \sum_{i=1}^{k-1} \mathbb{P}((X=k-i) \cap (Y=i)) \quad \text{par incompatibilité} \\
&= \sum_{i=1}^{k-1} \mathbb{P}(X=k-i)\, \mathbb{P}(Y=i) \quad \text{par indépendance de } X \text{ et } Y \\
&= \sum_{i=1}^{k-1} p(1-p)^{k-i-1} \cdot \frac{1}{6} \\
&= \frac{p}{6} \sum_{i=1}^{k-1} (1-p)^{k-i-1} \\
&= \frac{p}{6} \sum_{j=0}^{k-2} (1-p)^j \quad \text{en posant } j = k-i-1
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{p}{6} \cdot \frac{1-(1-p)^{k-1}}{1-(1-p)} \\
&= \frac{1}{6}(1-(1-p)^{k-1}).
\end{aligned}
$$

Correction de l'exercice 30.

1. En chaque état du monde $\omega \in \Omega$, $Y(\omega)$ est la valeur minimale entre les $X_1(\omega), \ldots, X_n(\omega)$ (qui sont positifs), donc *a fortiori* $0 \leqslant Y(\omega) \leqslant X_1(\omega)$: ainsi, on a $0 \leqslant Y \leqslant X_1$. De même, pour tout $\omega \in \Omega$ on a $Z(\omega) = \max(X_1(\omega), \ldots, X_n(\omega))$ donc, comme les $X_i(\omega)$ sont positifs :
$$X_1(\omega) \leqslant Z(\omega) \leqslant X_1(\omega) + \ldots + X_n(\omega),$$
d'où la double inégalité $X_1 \leqslant Z \leqslant X_1 + \ldots + X_n$.

   Comme les n VARD X_i admettent une espérance, c'est aussi le cas de leur somme $X_1 + \ldots + X_n$. Ainsi, Y et Z sont toutes deux des VARD positives majorées par la VARD $X_1 + \ldots + X_n$ qui admet une espérance : elles admettent donc elles-mêmes une espérance.

2. (a) Il s'agit ici de redémontrer la formule donnée dans le lemme 51 du cours, qui est la version « infinie » de la formule utilisée dans l'exercice 28. Les probabilités $\mathbb{P}(X = j)$ étant positives, on peut appliquer le théorème de Fubini et écrire :
$$
\begin{aligned}
\sum_{k=0}^{+\infty} \mathbb{P}(Y > k) &= \sum_{k=0}^{+\infty} \sum_{j=k+1}^{+\infty} \mathbb{P}(Y = j) \\
&= \sum_{j=1}^{+\infty} \sum_{k=0}^{j-1} \mathbb{P}(Y = j) \\
&= \sum_{j=1}^{+\infty} j\, \mathbb{P}(Y = k) \\
&= \mathbb{E}(Y),
\end{aligned}
$$
ce qui est le résultat attendu.

   (b) Remarquons que cette question a déjà été traitée dans l'exercice 11 avec d'autres notations.

   Soit $k \in \mathbb{N}$. Comme $Y = \min(X_1, \ldots, X_n)$, on a
$$\mathbb{P}(Y > k) = \mathbb{P}\left(\bigcap_{i=1}^{n} (X_i > k)\right) = \prod_{i=1}^{n} \mathbb{P}(X_i > k)$$
par indépendance des X_i. Or pour tout i, on peut [4] calculer $\mathbb{P}(X_i > k)$ en interprétant X_i comme le temps d'attente du premier succès lors de

4. Alternativement, un calcul direct *via* le reste d'une série géométrique aurait été possible :
$$\mathbb{P}(X_i > k) = \sum_{j=k+1}^{+\infty} \mathbb{P}(X = j) = \sum_{j=k+1}^{+\infty} p(1-p)^{j-1} = (1-p)^k.$$

la répétition d'expériences de Bernoulli indépendantes de probabilité de succès p, si bien que $\mathbb{P}(X_i > k)$ est la probabilité que les k premières expériences se soient soldées par un échec et vaut donc $(1-p)^k$. Ainsi, on a

$$\mathbb{P}(Y > k) = \prod_{i=1}^{n} (1-p)^k = (1-p)^{nk}.$$

Déterminons à présent la loi de Y. On a $Y(\Omega) = \mathbb{N}^*$ puisque les X_i peuvent prendre n'importe quelle valeur de $\mathbb{N}^*$, éventuellement simultanément. Pour tout $k \in \mathbb{N}$, on a

$$\begin{aligned}
\mathbb{P}(Y = k) &= \mathbb{P}(Y > k-1) - \mathbb{P}(Y > k) \\
&= (1-p)^{n(k-1)} - (1-p)^{nk} \quad \text{d'après la question précédente} \\
&= (1-(1-p)^n)((1-p)^n)^{k-1}.
\end{aligned}$$

Ainsi, la variable aléatoire Y suit la loi $\mathcal{G}(1-(1-p)^n)$.

Évoquons au passage une interprétation de ce résultat. Considérons n basketteurs qui réalisent une fois par minute des tirs au panier simultanés et ont une probabilité p de marquer à chaque tentative, indépendamment des autres tentatives et des autres basketteurs. Notons $X_1, X_2, \ldots, X_n$ les temps d'attente respectifs du premier succès du basketteur $1, 2, \ldots, n$; les X_i sont alors mutuellement indépendantes et de même loi $\mathcal{G}(p)$. Dans ce cas, $Y = \min(X_1, \ldots, X_n)$ est la variable aléatoire indiquant le temps d'attente du premier succès observé. Or l'expérience menée revient à répéter à chaque minute et de façon indépendante une unique expérience de Bernoulli consistant à savoir si *au moins un panier* est marqué par l'un des n basketteurs ; les expériences successives ainsi considérées étant indépendantes entre elles et de probabilité de succès $1-(1-p)^n$ (soit la probabilité que tous les basketteurs ne ratent pas leur lancer), on a bien $Y \sim \mathcal{G}(1-(1-p)^n)$.

3. (a) Pour tout $k \in \mathbb{N}$, on a

$$\begin{aligned}
\mathbb{P}(Z \leqslant k) &= \mathbb{P}\left(\bigcap_{i=1}^{n} (X_i \leqslant k) \right) \\
&= \prod_{i=1}^{n} \mathbb{P}(X_i \leqslant k) \quad \text{par indépendance des } X_i \\
&= \left(1-(1-p)^k\right)^n.
\end{aligned}$$

Déterminons à présent la loi de Z. On a $Z(\Omega) = \mathbb{N}^*$ pour la même raison que dans le cas de Y, et pour tout $k \geqslant 1$, on peut écrire

$$\begin{aligned}
\mathbb{P}(Z = k) &= \mathbb{P}(Z \leqslant k) - \mathbb{P}(Z \leqslant k-1) \\
&= \left(1-(1-p)^k\right)^n - \left(1-(1-p)^{k-1}\right)^n.
\end{aligned}$$

(b) L'espérance de Z est difficile à calculer dans le cas général, quelle que soit la formule utilisée. En revanche, si $n = 2$, on peut écrire que

$$X_1 + X_2 = \min(X_1, X_2) + \max(X_1, X_2) = Y + Z$$

pour trouver

$$\begin{aligned}\mathbb{E}(Z) &= \mathbb{E}(X_1 + X_2 - Y) = \mathbb{E}(X_1) + \mathbb{E}(X_2) - \mathbb{E}(Y) \quad \text{par linéarité de l'espérance}\\ &= \frac{1}{p} + \frac{1}{p} - \frac{1}{1-(1-p)^2} = \frac{3-2p}{2p-p^2}.\end{aligned}$$

Correction de l'exercice 31.

1. (a) On a $X(\Omega) = [\![1, n]\!]$, et pour tout $k \in [\![1, n]\!]$ on a $\mathbb{P}(X = k) = k\,\mathbb{P}(X_1)$ puisque l'événement $(X = k)$ se produit avec une probabilité k fois supérieure à celle de l'événement $(X = 1)$. Ainsi, on a

$$1 = \sum_{k=1}^{n} \mathbb{P}(X = k) = \mathbb{P}(X = 1) \sum_{k=1}^{n} k = \mathbb{P}(X = 1)\frac{n(n+1)}{2},$$

d'où $\mathbb{P}(X = 1) = \frac{2}{n(n+1)}$ et enfin :

$$\forall k \in [\![1, n]\!], \quad \mathbb{P}(X = k) = \frac{2k}{n(n+1)}.$$

(b) Comme X prend un nombre fini de valeurs, elle admet des moments de tous ordres.

Son espérance est

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^{n} k \cdot \frac{2k}{n(n+1)} = \frac{2}{n(n+1)} \sum_{k=1}^{n} k^2\\ &= \frac{2}{n(n+1)} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{2n+1}{3}.\end{aligned}$$

Son moment d'ordre 2 est

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{k=1}^{n} k^2 \cdot \frac{2k}{n(n+1)} = \frac{2}{n(n+1)} \sum_{k=1}^{n} k^3\\ &= \frac{2}{n(n+1)} \cdot \left(\frac{n(n+1)}{2}\right)^2 = \frac{n(n+1)}{2},\end{aligned}$$

donc la formule de Koenig donne

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \frac{n(n+1)}{2} - \frac{(2n+1)^2}{9} = \frac{(n-1)(n+2)}{18}.$$

2. (a) On a $Y(\Omega) = \{1, \frac{1}{2}, \ldots, \frac{1}{n}\}$, et pour tout $k \in [\![1, n]\!]$,

$$\mathbb{P}\left(Y = \frac{1}{k}\right) = \mathbb{P}(X = k) = \frac{2k}{n(n+1)}.$$

(b) Comme Y prend un nombre fini de valeurs, elle admet une espérance, et on a d'après le théorème de transfert :

$$\mathbb{E}(Y) = \mathbb{E}\left(\frac{1}{X}\right) = \sum_{k=1}^{n} \frac{1}{k} \cdot \frac{2k}{n(n+1)} = \sum_{k=1}^{n} \frac{2}{n(n+1)} = \frac{2}{n+1}.$$

Correction de l'exercice 32. Notons X le numéro de l'urne choisie lors de la première étape. On a donc $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$, et pour tous $i, k \in \llbracket 1, n \rrbracket$:

$$\mathbb{P}(Y = k \mid X = i) = \begin{cases} \frac{1}{i} & \text{si} \quad 0 \leqslant k \leqslant i \\ 0 & \text{si} \quad k > i. \end{cases}$$

1. N'importe quel numéro de boule entre 1 et n peut être tiré, donc $Y(\Omega) = \llbracket 1, n \rrbracket$. Si $k \in \llbracket 1, n \rrbracket$, en appliquant la formule des probabilités totales au système complet d'événements de probabilités non nulles $((X = 1), \ldots, (X = n))$, on obtient :
$$\begin{aligned} \mathbb{P}(Y = k) &= \sum_{i=1}^{n} \mathbb{P}(Y = k \mid X = i)\, \mathbb{P}(X = i) \\ &= \sum_{i=k}^{n} \frac{1}{i} \cdot \frac{1}{n} \\ &= \frac{1}{n} \sum_{i=k}^{n} \frac{1}{i}. \end{aligned}$$

2. Comme Y prend un nombre fini de valeurs, elle admet une espérance, et on a :
$$\begin{aligned} \mathbb{E}(Y) &= \sum_{k=1}^{n} k\, \mathbb{P}(Y = k) = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=k}^{n} \frac{k}{i} \quad \text{d'après la question précédente} \\ &= \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{i} \frac{k}{i} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{i} \sum_{k=1}^{i} k \\ &= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{i} \cdot \frac{i(i+1)}{2} = \frac{1}{n} \sum_{i=1}^{n} \frac{i+1}{2} \\ &= \frac{1}{2n} \left(\sum_{i=1}^{n} i + \sum_{i=1}^{n} 1 \right) = \frac{1}{2n} \left(\frac{n(n+1)}{2} + n \right) = \frac{n+3}{4}. \end{aligned}$$
On remarque que cette valeur moyenne est comprise entre 0 et n (puisque $n \geqslant 1$), mais aussi qu'elle est de l'ordre de $\frac{n}{4}$, donc plus petite que la valeur moyenne d'un tirage uniforme entre 1 et n. Ce fait est conforme à l'intuition puisque le protocole utilisé revient à tirer un numéro de boule en favorisant les boules à faibles numéros, présentes dans de nombreuses urnes.

Correction de l'exercice 33.

1. La variable aléatoire U_2 prend la valeur 0 si la deuxième boule tirée est la même que la première et 1 sinon. En utilisant le fait que les tirages sont indépendants et réalisés selon une loi uniforme, on obtient
$$\begin{aligned} &\mathbb{P}(U_2 = 0) = \mathbb{P}(T_2 = T_1) \\ &\quad = \sum_{k=1}^{n} \mathbb{P}\left((T_1 = k) \cap (T_2 = k)\right) = \sum_{k=1}^{n} \mathbb{P}(T_1 = k)\, \mathbb{P}(T_2 = k) = \sum_{k=1}^{n} \frac{1}{n^2} = \frac{1}{n}. \end{aligned}$$
Ainsi, U_2 suit une loi de Bernoulli $\mathcal{B}\left(1 - \frac{1}{n}\right)$.

2. (a) Pour tout $k \in \mathbb{N}^*$, T_k suit une loi uniforme sur $[\![1, n]\!]$ puisque les boules sont indiscernables au toucher et puisque les tirages ont lieu avec remise.

   (b) Soit $k \in \mathbb{N}^*$. En appliquant la formule des probabilités totales au système complet d'événements de probabilités non nulles $((T_k = i))_{i \in [\![1,n]\!]}$, on obtient

$$\begin{aligned}\mathbb{P}(U_k = 1) &= \sum_{i=1}^{n} \mathbb{P}(U_k = 1 \mid T_k = i)\, \mathbb{P}(T_k = i) \\ &= \sum_{i=1}^{n} \mathbb{P}\left(\bigcap_{j=1}^{k-1} (T_j \neq i)\right) \mathbb{P}(T_k = i)\end{aligned}$$

   d'où, par indépendance des T_j :

$$\mathbb{P}(U_k = 1) = \sum_{i=1}^{n} \left(\frac{n-1}{n}\right)^{k-1} \frac{1}{n} = \left(\frac{n-1}{n}\right)^{k-1} = \left(1 - \frac{1}{n}\right)^{k-1}$$

3. (a) $V_N(n)$ est le nombre de variables U_k, avec $k \in [\![1, N]\!]$, qui prennent la valeur 1. C'est donc aussi la somme [5] des variables U_k avec $k \in [\![1, N]\!]$ puisque les U_k qui ne prennent pas la valeur 1 prennent la valeur 0. On a donc

$$V_N(n) = \sum_{k=1}^{N} U_k.$$

   (b) On obtient en utilisant la question précédente

$$\mathbb{E}(V_N(n)) = \sum_{k=1}^{N} \mathbb{E}(U_k)$$

   par linéarité de l'espérance, d'où, en utilisant le fait que les U_k sont des variables de Bernoulli et la question 2(b) :

$$\begin{aligned}\mathbb{E}(V_N(n)) &= \sum_{k=1}^{N} \left(1 - \frac{1}{n}\right)^{k-1} = \sum_{k=0}^{N-1} \left(1 - \frac{1}{n}\right)^{k} \\ &= \frac{1 - \left(1 - \frac{1}{n}\right)^N}{1 - \left(1 - \frac{1}{n}\right)} = n\left(1 - \left(1 - \frac{1}{n}\right)^N\right).\end{aligned}$$

   (c) Lorsque $n \to +\infty$, la quantité ci-dessus donne une forme indéterminée $(\infty \times 0)$ que l'on résout grâce au développement limité de premier ordre $(1+u)^N \underset{u \to 0}{=} 1 + Nu + o(u)$:

$$n\left(1 - \left(1 - \frac{1}{n}\right)^N\right) = n\left(1 - \left(1 - \frac{N}{n} + o\left(\frac{1}{n}\right)\right)\right) = N + o(1) \xrightarrow[n \to +\infty]{} N.$$

   Ce résultat n'a rien de choquant puisque lorsque n est très grand, les N tirages amènent N boules différentes avec haute probabilité.

5. En d'autres termes, $V_N(n)$ est la variable de comptage associée aux variables de Bernoulli $U_1, \ldots, U_N$.

Correction de l'exercice 34.

1. Si le joueur possède N vignettes différentes, il n'est pas nécessaire qu'il en achète pour compléter sa collection puisque celle-ci est déjà complète ; on a donc $s_N = 0$.
2. Soit $i \in [\![0, N-1]\!]$. Considérons que le collectionneur possède i vignettes différentes et notons X le nombre d'achats qui lui seront nécessaires pour compléter sa collection. Posons aussi des variables aléatoires $X_1, \ldots, X_N$ telles que pour tout $j \in [\![1, N]\!]$, X_j représente le nombre de vignettes nécessaires pour que le collectionneur trouve la vignette numéro j parmi les vignettes achetées (même s'il la possédait déjà au départ). Comme les X_j représentent des temps d'attente d'un succès lors de la répétition d'expériences de Bernoulli indépendantes de probabilité de succès $\frac{1}{N}$, elles sont toutes de loi géométrique $\mathcal{G}(\frac{1}{N})$; or il est clair que l'on a
$$0 \leqslant X \leqslant \max(X_1, \ldots, X_N) \leqslant X_1 + \ldots + X_N$$
puisque $\max(X_1, \ldots, X_N)$ représente le nombre d'achats permettant au collectionneur de retrouver l'intégralité de la collection (indépendamment des vignettes déjà possédées au départ). Comme $X_1 + \ldots + X_N$ admet une espérance en tant que somme de VARD admettant une espérance, c'est donc aussi le cas de X ; cette espérance vaut par définition s_i, ce qui justifie la définition donnée dans l'énoncé.
3. Considérons toujours que i est fixé, que le collectionneur possède i vignettes différentes et que X est définie comme dans la question précédente. Soit à présent T la variable aléatoire représentant le nombre d'achats nécessaire au collectionneur pour obtenir *une* vignette qu'il ne possédait pas encore. La variable aléatoire T représente alors le temps d'obtention du premier succès lors de la répétition d'expériences de Bernoulli indépendantes de même probabilité de succès $\frac{N-i}{N}$ (puisque $N-i$ vignettes sont nouvelles pour le collectionneur et puisque les vignettes sont tirées uniformément parmi les N vignettes possibles) ; on a donc $T \sim \mathcal{G}\left(\frac{N-i}{N}\right)$. Or après l'obtention d'une première vignette non possédée, le collectionneur possède $i+1$ vignettes différentes et reprend ses achats : ainsi, la variable $X - T$, qui représente le temps nécessaire à la complétion de la collection après le temps T, suit la loi qu'aurait suivi X si le collectionneur avait initialement possédé $i+1$ vignettes différentes. On a donc $\mathbb{E}(X-T) = s_{i+1}$, si bien que
$$s_i = \mathbb{E}(X) = \mathbb{E}(T + X - T) = \mathbb{E}(T) + \mathbb{E}(X-T) = \frac{N}{N-i} + s_{i+1}$$
par linéarité de l'espérance, ce qu'il fallait démontrer [6].

---

6. Ce raisonnement délicat se résume de manière plus informelle en disant que si le collectionneur possède initialement i vignettes, il doit acheter suffisamment de vignettes pour en trouver une nouvelle (ce qui arrive après un temps en moyenne égal à $\frac{N}{N-i}$) puis pour compléter sa collection depuis une dotation initiale de $i+1$ vignettes différentes (ce qui arrive après un temps en moyenne égal à s_{N+1}).

4. D'après les deux questions précédentes, pour tout $i \in [\![0, N]\!]$ on a

$$\begin{aligned} s_i = s_i - s_N &= \sum_{k=i}^{N-1} (s_k - s_{k+1}) \\ &= \sum_{k=i}^{N-1} \frac{N}{N-k} = \sum_{j=1}^{j=N-i} \frac{N}{j} = N \sum_{j=1}^{N-i} \frac{1}{j}, \end{aligned}$$

où l'on a posé $j = N - k$. Or, à i fixé, on a l'équivalent classique [7]

$$\sum_{j=1}^{N-i} \frac{1}{j} \underset{N\to+\infty}{\sim} \ln(N-i) \underset{N\to+\infty}{\sim} \ln(N),$$

d'où

$$s_i \underset{N\to+\infty}{\sim} N\ln(N).$$

On retiendra par exemple que le temps de complétion d'une collection de N vignettes à partir d'une collection vide est de l'ordre de $N\ln(N)$ lorsque N est important.

Correction de l'exercice 35. D'après le point (iii) de l'exercice 5, X admet une espérance donnée par $\mathbb{E}(X) = e - 1$. On va réaliser un calcul similaire pour déterminer son moment d'ordre 2. Pour tout $k \geqslant 2$, on a

$$\begin{aligned} k^2\,\mathbb{P}(X = k) &= \frac{k^3}{(k+1)!} = \frac{(k+1)k(k-1)+k}{(k+1)!} \\ &= \frac{(k+1)k(k-1)}{(k+1)!} + \frac{k+1}{(k+1)!} - \frac{1}{(k+1)!} = \frac{1}{(k-2)!} + \frac{1}{k!} - \frac{1}{(k+1)!}. \end{aligned}$$

Ainsi, pour tout $n \geqslant 1$ on peut écrire

$$\begin{aligned} \sum_{k=1}^{n} k^2\,\mathbb{P}(X = k) &= \mathbb{P}(X = 1) + \sum_{k=2}^{n} k^2\mathbb{P}(X = k) \\ &= \frac{1}{2} + \sum_{k=2}^{n} \frac{1}{(k-2)!} + \sum_{k=2}^{n} \left(\frac{1}{k!} - \frac{1}{(k+1)!}\right) \\ &= \frac{1}{2} + \sum_{j=0}^{n-2} \frac{1}{j!} + \frac{1}{2} - \frac{1}{(n+1)!} \end{aligned}$$

7. La relation entre la série harmonique et le logarithme qui donne l'équivalent

$$\sum_{j=1}^{N-i} \frac{1}{j} \underset{N\to+\infty}{\sim} \ln(N-i)$$

se démontre par comparaison série-intégrale ou par encadrement (voir par exemple l'exercice 19 du chapitre sur les séries dans le tome de première année). Le deuxième équivalent s'établit en écrivant

$$\frac{\ln(N-i)}{\ln(N)} = \frac{\ln(N) + \ln(N-i) - \ln(N)}{\ln(N)} = 1 + \frac{\ln\left(\frac{N-i}{N}\right)}{\ln(N)} \xrightarrow[N\to+\infty]{} 1.$$

en repérant une somme télescopique et en posant $j = k - 2$, donc

$$\sum_{k=1}^{n} k^2 \mathbb{P}(X = k) \xrightarrow[n \to +\infty]{} e + 1.$$

Ainsi, X^2 admet une espérance et $\mathbb{E}(X^2) = e + 1$.

D'après la formule de Koenig, on a donc

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = e + 1 - (e - 1)^2 = 3e - e^2 = (3 - e)e.$$

Correction de l'exercice 36. Comme les X_k admettent un moment d'ordre 2, c'est aussi le cas de leur somme et on a

$$\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{Cov}(X_i, X_j)$$

d'après la formule du cours. Mais pour tout $(i, j) \in [\![1, n]\!]^2$ on a

$$\operatorname{Cov}(X_i, X_j) \leqslant \sqrt{\mathbb{V}(X_i)}\sqrt{\mathbb{V}(X_j)}$$

d'après l'inégalité de Cauchy-Schwarz, soit $\operatorname{Cov}(X_i, X_j) \leqslant \mathbb{V}(X_1)$ puisque toutes les X_k ont même loi, donc

$$\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{V}(X_1) = n^2 \mathbb{V}(X_1).$$

Remarquons que le cas d'égalité est atteint lorsque toutes les X_k sont égales : dans ce cas, les variations des X_k s'ajoutent les unes aux autres et la variance de leur somme des $\mathbb{V}(nX_1) = n^2\,\mathbb{V}(X_1)$.

Correction de l'exercice 37. Rappelons que si Z est une variable aléatoire suivant une loi de Bernoulli, alors $Z^2 = Z$ puisque Z ne prend que les valeurs 0 et 1.

Les VARD X, X_n et X_{n+1} sont des variables de Bernoulli donc admettent des moments de tous ordres ; ainsi, les quantités $\mathbb{V}(X)$, $\operatorname{Cov}(X, X_n)$ et $\operatorname{Cov}(X, X_{n+1})$ existent.

Comme X prend la valeur 1 si toutes les X_k prennent la valeur 1 et prend la valeur 0 sinon, on a

$$\mathbb{P}(X = 1) = \mathbb{P}(X_1 = 1) \cdots \mathbb{P}(X_n = 1) = p^n$$

par indépendance des X_k, et donc $X = \mathcal{B}(p^n)$. Ainsi, $\mathbb{V}(X) = p^n(1 - p^n)$.

On a ensuite

$$XX_n = X_1 \cdots X_{n-1} X_n^2 = X_1 \cdots X_{n-1} X_n = X$$

puisque $X_n^2 = X_n$, donc,

$$\mathbb{E}(XX_n) = \mathbb{E}(X) = p^n.$$

D'après la formule de Koenig, on a donc

$$\operatorname{Cov}(X, X_n) = \mathbb{E}(XX_n) - \mathbb{E}(X)\mathbb{E}(X_n) = p^n - p^n \cdot p = p^n(1 - p).$$

Conformément à ce que l'on pouvait attendre, cette covariance est positive puisque le fait que X_n prenne une valeur faible (c'est-à-dire nulle) implique que c'est aussi le cas pour X.

Enfin, comme les X_k sont mutuellement indépendantes, $X = X_1 \cdots X_n$ et X_{n+1} le sont aussi par le lemme des coalitions ; ainsi, on a $\mathrm{Cov}(X, X_{n+1}) = 0$.

Correction de l'exercice 38.

1. Les VARD $X + Y$ et XY prennent toutes deux des valeurs élevées quand X et Y prennent des valeurs élevées et prennent des valeurs plus faibles sinon ; on s'attend donc à ce qu'elles soient positivement corrélées, c'est-à-dire que l'on pressent que $\rho_{X+Y,XY} > 0$.
2. Comme X et Y admettent un moment d'ordre 2, et donc une variance, c'est aussi le cas de $X + Y$, et par indépendance de X et Y, on a

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) = \frac{2(1-p)}{p^2}.$$

On a par ailleurs $(XY)^2 = X^2Y^2$, or X^2 et Y^2 sont indépendantes d'après le lemme des coalitions donc $(XY)^2$ est le produit de deux VARD admettant une espérance, donc elle admet elle-même une espérance qui vaut

$$\mathbb{E}((XY)^2) = \mathbb{E}(X^2)\,\mathbb{E}(Y^2).$$

Or la formule de Koenig donne[8]

$$\mathbb{E}(X^2) = \mathbb{V}(X) + \mathbb{E}(X)^2 = \frac{1-p}{p^2} + \frac{1}{p^2} = \frac{2-p}{p^2},$$

et de même $\mathbb{E}(Y^2) = \frac{2-p}{p^2}$, d'où

$$\mathbb{E}((XY)^2) = \left(\frac{2-p}{p^2}\right)^2.$$

Ainsi, la variance de XY (qui existe puisque XY admet un moment d'ordre 2) est égale, d'après la formule de Koenig, à

$$\begin{aligned}\mathbb{V}(XY) &= \mathbb{E}((XY)^2) - \mathbb{E}(XY)^2 = \mathbb{E}(X^2Y^2) - \mathbb{E}(XY)^2 \\ &= \mathbb{E}(X^2Y^2) - (\mathbb{E}(X)\mathbb{E}(Y))^2 \quad \text{par indépendance de } X \text{ et } Y \\ &= \left(\frac{2-p}{p^2}\right)^2 - \left(\frac{1}{p}\right)^2\left(\frac{1}{p}\right)^2 = \frac{(1-p)(3-p)}{p^4}.\end{aligned}$$

Enfin, on a

$$\mathrm{Cov}(X+Y, XY) = \mathrm{Cov}(X, XY) + \mathrm{Cov}(Y, XY) = 2\,\mathrm{Cov}(X, XY)$$

8. On utilise ici la formule de Koenig pour retrouver $\mathbb{E}(X^2)$ à partir de $\mathbb{E}(X)$ et $\mathbb{V}(X)$ dont on connaît l'expression par cœur. Cette astuce un peu contre-intuitive (puisque la valeur de $\mathbb{V}(X)$ a découlé du calcul de $\mathbb{E}(X^2)$ pour la plupart des lois, dont la loi géométrique étudiée ici) est néanmoins très courante.

où la dernière égalité provient de la symétrie du problème en les variables X et Y (qui ont même loi), d'où, en utilisant la formule de Koenig ainsi que l'indépendance de X^2 et Y d'une part, et celle de X et Y d'autre part :

$$\begin{aligned}\operatorname{Cov}(X+Y, XY) &= 2(\mathbb{E}(XXY) - \mathbb{E}(X)\,\mathbb{E}(XY)) \\ &= 2\,(\mathbb{E}(X^2)\,\mathbb{E}(Y) - \mathbb{E}(X)^2\,\mathbb{E}(Y) \\ &= 2\,(\mathbb{E}(X^2) - \mathbb{E}(X)^2)\,\mathbb{E}(Y) \\ &= 2\,\mathbb{V}(X)\,\mathbb{E}(Y) = 2 \cdot \frac{1-p}{p^2} \cdot \frac{1}{p} \\ &= \frac{2(1-p)}{p^3}.\end{aligned}$$

On remarque que cette covariance est strictement positive : les VARD $X+Y$ et XY sont bien positivement corrélées. La valeur de leur coefficient de corrélation est égale à

$$\rho_{X+Y,XY} = \frac{\operatorname{Cov}(X+Y, XY)}{\sqrt{\mathbb{V}(X+Y)}\sqrt{\mathbb{V}(XY)}} = \frac{\frac{2(1-p)}{p^3}}{\sqrt{\frac{2(1-p)}{p^2}}\sqrt{\frac{(1-p)(3-p)}{p^4}}} = \sqrt{\frac{2}{3-p}}.$$

Lorsque p est proche de 1, les VARD X et Y prennent la valeur 1 avec haute probabilité. Lorsque ce n'est pas le cas, la valeur de $X+Y$ est fortement indicative de celle de XY : par exemple, lorsque X prend une valeur différente de 1, les VARD $X+Y$ et XY prennent avec haute probabilité (en l'occurrence $\mathbb{P}(Y=1)$) les mêmes valeurs que $X+1$ et X, donc des valeurs simultanément hautes ou simultanément basses. Cela explique que la corrélation entre $X+Y$ et XY soit particulièrement forte lorsque $p \to 1$, ce qu'exprime la relation $\rho_{X+Y,XY} \underset{p \to 1}{\longrightarrow} 1$.

Correction de l'exercice 39.

1. Les boîtes contenant initialement N allumettes, le nombre d'allumettes présentes dans la boîte qui n'est pas celle dans laquelle le fumeur pioche pour la dernière fois est un entier entre 0 et N. Par ailleurs, pour tout $i \in [\![0, N]\!]$ il est possible d'imaginer une situation dans laquelle X_N prend la valeur i : il suffit que le fumeur choisisse $N-i$ allumettes dans sa poche gauche puis N allumettes dans sa poche droite, puis tente à nouveau de piocher dans sa poche droite. On a donc $X_N(\Omega) = [\![0, N]\!]$.
2. Pour tout $j \in \mathbb{N}^*$, la VARD Z_j est égale au nombre de succès obtenus lors de la répétition de j expériences de Bernoulli indépendantes de probabilité de succès (c'est-à-dire de probabilité de piocher à gauche) égale à $\frac{1}{2}$, donc $Z_j \sim \mathcal{B}(j, \frac{1}{2})$.
3. Soit $k \in [\![0, N]\!]$. L'événement $(X_N = k)$ est réalisé dans les deux cas suivants (et uniquement dans ces deux cas) :
   - Le fumeur, pour allumer ses $N+N-k = 2N-k$ premières pipes, a pioché N allumettes à gauche et $N-k$ à droite, et il allume sa $(2N-k+1)$-ième pipe en piochant une allumette à gauche (dans ce cas, il trouve la boîte de gauche vide et il reste k allumettes à droite).

- Le fumeur, pour allumer ses $2N-k$ premières pipes, a pioché N allumettes à droite et $N-k$ à gauche, et il allume sa $(2N-k+1)$-ième pipe en piochant une allumette à droite (dans ce cas, il trouve la boîte de droite vide et il reste k allumettes à gauche).

On a donc

$$(X_N = k) = \Big((Z_{2N-k} = N) \cap G_{2N-k+1}\Big) \sqcup \Big((Z_{2N-k} = N-k) \cap \overline{G_{2N-k+1}}\Big).$$

Ainsi, on a par incompatibilité, puis par indépendance des tirages :

$$\begin{aligned}
\mathbb{P}(X_N = k) &= \mathbb{P}\left((Z_{2N-k} = N) \cap G_{2N-k+1}\right) + \mathbb{P}\left((Z_{2N-k} = N-k) \cap \overline{G_{2N-k+1}}\right) \\
&= \mathbb{P}(Z_{2N-k} = N)\,\mathbb{P}(G_{2N-k+1}) + \mathbb{P}(Z_{2N-k} = N-k)\,\mathbb{P}(\overline{G_{2N-k+1}}) \\
&= \binom{2N-k}{N}\frac{1}{2^{2N-k}} \cdot \frac{1}{2} + \binom{2N-k}{N-k}\frac{1}{2^{2N-k}} \cdot \frac{1}{2},
\end{aligned}$$

or $\binom{2N-k}{N} = \binom{2N-k}{N-k}$ par symétrie des coefficients binomiaux, d'où

$$\mathbb{P}(X_N = k) = \binom{2N-k}{N}\frac{1}{2^{2N-k}}.$$

4. Soit $k \in [\![0, N-1]\!]$. D'après la question précédente, on a alors

$$\begin{aligned}
\frac{\mathbb{P}(X_N = k+1)}{\mathbb{P}(X_N = k)} &= \frac{\binom{2N-k-1}{N}\frac{1}{2^{2N-k-1}}}{\binom{2N-k}{N}\frac{1}{2^{2N-k}}} = \frac{2\binom{2N-k-1}{N}}{\binom{2N-k}{N}} \\
&= \frac{2\frac{(2N-k-1)!}{N!(N-k-1)!}}{\frac{(2N-k)!}{N!(N-k)!}} = \frac{2(N-k)}{2N-k}.
\end{aligned}$$

Ainsi, on a bien

$$(2N-k)\,\mathbb{P}(X_N = k+1) = 2(N-k)\,\mathbb{P}(X_N = k).$$

5. Sommons la relation précédente pour $k \in [\![0, N-1]\!]$. On obtient

$$\sum_{k=0}^{N-1}(2N-k)\,\mathbb{P}(X_N = k+1) = \sum_{k=0}^{N-1} 2(N-k)\,\mathbb{P}(X_N = k)$$

soit, en posant $j = k+1$:

$$\sum_{j=1}^{N}(2N-j+1)\,\mathbb{P}(X_N = j) = \sum_{k=0}^{N-1} 2(N-k)\,\mathbb{P}(X_N = k).$$

En utilisant la linéarité de la somme de part et d'autre, on obtient

$$(2N+1)\sum_{j=1}^{N}\mathbb{P}(X_N = j) - \sum_{j=1}^{N} j\,\mathbb{P}(X_N = j) = 2N\sum_{k=0}^{N-1}\mathbb{P}(X_N = k) - 2\sum_{k=0}^{N-1} k\,\mathbb{P}(X_N = k),$$

or $\displaystyle\sum_{k=0}^{N} \mathbb{P}(X_N = k) = 1$ et $\displaystyle\mathbb{E}(X_N) = \sum_{k=0}^{N} k\,\mathbb{P}(X_N = k) = \sum_{k=1}^{N} k\,\mathbb{P}(X_N = k)$ donc

$$\begin{aligned}(2N+1)\big(1 - \mathbb{P}(X_N = 0)\big) &- \mathbb{E}(X_N) \\ &= 2N(1 - \mathbb{P}(X_N = N)) - 2\Big(\mathbb{E}(X_N) - N\mathbb{P}(X_N = N)\Big)\end{aligned}$$

soit

$$(2N+1)\left(1 - \binom{2N}{N}\frac{1}{2^{2N}}\right) - \mathbb{E}(X_N) = 2N\left(1 - \frac{1}{2^N}\right) - 2\left(\mathbb{E}(X_N) - \frac{N}{2^N}\right)$$

donc

$$\mathbb{E}(X_N) = 2N\left(1 - \frac{1}{2^N}\right) + \frac{2N}{2^N} - (2N+1)\left(1 - \binom{2N}{N}\frac{1}{2^{2N}}\right)$$

d'où enfin

$$\mathbb{E}(X_N) = \binom{2N}{N}\frac{2N+1}{2^{2N}} - 1.$$

Correction de l'exercice 40.

1. Soient $n \in \mathbb{N}$ et $k \in [\![0, n]\!]$. Conditionnellement à $(X = n)$, c'est-à-dire au fait que n œufs ont été pondus, le nombre d'œufs éclos suit une loi binomiale $\mathcal{B}(n, p)$: en effet, il s'agit du nombre de succès obtenus lors de la répétition de n épreuves de Bernoulli indépendantes de probabilité de réussite p. On a donc

$$\mathbb{P}(Y = k \mid X = n) = \binom{n}{k} p^k (1-p)^{n-k}.$$

2. On a $Y(\Omega) = \mathbb{N}$ puisque le nombre d'œufs pondus est un entier positif et puisqu'il est possible pour tout $k \in \mathbb{N}$ d'envisager une situation de probabilité strictement positive telle que k œufs éclosent (par exemple la situation dans laquelle k œufs sont pondus et tous éclosent). Si $k \in \mathbb{N}$, on a d'après la formule des probabilités totales appliquée au système complet d'événements $\big((X = n)\big)_{n \in \mathbb{N}}$:

$$\begin{aligned}\mathbb{P}(Y = k) &= \sum_{n=0}^{+\infty} \mathbb{P}(X = n)\,\mathbb{P}(Y = k \mid X = n) \\ &= \sum_{n=0}^{+\infty} e^{-\lambda}\frac{\lambda^n}{n!}\mathbb{P}(Y = k \mid X = n).\end{aligned}$$

Or d'après le résultat de la question précédente, pour tous $k, n \in \mathbb{N}$, on a

$$\mathbb{P}(Y = k \mid X = n) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } k \leqslant n \\ 0 & \text{sinon.}\end{cases}$$

On en déduit que pour tout $k \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{P}(Y=k) &= \sum_{n=k}^{+\infty} e^{-\lambda}\frac{\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k} \\
&= e^{-\lambda}p^k\sum_{n=k}^{+\infty}\frac{\lambda^n}{k!(n-k)!}(1-p)^{n-k} \\
&= e^{-\lambda}\frac{1}{k!}p^k\sum_{j=0}^{+\infty}\frac{\lambda^{j+k}}{j!}(1-p)^j \quad \text{en posant } j=n-k \\
&= e^{-\lambda}\frac{\lambda^k}{k!}p^k\sum_{j=0}^{+\infty}\frac{(\lambda(1-p))^j}{j!} \\
&= e^{-\lambda}\frac{(\lambda p)^k}{k!}e^{\lambda(1-p)} = e^{-\lambda p}\frac{(\lambda p)^k}{k!},
\end{aligned}$$

donc $Y \sim \mathcal{P}(\lambda p)$.

On est en présence d'un *processus de Poisson marqué* : parmi les œufs pondus selon un flux d'intensité λ, une proportion p est vouée à éclore, ce qui correspond à des éclosions selon un flux d'intensité λp, d'où la loi de Y.

3. Par la même logique que dans la question précédente, le nombre d'œufs non éclos $X-Y$ suit la loi $\mathcal{P}(\lambda(1-p))$ puisque la probabilité de non-éclosion d'un œuf donné est $1-p$.

   Soient à présent $i, j \in \mathbb{N}$. On a alors

$$\begin{aligned}
\mathbb{P}((Y=i)\cap(X-Y=j)) &= \mathbb{P}((Y=i)\cap(X=i+j)) \\
&= \mathbb{P}(X=i+j)\,\mathbb{P}(Y=i\,|\,X=i+j) \\
&= e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}\binom{i+j}{i}p^i(1-p)^j
\end{aligned}$$

   puisque la loi de Y conditionnellement à $(X=i+j)$ est une loi binomiale de paramètres $i+j$ et p, si bien que

$$\mathbb{P}((Y=i)\cap(X-Y=j)) = e^{-\lambda}\frac{\lambda^{i+j}}{i!j!}p^i(1-p)^j.$$

   Or

$$\mathbb{P}(Y=i)\,\mathbb{P}(X-Y=j) = e^{-\lambda p}\frac{(\lambda p)^i}{i!}e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}$$

   puisque $Y \sim \mathcal{P}(\lambda p)$ et $X-Y \sim \mathcal{P}(\lambda(1-p))$, donc

$$\begin{aligned}
\mathbb{P}(Y=i)\,\mathbb{P}(X-Y=j) &= e^{-\lambda}\frac{\lambda^{i+j}}{i!j!}p^i(1-p)^j \\
&= \mathbb{P}((Y=i)\cap(X-Y=j)).
\end{aligned}$$

   Les variables aléatoires Y et $X-Y$ sont donc bien indépendantes.

4. On aurait pu s'attendre à l'existence d'une corrélation non nulle entre Y et $X-Y$ pour deux raisons.

- Le nombre d'œufs pondus étant aléatoire, on s'attendrait à ce qu'un nombre élevé d'œufs non éclos soit indicatif d'un nombre élevé d'œufs pondus et donc d'un nombre élevé d'œufs éclos, et vice versa.
- Inversement, pour un nombre d'œufs pondus donné[9], tout œuf qui éclot est compté par la variable aléatoire Y et n'apparaît pas dans le compte de la variable aléatoire $X - Y$, donc une valeur élevée de Y implique mécaniquement une valeur faible de $X - Y$ et réciproquement.

Le résultat d'indépendance obtenu montre que ces deux logiques agissant dans des sens opposés s'équilibrent parfaitement dans le cas où le flux d'œufs pondus suit une loi de Poisson.

Correction de l'exercice 41.

1. Il suffit d'appliquer la propriété d'absence de mémoire avec $m = 1$: pour tout $n \in \mathbb{N}^*$, on a $\mathbb{P}(X > n+1 \,|\, X > n) = \mathbb{P}(X > 1)$, or
$$\mathbb{P}(X > n+1 \,|\, X > n) = \frac{\mathbb{P}((X > n+1) \cap (X > n))}{\mathbb{P}(X > n)} = \frac{\mathbb{P}(X > n+1)}{\mathbb{P}(X > n)}$$
puisque $(X > n+1) \subset (X > n)$, d'où le résultat attendu.
2. D'après la question précédente, la suite $(\mathbb{P}(X > n))_{n \in \mathbb{N}^*}$ est géométrique de raison $\mathbb{P}(X > 1)$. Comme son terme d'indice 1 est $\mathbb{P}(X > 1)$, on a :
$$\forall n \in \mathbb{N}^*, \quad \mathbb{P}(X > n) = \mathbb{P}(X > 1)^{n-1}\mathbb{P}(X > 1) = \mathbb{P}(X > 1)^n.$$
On remarque que cette relation est aussi valide pour $n = 0$, ce qui sera utile dans la question suivante.
3. Comme X est à valeurs entières, pour tout $k \in \mathbb{N}^*$ on a d'après la question précédente
$$\begin{aligned}\mathbb{P}(X = k) &= \mathbb{P}(X > k-1) - \mathbb{P}(X > k) \\ &= \mathbb{P}(X > 1)^{k-1} - \mathbb{P}(X > 1)^k \\ &= \mathbb{P}(X > 1)^{k-1}(1 - \mathbb{P}(X > 1)),\end{aligned}$$
ce qui montre que X suit une loi géométrique de paramètre[10] $1 - \mathbb{P}(X > 1)$, c'est-à-dire à $\mathbb{P}(X = 1)$.

Correction de l'exercice 42.

1. On a clairement $X_A(\Omega) = [\![1, 10]\!]$ puisque le gardien à jeun peut ouvrir la porte en n'importe quel nombre entier d'essais entre 1 et 10. Pour tout $k \in [\![1, 10]\!]$, l'événement $(X_A = k)$ est réalisé si les $k - 1$ premiers essais du gardien sont

---

9. Si X, au lieu de suivre une loi de Poisson, était une quantité déterministe égale à un entier N, alors Y et $X - Y = N - Y$ seraient linéairement corrélées et on aurait $\rho_{Y,X-Y} = -1$.

10. Remarquons que ce fait est raisonnable puisque réciproquement, si X suit une loi géométrique de paramètre $p \in]0, 1[$, alors ce paramètre p est précisément égal à la probabilité $\mathbb{P}(X = 1)$

infructueux et si son k-ième essai conduit à utiliser la clé adaptée. Introduisons à présent, pour tout $i \in [\![1, 10]\!]$, l'événement A_i : « le i-ième essai est un succès ». Pour tout $k \in [\![1, 10]\!]$, on obtient donc, grâce à la formule des probabilités composées [11] :

$$\begin{aligned}
\mathbb{P}(X_A = k) &= \mathbb{P}\left(\overline{A_1} \cap \ldots \cap \overline{A_{k-1}} \cap A_k\right) \\
&= \mathbb{P}(\overline{A_1})\,\mathbb{P}(\overline{A_2} \mid \overline{A_1}) \cdots \mathbb{P}(\overline{A_{k-1}} \mid \overline{A_1} \cap \ldots \cap \overline{A_{k-2}})\,\mathbb{P}(A_k \mid \overline{A_1} \cap \ldots \cap \overline{A_{k-1}}) \\
&= \frac{9}{10} \cdot \frac{8}{9} \cdots \frac{10-(k-1)}{10-(k-2)} \cdot \frac{1}{10-(k-1)} \\
&= \frac{1}{10}.
\end{aligned}$$

Ainsi, X_A suit la loi $\mathcal{U}([\![1, 10]\!])$. On a ainsi $\mathbb{E}(X_A) = \frac{1+10}{2} = \frac{11}{2}$.

Remarquons qu'il s'agit d'un résultat raisonnable : comme les clés sont essayées dans un ordre aléatoire, il n'est pas plus probable que la bonne clé arrive en première position qu'en n'importe quelle autre position.

2. La variable aléatoire X_B représente le temps d'attente du premier succès lors de la répétition d'expériences de Bernoulli indépendantes de probabilité de succès $\frac{1}{10}$; on a donc $X \sim \mathcal{G}(\frac{1}{10})$ et $\mathbb{E}(X) = 10$.
3. Notons I l'événement « le gardien est ivre » et X la variable aléatoire donnant le nombre d'essais nécessaires au gardien pour ouvrir la porte. Conditionnellement à I, la VARD X suit la même loi que X_B, et conditionnellement à $\overline{I}$, elle suit la même loi que X_A. Or la probabilité que l'on souhaite calculer se réécrit grâce à la formule de Bayes : il s'agit de

$$\mathbb{P}(I \mid X > 8) = \frac{\mathbb{P}(I \cap (X > 8))}{\mathbb{P}(X > 8)} = \frac{\mathbb{P}(X > 8 \mid I)\,\mathbb{P}(I)}{\mathbb{P}(X > 8 \mid I)\,\mathbb{P}(I) + \mathbb{P}(X > 8 \mid \overline{I})\,\mathbb{P}(\overline{I})},$$

où l'on a utilisé la formule des probabilités totales sur le système complet d'événements $(I, \overline{I})$. On a donc

$$\begin{aligned}
\mathbb{P}(I \mid X > 8) &= \frac{\mathbb{P}(X_B > 8)\,\mathbb{P}(I)}{\mathbb{P}(X_B > 8)\,\mathbb{P}(I) + \mathbb{P}(X_A > 8)\,\mathbb{P}(\overline{I})} \\
&= \frac{\left(\frac{9}{10}\right)^8 \cdot \frac{1}{4}}{\left(\frac{9}{10}\right)^8 \cdot \frac{1}{4} + \frac{2}{10} \cdot \frac{3}{4}} \\
&= \frac{14\,348\,307}{34\,348\,907} \\
&= 0{,}4177\ldots
\end{aligned}$$

Ce résultat est conforme à l'intuition : les nombreux échecs du gardien rendent crédible l'hypothèse selon laquelle il est ivre, ce qui explique que la probabilité obtenue après avoir observé l'événement $(X > 8)$ soit significativement supérieure à la probabilité d'ivresse *a priori*, égale à $\frac{1}{4}$.

---

11. L'exercice proposé est un sujet d'oral. Lors de la présentation d'un tel sujet, il n'est pas nécessaire de recourir au formalisme de la formule des probabilités composées dans toute sa rigueur et il peut suffire de détailler oralement le calcul de $\mathbb{P}(X_A = 1)$, $\mathbb{P}(X_A = 2)$ et $\mathbb{P}(X_A = 3)$ en prenant soin de justifier les différents facteurs qui interviennent dans les produits, puis de généraliser le résultat à $\mathbb{P}(X_A = k)$.

Correction de l'exercice 43.

1. L'une contient initialement deux boules, et on lui rajoute deux boules par étape. Ainsi, pour tout $n \geqslant 1$, le nombre de boules dans l'urne à l'issue de l'étape n est $2 + 2n = 2(n+1)$.
2. L'urne contient initialement une boule verte, et à chaque étape on lui ajoute 0 ou 1 boule verte ; ainsi, pour tout $n \geqslant 0$ on a $V_n(\Omega) = [\![1, n+1]\!]$.
3. Soient $n \geqslant 1$ et $k \in [\![1, n]\!]$. Conditionnellement à l'événement $(V_{n-1} = k)$, au terme de la $(n-1)$-ième étape l'urne contient $2n$ boules dont k boules vertes, si bien que la probabilité d'y tirer une boule verte au début de l'étape suivante est égale à $\frac{k}{2n}$ par équiprobabilité. Ainsi, on a
$$\mathbb{P}(E_n \,|\, V_{n-1} = k) = \frac{k}{2n}.$$
4. Soit $n \geqslant 1$. En utilisant la formule des probabilités totales sur le système complet d'événements $(V_{n-1} = k)_{k \in [\![1,n]\!]}$, on a, en utilisant la question précédente :
$$\begin{aligned}\mathbb{P}(E_n) &= \sum_{k=1}^{n} \mathbb{P}(E_n \,|\, V_{n-1} = k)\, \mathbb{P}(V_{n-1} = k) \\ &= \sum_{k=1}^{n} \frac{k}{2n} \mathbb{P}(V_{n-1} = k) \\ &= \frac{1}{2n} \sum_{k=1}^{n} k\, \mathbb{P}(V_{n-1} = k) \\ &= \frac{1}{2n} \mathbb{E}(V_{n-1}),\end{aligned}$$
ce qui est le résultat attendu.
5. Soit $n \geqslant 1$. Conformément à l'indication, on considère l'indicatrice $\mathbb{1}_{E_n}$, c'est-à-dire la variable aléatoire qui prend la valeur 1 si E_n est réalisé et 0 sinon ; on a alors $\mathbb{1}_{E_n} \sim \mathcal{B}(\mathbb{P}(E_n))$. Or le lien entre V_{n-1} et V_n est donné par
$$V_n = V_{n-1} + \mathbb{1}_{E_n}$$
puisque V_n prend la même valeur que V_{n-1} si E_n n'est pas réalisé et la même valeur que $V_{n-1} + 1$ si E_n est réalisé. Ainsi, on a par linéarité de l'espérance
$$\mathbb{E}(V_n) = \mathbb{E}(V_{n-1}) + \mathbb{E}(\mathbb{1}_{E_n}) = \mathbb{E}(V_{n-1}) + \mathbb{P}(E_n),$$
soit
$$\mathbb{E}(V_n) = \mathbb{E}(V_{n-1}) \left(1 + \frac{1}{2n}\right)$$
d'après la question précédente.
6. La relation de récurrence obtenue dans la question précédente se réécrit de la façon suivante :
$$\forall n \geqslant 1, \quad \mathbb{E}(V_n) = \frac{2n+1}{2n}\, \mathbb{E}(V_{n-1}).$$

En calculant $\mathbb{E}(V_1) = \frac{3}{2}$ (puisque V_1 prend les valeurs 1 et 2 de façon équiprobable), $\mathbb{E}(V_2) = \frac{5}{4}\mathbb{E}(V_1) = \frac{5}{4} \cdot \frac{3}{2}$ puis $\mathbb{E}(V_3) = \frac{7}{6}\mathbb{E}(V_2) = \frac{7}{6} \cdot \frac{5}{4} \cdot \frac{3}{2}$, on acquiert l'intuition de la formule suivante[12] :

$$\forall n \geqslant 0, \quad \mathbb{E}(V_n) = \prod_{k=1}^{n} \frac{2k+1}{2k},$$

que l'on démontre aisément par récurrence (voir l'exercice 26 du chapitre 14 du tome de première année, qui porte sur les intégrales de Wallis, pour un calcul similaire).

On peut réécrire cette formule par un stratagème classique, en complétant le numérateur pour faire apparaître une factorielle et en remarquant que le dénominateur se calcule par factorisation par 2 : pour tout $n \geqslant 1$,

$$\begin{aligned}
\mathbb{E}(V_n) &= \frac{(2n+1)(2n-1)\cdots 3 \cdot 1}{2n(2n-2)\cdots 4 \cdot 2} \\
&= \frac{(2n+1)2n(2n-1)\cdots 3 \cdot 2 \cdot 1}{(2n(2n-2)\cdots 4 \cdot 2)^2} \\
&= \frac{(2n+1)!}{(2^n \cdot n(n-1)\cdots 2 \cdot 1)^2} \\
&= \frac{(2n+1)!}{2^{2n} n!^2}.
\end{aligned}$$

7. La formule de Stirling donne

$$(2n+1)! \underset{n\to+\infty}{\sim} \sqrt{2\pi(2n+1)} \left(\frac{2n+1}{e}\right)^{2n+1}$$

et

$$n!^2 \underset{n\to+\infty}{\sim} 2\pi n \left(\frac{n}{e}\right)^{2n},$$

donc

$$\begin{aligned}
\mathbb{E}(V_n) = \frac{(2n+1)!}{2^{2n} n!^2} &\underset{n\to+\infty}{\sim} \frac{\sqrt{2\pi(2n+1)} \left(\frac{2n+1}{e}\right)^{2n+1}}{2^{2n} \cdot 2\pi n \left(\frac{n}{e}\right)^{2n}} \\
&= \frac{\sqrt{2\pi(2n+1)}}{2\pi n} \left(\frac{2n+1}{2n}\right)^{2n} \frac{2n+1}{e} \\
&\underset{n\to+\infty}{\sim} \frac{\sqrt{2\pi \cdot 2n}}{2\pi n} \left(\frac{2n+1}{2n}\right)^{2n} \cdot \frac{2n}{e} \\
&= \frac{2\sqrt{n}}{\sqrt{\pi} e} \left(\frac{2n+1}{2n}\right)^{2n}.
\end{aligned}$$

12. On rappelle que par convention, le produit $\displaystyle\prod_{k=1}^{0} \frac{2k+1}{2k}$ vaut 1.

Or

$$\left(\frac{2n+1}{2n}\right)^{2n} = \left(1+\frac{1}{2n}\right)^{2n} = e^{2n\ln\left(1+\frac{1}{2n}\right)}$$
$$\underset{n\to+\infty}{=} e^{2n\left(\frac{1}{2n}+o\left(\frac{1}{2n}\right)\right)} = e^{1+o(1)} \xrightarrow[n\to+\infty]{} e$$

grâce au développement limité $\ln(1+x) \underset{x\to 0}{=} x + o(x)$ et par continuité de l'exponentielle, d'où

$$\mathbb{E}(V_n) \underset{n\to+\infty}{\sim} \frac{2\sqrt{n}}{\sqrt{\pi}e}\left(\frac{2n+1}{2n}\right)^{2n} \underset{n\to+\infty}{\sim} \frac{2\sqrt{n}}{\sqrt{\pi}e}\cdot e = 2\sqrt{\frac{n}{\pi}},$$

ce qui est le résultat attendu avec $C = \dfrac{2}{\sqrt{\pi}}$.

Correction de l'exercice 44.

1. Les variables aléatoires B et H étant à valeurs dans $\mathbb{N}^*$ et indépendantes, la probabilité recherchée est
$$\begin{aligned}\mathbb{P}(BH=2) &= \mathbb{P}\Big(\big((B=1)\cap(H=2)\big)\sqcup\big((B=2)\cap(H=1)\big)\Big)\\ &= \mathbb{P}((B=1)\cap(H=2)) + \mathbb{P}((B=2)\cap(H=1)) \quad \text{par incompatibilité}\\ &= \mathbb{P}(B=1)\,\mathbb{P}(H=2) + \mathbb{P}(B=2)\,\mathbb{P}(H=1) \quad \text{par indépendance}\\ &= p\cdot(1-p)p + (1-p)p\cdot p\\ &= 2p^2(1-p).\end{aligned}$$

2. Les VARD B et H admettent une espérance égale à $\frac{1}{p}$ et sont indépendantes, donc l'aire BH du rectangle admet une espérance donnée par
$$\mathbb{E}(BH) = \mathbb{E}(B)\,\mathbb{E}(H) = \frac{1}{p^2}.$$

3. La probabilité d'obtenir un carré est la probabilité de l'événement $(B=H)$, qui vaut
$$\begin{aligned}\mathbb{P}(B=H) &= \mathbb{P}\left(\bigsqcup_{k=1}^{+\infty}(B=k)\cap(H=k)\right)\\ &= \sum_{k=1}^{+\infty}\mathbb{P}((B=k)\cap(H=k)) \quad \text{par incompatibilité}\\ &= \sum_{k=1}^{+\infty}\mathbb{P}(B=k)\,\mathbb{P}(H=k) \quad \text{par indépendance}\\ &= \sum_{k=1}^{+\infty}(1-p)^{k-1}p(1-p)^{k-1}p\\ &= p^2\sum_{j=0}^{+\infty}((1-p)^2)^j \quad \text{en posant } j=k-1\\ &= p^2\cdot\frac{1}{1-(1-p)^2} \quad \text{car } (1-p)^2\in[0,1[\\ &= \frac{p}{2-p}.\end{aligned}$$

Correction de l'exercice 45.

1. Pour tout $k \in [\![1,n]\!]$, on a $Z_k \sim \mathcal{U}([\![1,N]\!])$, et les Z_k sont indépendantes puisque les tirages ont lieu avec remise.

   Les variables aléatoires S_n et T_n admettent une espérance et une variance en tant que VARD bornées (ou en tant que combinaisons linéaires des Z_k qui admettent un moment d'ordre 2). Par linéarité de l'espérance, on a

   $$\mathbb{E}(S_n) = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(Z_k) = \frac{1}{n}\sum_{k=1}^{n}\frac{N+1}{2} = \frac{N+1}{2}$$

   donc

   $$\mathbb{E}(T_n) = 2\,\mathbb{E}(S_n) - 1 = N.$$

   On a par ailleurs

   $$\mathbb{V}(S_n) = \frac{1}{n^2}\mathbb{V}\left(\sum_{k=1}^{n} Z_k\right) = \frac{1}{n^2}\sum_{k=1}^{n}\mathbb{V}(Z_k)$$

   par indépendance des Z_k, soit, comme les Z_k suivent une même loi uniforme sur $[\![1,N]\!]$:

   $$\mathbb{V}(S_n) = \frac{1}{n}\mathbb{V}(Z_1) = \frac{N^2-1}{12n},$$

   d'où

   $$\mathbb{V}(T_n) = \mathbb{V}(2S_n - 1) = 4\mathbb{V}(S_n) = \frac{N^2-1}{3n}.$$

2. On a $M_n(\Omega) = [\![1,N]\!]$ puisque M_n, comme les Z_k, peut prendre n'importe quelle valeur entière entre 1 et N. Pour tout $i \in [\![0,N]\!]$, on a

   $$\begin{aligned}
   \mathbb{P}(M_n \leqslant i) &= \mathbb{P}\left(\bigcap_{k=1}^{n}(Z_k \leqslant i)\right) \\
   &= \prod_{k=1}^{n}\mathbb{P}(Z_k \leqslant i) \quad \text{par indépendance des } Z_k \\
   &= \prod_{k=1}^{n}\frac{i}{N} \\
   &= \left(\frac{i}{N}\right)^n,
   \end{aligned}$$

   donc pour tout $j \in [\![1,N]\!]$ on peut écrire :

   $$\mathbb{P}(Z = j) = \mathbb{P}(Z \leqslant j) - \mathbb{P}(Z \leqslant j-1) = \left(\frac{j}{N}\right)^n - \left(\frac{j-1}{N}\right)^n.$$

3. Comme M_n prend un nombre fini de valeurs, elle admet une espérance donnée par

   $$\mathbb{E}(M_n) = \sum_{k=1}^{N} k\,\mathbb{P}(Z = k)$$

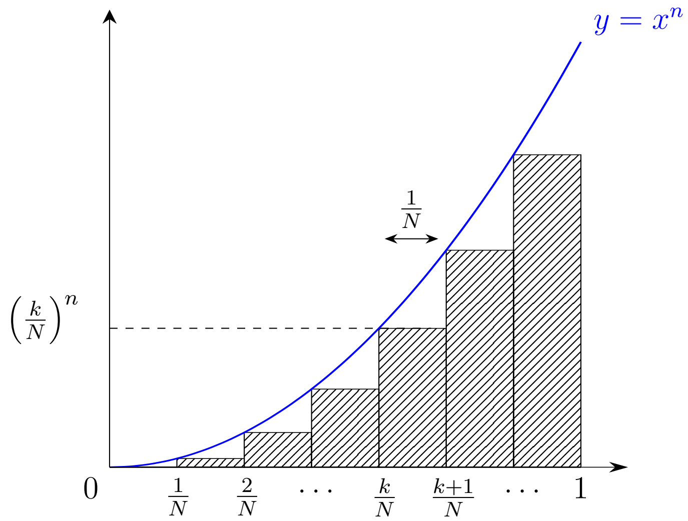
$$\begin{aligned}
&= \sum_{k=1}^{N} k\left(\left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n\right) \\
&= \sum_{k=1}^{N} \left[(k+1)\left(\frac{k}{N}\right)^n - k\left(\frac{k-1}{N}\right)^n\right] - \sum_{k=1}^{N}\left(\frac{k}{N}\right)^n \\
&= (N+1) - \sum_{k=1}^{N}\left(\frac{k}{N}\right)^n \quad \text{en repérant une somme télescopique} \\
&= N - \sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n,
\end{aligned}$$

ce qu'il fallait démontrer.

4. D'après la question précédente, il suffit d'établir l'inégalité

$$\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n \leqslant \frac{N}{n+1}.$$

Pour ce faire, on va comparer la somme $\frac{1}{N}\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n$ (qui est l'aire totale des rectangles de largeur $\frac{1}{N}$ situés sous la courbe de la fonction $t \mapsto t^n$ entre 0 et 1) avec l'intégrale $\int_0^1 t^n \mathrm{d}t$ (qui est l'aire de la surface comprise sous cette courbe) :



Pour tout $k \in [\![1, N-1]\!]$ et tout $t \in \left[\frac{k}{N}, \frac{k+1}{N}\right[$, on a

$$\left(\frac{k}{N}\right)^n \leqslant t^n$$

par croissance de la fonction $t \mapsto t^n$ sur $\mathbb{R}_+$, donc, en intégrant cette inégalité sur $\left[\frac{k}{N}, \frac{k+1}{N}\right[$:

$$\frac{1}{N}\left(\frac{k}{N}\right)^n \leqslant \int_{\frac{k}{N}}^{\frac{k+1}{N}} t^n \,\mathrm{d}t.$$

On somme ensuite l'inégalité obtenue pour $k \in [\![1, N-1]\!]$, ce qui donne

$$\frac{1}{N}\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n \leqslant \sum_{k=1}^{N-1}\int_{\frac{k}{N}}^{\frac{k+1}{N}} t^n \,\mathrm{d}t,$$

soit

$$\frac{1}{N}\sum_{k=1}^{N-1}\left(\frac{k}{N}\right)^n \leqslant \int_0^1 t^n \,\mathrm{d}t = \frac{1}{n+1}$$

d'après la relation de Chasles, ce qui démontre l'inégalité recherchée et clôt la preuve.

5. Comme M_n est à valeurs dans $[\![1, N]\!]$, on peut réécrire le résultat de la question précédente sous la forme d'un encadrement :

$$\forall n \geqslant 1, \quad N - \frac{N}{n+1} \leqslant \mathbb{E}(M_n) \leqslant N.$$

Comme $N - \frac{N}{n+1} \xrightarrow[n\to+\infty]{} N$, on a bien $\mathbb{E}(M_n) \xrightarrow[n\to+\infty]{} N$ d'après le théorème des gendarmes.

Lorsque n est grand, il devient très probable d'avoir tiré au moins une fois la boule numérotée N au cours des n premiers tirages, donc il est très probable que M_n prenne la valeur N. Cette « convergence [13] de M_n vers N » se traduit ici par la convergence de la valeur moyenne de M_n (c'est-à-dire $\mathbb{E}(M_n)$) vers N.

Correction de l'exercice 46.

1. Si les B_k sont mutuellement indépendants, alors S est la somme de 100 variables de Bernoulli indépendantes de paramètre $\frac{1}{100}$, donc S suit la loi binomiale $\mathcal{B}(100, \frac{1}{100})$.
2. Si les B_k sont deux à deux incompatibles, alors presque sûrement l'un des B_k (et un seul) est réalisé puisque l'on a, par incompatibilité :

$$\mathbb{P}\left(\bigsqcup_{k=1}^{100} B_k\right) = \sum_{k=1}^{100}\mathbb{P}(B_k) = 100 \cdot \frac{1}{100} = 1.$$

Ainsi, avec probabilité 1 la variable aléatoire S prend la valeur 1 : elle suit donc la loi de Dirac de paramètre 1.

3. Comme S est la somme de 100 variables de Bernoulli de paramètre $\frac{1}{100}$, qui admettent toutes $\frac{1}{100}$ pour espérance, alors S admet une espérance et on a, par linéarité de l'espérance :

$$\mathbb{E}(S) = \sum_{k=1}^{100}\mathbb{E}(X_k) = 100 \cdot \frac{1}{100} = 1.$$

13. On verra dans le chapitre sur la convergence des suites de variables aléatoires qu'il est possible de qualifier de diverses façons le fait que la suite $(M_n)_{n\in\mathbb{N}^*}$ converge vers N. Dans le cas qui nous occupe, on observe une convergence presque sûre, en probabilité, en loi et en espérance.

4. Pour tous $i, j \in [\![1, 100]\!]$, les VARD X_i et X_j suivent des lois de Bernoulli donc elles admettent des moments d'ordre 2, et d'après l'inégalité de Cauchy-Schwarz on a

$$|\mathrm{Cov}(X_i, X_j)| \leqslant \sqrt{\mathbb{V}(X_i)}\sqrt{\mathbb{V}(X_j)} = \sqrt{\frac{1}{100} \cdot \frac{99}{100}}\sqrt{\frac{1}{100} \cdot \frac{99}{100}} = \frac{99}{10\,000},$$

d'où le résultat attendu.

5. Comme les X_k admettent toutes un moment d'ordre 2, c'est aussi le cas de leur somme S et on a

$$\begin{aligned}
\mathbb{E}(S^2) &= \mathbb{E}\left(\sum_{i=1}^{100}\sum_{j=1}^{100} X_i X_j\right) \\
&= \sum_{i=1}^{100}\sum_{j=1}^{100} \mathbb{E}(X_i X_j) \quad \text{par linéarité de l'espérance} \\
&= \sum_{i=1}^{100}\sum_{j=1}^{100} (\mathrm{Cov}(X_i, X_j) + \mathbb{E}(X_i)\,\mathbb{E}(X_j)) \quad \text{d'après la formule de Koenig} \\
&\leqslant \sum_{i=1}^{100}\sum_{j=1}^{100} \left(\frac{99}{10\,000} + \frac{1}{10\,000}\right) \quad \text{d'après la question précédente} \\
&= 10\,000 \cdot \frac{100}{10\,000} = 100,
\end{aligned}$$

ce qu'il fallait démontrer.

L'égalité est par exemple réalisée lorsque les X_i sont toutes égales à X_1, puisque l'on a dans ce cas $S = 100X_1$, donc $S^2 = 10\,000X_1^2 = 10\,000X_1$, ce qui implique que $\mathbb{E}(S^2) = 10\,000\,\mathbb{E}(X_1) = 100$.

Allons plus loin : on va montrer que l'égalité n'est réalisée que dans le cas où les X_i sont presque sûrement égales les unes aux autres.

Supposons que l'égalité $\mathbb{E}(S^2) = 100$ soit réalisée. Dans ce cas, toutes les inégalités utilisées pour établir ci-dessus que $\mathbb{E}(S^2) \leqslant 100$ sont saturées (c'est-à-dire sont des égalités). Soient $i, j \in [\![1, 100]\!]$; on a alors $\mathrm{Cov}(X_i, X_j) = \frac{99}{10\,000}$, donc l'inégalité de Cauchy-Schwarz $\mathrm{Cov}(X_i, X_j) \leqslant \sqrt{\mathbb{V}(X_i)}\sqrt{\mathbb{V}(X_j)}$ est une égalité (avec une covariance strictement positive). Ainsi, il existe donc $a \in \mathbb{R}_+^*$ et $b \in \mathbb{R}$ tels que $X_j = aX_i + b$ avec probabilité 1. Or en considérant la variance des deux termes, on obtient $\frac{99}{10\,000} = a^2 \cdot \frac{99}{10\,000}$, d'où $a^2 = 1$ et donc $a = 1$, et en considérant l'espérance de X_j et $X_i + b$, on obtient que $\frac{1}{100} = \frac{1}{100} + b$, soit $b = 0$. Ainsi, on a $X_i = X_j$ avec probabilité 1. Comme i et j sont quelconques, on a démontré que les X_i sont presque sûrement égales les unes aux autres, ce qui clôt la preuve.

Correction de l'exercice 47.

1. On a clairement $L_1(\Omega) = \mathbb{N}^*$ puisque la longueur de la première série est un entier pouvant prendre n'importe quelle valeur strictement positive. Remarquons à présent que si le premier lancer amène Pile, alors L_1 est la longueur

de la première série de Pile ; s'il amène Face, alors L_1 est la longueur de la première série de Face. Introduisons l'événement A : « le premier lancer amène Pile ». Pour tout $n \geqslant 1$, en utilisant la formule des probabilités totales sur le système complet d'événements $(A, \overline{A})$, on obtient

$$\begin{aligned}\mathbb{P}(L_1 = n) &= \mathbb{P}((L_1 = n) \cap A) + \mathbb{P}(\overline{A})\,\mathbb{P}((L_1 = n) \cap \overline{A}) \\ &= p^n q + q^n p,\end{aligned}$$

où le calcul de $\mathbb{P}((L_1 = n) \cap A)$, par exemple, résulte du fait que l'événement $(L_1 = n) \cap A$ est réalisé si et seulement les lancers 1 à n amènent Pile et le $(n+1)$-ième lancer amène Face, et du fait que tous les lancers sont indépendants.

On va à présent calculer l'espérance et la variance de L_1 en faisant apparaître des séries géométriques dérivées et dérivées deux fois, comme on l'a fait dans le cours pour des lois géométriques. Comme $p, q \in \,]0,1[$, les séries géométriques dérivées de paramètres p et q convergent, donc L_1 admet une espérance donnée par

$$\begin{aligned}\mathbb{E}(L_1) &= \sum_{n=1}^{+\infty} pq(np^{n-1} + nq^{n-1}) \\ &= pq\left(\sum_{n=1}^{+\infty} np^{n-1} + \sum_{n=1}^{+\infty} nq^{n-1}\right) \\ &= pq\left(\frac{1}{q^2} + \frac{1}{p^2}\right) = \frac{p}{q} + \frac{q}{p} \\ &= \frac{p^2+q^2}{pq}.\end{aligned}$$

À présent, la série géométrique dérivée deux fois de paramètre p converge puisque $p \in \,]0,1[$, et on a, par linéarité de la somme des séries convergentes :

$$\begin{aligned}\sum_{n=1}^{+\infty} n^2\, p^{n-1} &= \sum_{n=1}^{+\infty} (n(n-1) + n)p^{n-1} \\ &= p\sum_{n=2}^{+\infty} n(n-1)p^{n-2} + \sum_{n=1}^{+\infty} np^{n-1} \\ &= p \cdot \frac{2}{q^3} + \frac{1}{q^2} \\ &= \frac{2p+q}{q^3},\end{aligned}$$

et de même

$$\sum_{n=1}^{+\infty} n^2 q^{n-1} = \frac{2q+p}{p^3}$$

puisque $q \in \,]0,1[$. Ainsi, L_1^2 admet une espérance donnée par

$$\mathbb{E}(L_1^2) = pq\sum_{n=1}^{+\infty}(n^2p^{n-1} + n^2q^{n-1})$$

$$
\begin{aligned}
&= pq\left(\frac{2p+q}{q^3}+\frac{2q+p}{p^3}\right)\\
&= \frac{2p^2+pq}{q^2}+\frac{2q^2+pq}{p^2}\\
&= \frac{2p^4+p^3q+2q^4+pq^3}{p^2q^2}
\end{aligned}
$$

D'après la formule de Koenig, L_1 admet donc une variance égale à

$$
\begin{aligned}
\mathbb{V}(L_1) &= \mathbb{E}(L_1^2)-\mathbb{E}(L_1)^2\\
&= \frac{2p^4+p^3q+2q^4+pq^3}{p^2q^2}-\left(\frac{p^2+q^2}{pq}\right)^2\\
&= \frac{2p^4+p^3q+2q^4+pq^3-(p^2+q^2)^2}{p^2q^2}\\
&= \frac{p^4+p^3q+q^4+pq^3-2p^2q^2}{p^2q^2}\\
&= \frac{p^3(p+q)+q^3(q+p)-2p^2q^2}{p^2q^2}\\
&= \frac{p^3+q^3-2p^2q^2}{p^2q^2} \quad \text{car } p+q=1.
\end{aligned}
$$

2. On a $L_2(\Omega)=\mathbb{N}^*$ comme précédemment, et pour tous $(i,j)\in\mathbb{N}^*$ on peut écrire, grâce à la formule des probabilités totales appliquée au système complet d'événements $(A,\overline{A})$:

$$
\begin{aligned}
\mathbb{P}((L_1=i)\cap(L_2=j)) &= \mathbb{P}((L_1=i)\cap(L_2=j)\cap A)+\mathbb{P}((L_1=i)\cap(L_2=j)\cap\overline{A})\\
&= p^iq^jp+q^ip^jq\\
&= p^{i+1}q^j+q^{i+1}p^j
\end{aligned}
$$

puisque l'événement $(L_1=i)\cap(L_2=j)\cap A$, par exemple, est réalisé si et seulement si les premiers lancers, qui sont indépendants, amènent i fois Pile puis j fois Face puis encore une fois Pile.

On en déduit la loi marginale de L_2 par sommation : pour tout $j\in\mathbb{N}^*$,

$$
\begin{aligned}
\mathbb{P}(L_2=j) &= \sum_{i=1}^{+\infty}\mathbb{P}((L_1=i)\cap(L_2=j))\\
&= \sum_{i=1}^{+\infty}\left(p^{i+1}q^j+q^{i+1}p^j\right)=pq^j\sum_{i=1}^{+\infty}p^i+qp^j\sum_{i=1}^{+\infty}q^i\\
&= pq^j\cdot\frac{p}{q}+qp^j\cdot\frac{q}{p}=p^2q^{j-1}+q^2p^{j-1}.
\end{aligned}
$$

3. On peut déterminer l'espérance de L_2 en faisant apparaître des séries géométriques dérivées : on trouve

$$
\mathbb{E}(L_2)=p^2\sum_{j=1}^{+\infty}jq^{j-1}+q^2\sum_{j=1}^{+\infty}jp^{j-1}=p^2\cdot\frac{1}{p^2}+q^2\cdot\frac{1}{q^2}=2.
$$

Il peut sembler surprenant que cette espérance ne dépende pas de p et q. Ce fait s'explique de manière informelle[14] en remarquant que la première série est une série de Pile avec probabilité p et de Face avec probabilité q, donc la deuxième est une série de Face avec probabilité p et de Pile avec probabilité q ; or la longueur moyenne d'une série de Pile est $\frac{1}{q}$ (l'espérance d'une variable aléatoire de loi $\mathcal{G}(q)$), et la longueur moyenne d'une série de Face est $\frac{1}{p}$ (l'espérance d'une variable aléatoire de loi $\mathcal{G}(p)$). Ainsi, la probabilité que la deuxième série soit constituée d'un résultat donné est exactement compensée par la longueur moyenne de la série associée à ce résultat : la longueur moyenne de L_2 est ainsi égale à $p \cdot \frac{1}{p} + q \cdot \frac{1}{q} = 2$.

Remarquons qu'un raisonnement similaire aurait permis d'écrire directement que $\mathbb{E}(L_1) = p \cdot \frac{1}{q} + q \cdot \frac{1}{p}$.

4. En vertu de la discussion que nous venons de mener, nous avons de bonnes raisons de suspecter que L_1 et L_2 ne sont pas indépendantes lorsque $p \neq q$: par exemple, si p est fort, obtenir une valeur importante pour L_1 tend à suggérer que la première série est une série de Pile, et donc que L_2, qui est la longueur d'une série de Face, prendra une valeur faible.

   Démontrons formellement cette absence d'indépendance. On suppose que $p \neq q$, c'est-à-dire que $p \neq \frac{1}{2}$. On a alors

$$\mathbb{P}((L_1 = 1) \cap (L_2 = 1)) = p^2 q + q^2 p = pq(p+q) = pq,$$

   tandis que

$$\mathbb{P}(L_1 = 1)\,\mathbb{P}(L_2 = 1) = 2pq \cdot (p^2 + q^2),$$

   d'où

$$\begin{aligned}\mathbb{P}(L_1 = 1)\,\mathbb{P}(L_2 = 1) - \mathbb{P}((L_1 = 1) \cap (L_2 = 1)) &= pq(2(p^2+q^2) - 1)\\ &= pq(2(p^2+q^2) - (p+q)^2)\\ &= pq(p^2 + q^2 - 2pq)\\ &= pq(p-q)^2 \neq 0,\end{aligned}$$

   donc

$$\mathbb{P}((L_1 = 1) \cap (L_2 = 1)) \neq \mathbb{P}(L_1 = 1)\,\mathbb{P}(L_2 = 1),$$

   ce qui montre que L_1 et L_2 ne sont pas indépendantes.

---

14. Le raisonnement présenté ici se formalise en utilisant la notion d'*espérance conditionnelle*, qui n'est pas au programme de B/L. En définissant l'espérance de L_2 conditionnellement à un événement E comme $\mathbb{E}(L_2 \,|\, A) := \sum_{k=1}^{+\infty} j\,\mathbb{P}(L_2 = j \,|\, E)$, on peut écrire la formule suivante, dite *formule de l'espérance totale*, dont on appréciera l'analogie avec la formule des probabilités totales :

$$\mathbb{E}(L_2) = \mathbb{P}(A)\mathbb{E}(L_2 \,|\, A) + \mathbb{P}(\overline{A})\mathbb{E}(L_2 \,|\, \overline{A}).$$

Cette relation permet de conclure.

Si $p = \frac{1}{2}$, en revanche, L_1 et L_2 sont indépendantes. Cela peut se voir par un calcul direct : les questions précédentes donnent en effet que dans ce cas, L_1 et L_2 suivent des lois géométriques de paramètre $\frac{1}{2}$, et que pour tous $(i, j) \in \mathbb{N}^*$ on a

$$\begin{aligned}\mathbb{P}((L_1 = i) \cap (L_2 = j)) &= \left(\frac{1}{2}\right)^{i+1} \left(\frac{1}{2}\right)^{j} + \left(\frac{1}{2}\right)^{i+1} \left(\frac{1}{2}\right)^{j} \\ &= \left(\frac{1}{2}\right)^{i+j} = \left(\frac{1}{2}\right)^{i} \left(\frac{1}{2}\right)^{j} \\ &= \mathbb{P}(L_1 = i)\,\mathbb{P}(L_2 = j).\end{aligned}$$

Cela peut aussi se comprendre en remarquant que puisque $p = \frac{1}{2}$, les Pile et les Face jouent un rôle symétrique, donc le résultat du premier lancer n'influe pas sur la loi de la longueur des séries, et tout se passe comme si l'expérience aléatoire était réinitialisée sans mémoire lors du début de la deuxième série.

Correction de l'exercice 48.

1. (a) Par indépendance des floraisons, chaque ligne a une probabilité p^n de fleurir intégralement, indépendamment des autres lignes ; ainsi, la variable aléatoire L, qui compte le nombre de succès obtenus lors de la répétition de n expériences de Bernoulli indépendantes de probabilité de succès p^n, suit une loi binomiale $\mathcal{B}(n, p^n)$. Il en va de même pour C par symétrie du problème.

   (b) Il est intuitivement clair que les variables L et C ne sont pas indépendantes puisque si l'événement $(L = n)$ est réalisé, alors tous les bulbes ont fleuri donc l'événement $(C = n)$ est nécessairement réalisé. Plus formellement, on a

   $$\mathbb{P}((L = n) \cap (C = n)) = \mathbb{P}(L = n) = (p^n)^n = p^{n^2}$$

   tandis que

   $$\mathbb{P}(L = n)\,\mathbb{P}(C = n) = (p^{n^2})(p^{n^2}) = p^{2n^2},$$

   or $p \neq 1$ et $n \geqslant 1$ donc $p^{n^2} \neq p^{2n^2}$, d'où

   $$\mathbb{P}((L = n) \cap (C = n)) \neq \mathbb{P}(L = n)\,\mathbb{P}(C = n),$$

   ce qui montre bien que L et C ne sont pas indépendantes.

2. L_i et C_j sont à valeurs dans $\{0, 1\}$, donc elles suivent une loi de Bernoulli. La ligne i et la colonne j étant chacune constituées de n bulbes à floraison indépendante, on a $\mathbb{P}(L_i = 1) = \mathbb{P}(C_j = 1) = p^n$, si bien que L_i et C_j suivent la loi $\mathcal{B}(p^n)$.

   Le produit L_iC_j vaut 1 si et seulement si tous les bulbes situés sur la ligne i et tous les bulbes situés sur la colonne j ont fleuri, et 0 sinon ; or ces bulbes sont au nombre de $2n - 1$ (puisqu'un unique bulbe est commun à la ligne et à la colonne), et leurs floraisons sont indépendantes, donc $L_iC_j \sim \mathcal{B}(p^{2n-1})$. Ainsi, on a d'après la formule de Koenig :

   $$\operatorname{Cov}(L_i, C_j) = \mathbb{E}(L_iC_j) - \mathbb{E}(L_i)\mathbb{E}(C_j) = p^{2n-1} - p^np^n = (1 - p)p^{2n-1}.$$

3. D'après la question précédente et la bilinéarité de la covariance, on peut écrire

$$\begin{aligned}\mathrm{Cov}(L,C) &= \mathrm{Cov}\left(\sum_{i=1}^{n} L_i, \sum_{j=1}^{n} C_j\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(L_i, C_j)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} (1-p)p^{2n-1}\\ &= n^2(1-p)p^{2n-1}.\end{aligned}$$

On remarque que cette covariance est positive, conformément à l'intuition puisque L et C tendent toutes deux à prendre des valeurs élevées lorsque le parterre est très fleuri et des valeurs faibles lorsqu'il ne l'est pas.

Correction de l'exercice 49.

1. (a) Si le premier résultat est négatif, alors le mélange ne présente pas la propriété, donc aucune des n personnes ne la présente ; dans ce cas, on dispose de toute l'information requise et il n'est pas nécessaire de réaliser une autre analyse.
   (b) Dans le cadre du protocole **A**, on réalise 1 ou $n+1$ tests, donc $A_n(\Omega) = \{1, n+1\}$.
   (c) La probabilité de n'avoir à réaliser qu'un test est la probabilité qu'aucune des personnes ne présente la propriété, c'est-à-dire q^n par indépendance. Ainsi, on a $\mathbb{P}(A_n = 1) = q^n$, et donc $\mathbb{P}(A_n = n+1) = 1 - q^n$.
   (d) Comme A_n prend un nombre fini de valeurs, elle admet une espérance donnée par

$$\begin{aligned}\mathbb{E}(A_n) &= 1 \cdot \mathbb{P}(A_n = 1) + (n+1)\,\mathbb{P}(A_n = n+1)\\ &= q^n + (n+1)(1-q^n) = n + 1 - nq^n.\end{aligned}$$

2. Notons $A = (A_n = n+1)$ l'événement « le résultat de l'analyse du mélange est positif », B l'événement « les k premières personnes sont toutes négatives » et C l'événement « les $n-k$ dernières personnes sont toutes négatives » (on considère que les n personnes sont ordonnées et possèdent ou non la propriété donc sont intrinsèquement « positives » ou « négatives », qu'elles soient individuellement testées ou pas). On a alors

$$\mathbb{P}(B \,|\, A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B \cap \overline{C})}{\mathbb{P}(A)}$$

puisque l'événement $B \cap A$ correspond au fait que les k premières personnes sont toutes négatives et qu'au moins une personne parmi les $n-k$ dernières est positive, c'est-à-dire à l'événement $B \cap \overline{C}$. Or B et $\overline{C}$ sont indépendants puisqu'ils portent sur des groupes disjoints de personnes et puisque les propriétés sont indépendantes entre les individus, donc

$$\mathbb{P}(B \,|\, A) = \frac{\mathbb{P}(B)\,\mathbb{P}(\overline{C})}{\mathbb{P}(A)}.$$

Mais on a vu que $\mathbb{P}(A) = q^n$, et par le même raisonnement $\mathbb{P}(B) = q^k$ et $\mathbb{P}(C) = q^{n-k}$, donc

$$\mathbb{P}(B \mid A) = \frac{q^k(1-q^{n-k})}{q^n} = \frac{q^k - q^n}{q^n}.$$

3. (a) Comme on l'a vu dans la question 1(d), le nombre moyen de tests effectués dans le cadre du protocole **A** est $n+1-nq^n$. Le nombre de tests effectués dans le cadre du protocole **B** est toujours n. Ainsi, on fait en moyenne moins de tests avec le protocole **A** qu'avec le protocole **B** si et seulement si on a $n+1-nq^n \leqslant n$, c'est-à-dire si et seulement si $nq^n \geqslant 1$.

   (b) On a
$$x^x = e^{x\ln(x)} \xrightarrow[x\to 0^+]{} e^0 = 1$$
   puisque $x\ln(x) \xrightarrow[x\to 0^+]{} 0$ par croissance comparée et puisque exp est continue en 0.

   (c) D'après la question 3(a), il s'agit de montrer que lorsque n est assez grand alors $nq^n < 1$. Or pour tout $n \geqslant 1$ on a
$$nq^n < 1 \iff q^n < \frac{1}{n} \iff q < \left(\frac{1}{n}\right)^{\frac{1}{n}},$$
   mais $\left(\frac{1}{n}\right)^{\frac{1}{n}} \xrightarrow[n\to+\infty]{} 1$ d'après la question précédente [15] et $q < 1$, donc on a $q < \left(\frac{1}{n}\right)^{\frac{1}{n}}$ à partir d'un certain rang. Ainsi, pour n assez grand on a bien $nq^n < 1$, c'est-à-dire que le protocole **B** est préférable au protocole **A** selon le critère du nombre moyen de tests à effectuer

Correction de l'exercice 50.

1. La fonction h est en fait une suite : pour tout $n \in \mathbb{N}$, le réel $h(n)$ est le terme traditionnellement noté h_n. En tant que suite croissante et majorée, elle converge donc d'après le théorème de la limite monotone [16], c'est-à-dire que h admet une limite finie en $+\infty$.
2. Comme h est bornée, les VARD $h(X)$ et $h(Y)$ le sont elles aussi ; elles admettent donc des espérances.
3. On a, en notant $L := \lim\limits_{+\infty} h$:
$$\sum_{n=1}^{N} (h(n) - h(n-1)) = h(N) - h(0) \xrightarrow[N\to+\infty]{} L - h(0)$$
   en repérant une somme télescopique, ce qui signifie que la série de terme général $h(n) - h(n-1)$ converge.

15. On réalise ici quelques contorsions pour utiliser la question précédente, mais il aurait été possible de directement remarquer que puisque $q \in]0,1[$ alors $nq^n \xrightarrow[n\to+\infty]{} 0$ par croissance comparée, et donc $nq^n < 1$ à partir d'un certain rang. Il s'agit bien sûr du même argument, qui n'est plus succinct que parce qu'il utilise un résultat du cours.

16. Le résultat aurait de toute façon été vrai pour une fonction définie sur une partie quelconque de $\mathbb{R}$.

4. Comme h est croissante, on peut écrire que pour tout $n \geqslant 1$ on a
$$h(n) - h(n-1) \geqslant 0$$
d'où, en multipliant par $\mathbb{P}(X \geqslant n) \in [0,1]$:
$$0 \leqslant (h(n) - h(n-1))\, \mathbb{P}(X \geqslant n) \leqslant h(n) - h(n-1).$$
Or la série de terme général $h(n) - h(n-1)$ converge, donc il en va de même pour la série de terme général $(h(n) - h(n-1))\mathbb{P}(X \geqslant n)$ par comparaison des séries à termes positifs.
5. La suite d'événements $((X \geqslant N+1))_{N \geqslant 0}$ est décroissante, donc on a d'après le théorème de la limite monotone :
$$\lim_{N \to +\infty} \mathbb{P}(X \geqslant N+1) = \mathbb{P}\left(\bigcap_{N=0}^{+\infty} (X \geqslant N+1)\right) = \mathbb{P}(\varnothing) = 0$$
puisque X ne peut prendre une valeur supérieure à tous les entiers naturels.
6. Soit $N \geqslant 1$. On a [17]
$$\begin{aligned}
&\sum_{n=1}^{N+1} (h(n) - h(n-1))\mathbb{P}(X \geqslant n) \\
&\qquad = \sum_{n=1}^{N+1} h(n)\mathbb{P}(X \geqslant n) - \sum_{n=1}^{N+1} h(n-1)\mathbb{P}(X \geqslant n) \\
&\qquad = \sum_{n=1}^{N+1} h(n)\mathbb{P}(X \geqslant n) - \sum_{j=0}^{N} h(j)\mathbb{P}(X \geqslant j+1) \quad \text{en posant } j = n-1 \\
&\qquad = h(N+1)\mathbb{P}(X \geqslant N+1) + \sum_{n=1}^{N} h(n)(\mathbb{P}(X \geqslant n) - \mathbb{P}(X \geqslant n+1)) - h(0)\mathbb{P}(X \geqslant 1) \\
&\qquad = h(N+1)\mathbb{P}(X \geqslant N+1) + \sum_{n=1}^{N} h(n)\, \mathbb{P}(X = n) - h(0)\mathbb{P}(X \geqslant 1),
\end{aligned}$$
d'où
$$\begin{aligned}
&\sum_{n=1}^{N} h(n)\mathbb{P}(X = n) \\
&\qquad = \sum_{n=1}^{N+1} (h(n) - h(n-1))\mathbb{P}(X \geqslant n) - h(N+1)\mathbb{P}(X \geqslant N+1) + h(0)\mathbb{P}(X \geqslant 1)
\end{aligned}$$
et donc, en ajoutant $h(0)\mathbb{P}(X = 0)$ de part et d'autre :
$$\begin{aligned}
&\sum_{n=0}^{N} h(n)\mathbb{P}(X = n) \\
&\qquad = \sum_{n=1}^{N+1} (h(n) - h(n-1))\mathbb{P}(X \geqslant n) - h(N+1)\mathbb{P}(X \geqslant N+1) + h(0).
\end{aligned}$$

17. La réécriture obtenue ici est une *transformation d'Abel* (voir l'exercice 32 du chapitre « Sommes et produits » dans le tome de première année). Il s'agit de l'équivalent discret d'une intégration par parties, où la *dérivée discrète* $h(n) - h(n-1)$ joue le rôle de la « dérivée » de $h(n)$ et où $\mathbb{P}(X \geqslant n) = \sum_{k=n}^{+\infty} \mathbb{P}(X = k)$ joue le rôle d'une primitive de $-\mathbb{P}(X = n)$.

Or h est bornée donc $h(N+1)\mathbb{P}(X \geqslant N+1) \xrightarrow[N\to+\infty]{} 0$ d'après la question précédente, et la série de terme général $(h(n)-h(n-1)\mathbb{P}(X \geqslant n)$ converge d'après la question 4 ; ainsi, on a

$$\sum_{n=0}^{N} h(n)\mathbb{P}(X=n) \xrightarrow[N\to+\infty]{} \sum_{n=1}^{+\infty}(h(n)-h(n-1))\mathbb{P}(X \geqslant n)+h(0).$$

On en déduit d'après le théorème de transfert que

$$\mathbb{E}(h(X)) = \sum_{n=0}^{+\infty} h(n)\mathbb{P}(X=n) = \sum_{n=1}^{+\infty}(h(n)-h(n-1))\mathbb{P}(X \geqslant n)+h(0),$$

ce qu'il fallait démontrer.

7. Supposons que $X \leqslant_{st} Y$. Pour tout $n \in \mathbb{N}^*$, on a alors $\mathbb{P}(X \geqslant n) \leqslant \mathbb{P}(Y \geqslant n)$, donc, comme h est croissante,

$$(h(n)-h(n-1))\mathbb{P}(X \geqslant n) \leqslant (h(n)-h(n-1))\mathbb{P}(Y \geqslant n).$$

Ainsi, la question précédente donne

$$\begin{aligned}\mathbb{E}(h(X)) &= h(0)+\sum_{n=1}^{+\infty}(h(n)-h(n-1))\mathbb{P}(X \geqslant n)\\ &\leqslant h(0)+\sum_{n=1}^{+\infty}(h(n)-h(n-1))\mathbb{P}(Y \geqslant n) = \mathbb{E}(h(Y)),\end{aligned}$$

ce que nous cherchions à obtenir.

8. Pour tout $i \in \mathbb{N}$, on a $h_n(i+1) = h_n(i)$ si $i \neq n-1$ et $h_n(i+1) = h_n(i)+1$ si $i = n-1$, donc dans tous les cas $h_n(i) \leqslant h_n(i+1)$; ainsi, h_n est croissante. Par ailleurs, comme h_n ne prend que les valeurs 0 et 1, elle est bornée.

9. Soit $n \in \mathbb{N}^*$. Comme h_n est croissante et bornée, on a par hypothèse $\mathbb{E}(h_n(X)) \leqslant \mathbb{E}(h_n(Y))$. Or $h_n(X)$ prend uniquement les valeurs 0 et 1, donc

$$\mathbb{E}(h_n(X)) = \mathbb{P}(h_n(X)=1) = \mathbb{P}(X \geqslant n),$$

et on a de même

$$\mathbb{E}(h_n(Y)) = \mathbb{P}(Y \geqslant n).$$

Ainsi, on a bien $\mathbb{P}(X \geqslant n) \leqslant \mathbb{P}(Y \geqslant n)$.

10. La relation $\mathbb{P}(X \geqslant t) \leqslant \mathbb{P}(Y \geqslant t)$ est vraie pour tout entier positif t d'après la question précédente. Si t est un réel strictement négatif, elle est vraie aussi puisque les deux probabilités $\mathbb{P}(X \geqslant t)$ et $\mathbb{P}(Y \geqslant t)$ sont nulles. Enfin, si $t \in \mathbb{R}_+$, comme X et Y sont à valeurs entières on a $\mathbb{P}(X \geqslant t) = \mathbb{P}(X \geqslant \lfloor t \rfloor)$ et $\mathbb{P}(Y \geqslant t) = \mathbb{P}(Y \geqslant \lfloor t \rfloor)$, d'où, en utilisant l'inégalité vérifiée pour les entiers positifs :

$$\mathbb{P}(X \geqslant t) = \mathbb{P}(X \geqslant \lfloor t \rfloor) \leqslant \mathbb{P}(Y \geqslant \lfloor t \rfloor) \leqslant \mathbb{P}(Y \geqslant t).$$

Ainsi, on a $\mathbb{P}(X \geqslant t) \leqslant \mathbb{P}(Y \geqslant t)$ pour tout $t \in \mathbb{R}$, ce qui montre bien que $X \leqslant_{st} Y$.

11. On vient de démontrer que $X \leqslant_{st} Y$ si et seulement si pour toute fonction $h : \mathbb{N} \to \mathbb{R}_+$ croissante et bornée on a $\mathbb{E}(h(X)) \leqslant \mathbb{E}(h(Y))$.
12. On a

$$\begin{aligned}
\mathbb{P}(X \leqslant Y) &= \mathbb{P}\left(\bigsqcup_{k=0}^{+\infty} \Big((X = k) \cap (Y \geqslant k)\Big)\right) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}((X = k) \cap (Y \geqslant k)) \quad \text{par incompatibilité} \\
&= \sum_{k=0}^{+\infty} \mathbb{P}(X = k)\, \mathbb{P}(Y \geqslant k) \quad \text{par indépendance,}
\end{aligned}$$

d'où le résultat attendu.
13. Comme $X \leqslant_{st} Y$ et comme X et Z ont même loi, on a

$$\mathbb{P}(Y \geqslant k) \geqslant \mathbb{P}(X \geqslant k) = \mathbb{P}(Z \geqslant k)$$

pour tout $k \in \mathbb{N}$. Ainsi, le résultat de la question précédente donne

$$\mathbb{P}(X \leqslant Y) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)\, \mathbb{P}(Y \geqslant k) \geqslant \sum_{k=0}^{+\infty} \mathbb{P}(X = k)\, \mathbb{P}(Z \geqslant k)$$

puisque les $\mathbb{P}(X = k)$ sont positifs.
14. L'argument est le même que dans la question 12 en remplaçant Y par Z.
15. Soit $n \in \mathbb{N}$ tel que $\mathbb{P}(X = n) > 0$ (un tel n existe nécessairement puisque la somme des $\mathbb{P}(X = n)$ vaut 1). Comme X et Z sont indépendantes et de même loi, on a

$$\begin{aligned}
\mathbb{P}(X = Z) &\geqslant \mathbb{P}((X = n) \cap (Z = n)) \\
&= \mathbb{P}(X = n)\, \mathbb{P}(Z = n) = \mathbb{P}(X = n)^2 > 0.
\end{aligned}$$

Démontrons à présent la deuxième relation. Comme X et Z suivent la même loi, on a, en reprenant le résultat de la question 14 :

$$\begin{aligned}
\mathbb{P}(Z \geqslant X) &= \sum_{k=0}^{+\infty} \mathbb{P}(X = k)\, \mathbb{P}(Z \geqslant k) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}(Z = k)\, \mathbb{P}(X \geqslant k) = \mathbb{P}(X \geqslant Z),
\end{aligned}$$

d'où

$$\mathbb{P}(Z \geqslant X) = \mathbb{P}(X > Z) + \mathbb{P}(X = Z),$$

soit

$$\mathbb{P}(Z \geqslant X) = 1 - \mathbb{P}(Z \geqslant X) + \mathbb{P}(X = Z),$$

donc

$$2\, \mathbb{P}(Z \geqslant X) = 1 + \mathbb{P}(X = Z),$$

et enfin

$$\mathbb{P}(Z \geqslant X) = \frac{1}{2} + \frac{\mathbb{P}(X = Z)}{2}.$$

Or $\mathbb{P}(X = Z) > 0$, donc $\mathbb{P}(Z \geqslant X) > \frac{1}{2}$.

16. D'après les questions 13, 14 et 15, on a

$$\mathbb{P}(X \leqslant Y) \geqslant \sum_{k=0}^{+\infty} \mathbb{P}(X = k)\,\mathbb{P}(Z \geqslant k) = \mathbb{P}(Z \geqslant X) > \frac{1}{2},$$

d'où le résultat attendu.

Correction de l'exercice 51. Commentons tout d'abord le formalisme utilisé. Le problème étudié est celui d'une marche aléatoire symétrique sur $\mathbb{Z}$, c'est-à-dire du mouvement aléatoire d'une particule commençant sur le point 0 et sautant de façon équiprobable à chaque étape sur l'entier inférieur ou sur l'entier supérieur, indépendamment du passé, jusqu'à avoir effectué N sauts. Un élément $\omega \in \Omega$ consiste en la liste des sauts opérés par la particule (à gauche ou à droite), et toutes ces listes sont équiprobables (de probabilité $\frac{1}{2^N}$). Pour une liste ω donnée et un $n \in [\![1, N]\!]$, l'entier $S_n(\omega)$ est la position de la particule après le n-ième saut, et l'entier $T_n(\omega)$ est le nombre de sauts « vers le haut [18] » faits par la particule lors des n premières étapes. Dans la suite, on s'autorisera ainsi à fonder certains raisonnements sur la description de la *trajectoire* aléatoire de la particule, qui démarre du point 0 et suit les directions données par le vecteur ω.

1. La variable aléatoire T_n compte le nombre de « pas vers le haut » dans une trajectoire de longueur n, c'est-à-dire le nombre de succès obtenus lors de la répétition de n expériences de Bernoulli indépendantes de probabilité de succès $\frac{1}{2}$: ainsi, on a $T_n \sim \mathcal{B}(n, \frac{1}{2})$.
2. Comme toute trajectoire de longueur n considérée part de 0 et effectue T_n « pas vers le haut » et $n - T_n$ « pas vers le bas », on a

$$S_n = (-1) \cdot (n - T_n) + 1 \cdot T_n = 2T_n - n.$$

Or T_n prend avec probabilité strictement positive toutes les valeurs de $[\![0, n]\!]$ (et uniquement celles-ci), donc l'ensemble des valeurs prises par S_n avec probabilité strictement positive est

$$S_n(\Omega) = \{2k - n \,:\, k \in [\![0, n]\!]\} = \{i \in [\![-n, n]\!] \,:\, i \text{ et } n \text{ ont même parité}\}.$$

3. Il est clair que si $x \notin S_n(\Omega)$, alors $p_{n,x} = \mathbb{P}(S_n = x) = 0$. À présent, si $x \in S_n(\Omega)$ et si $k \in [\![0, n]\!]$ est tel que l'on puisse écrire $x = 2k - n$, alors l'événement [19] $\{S_n = x\} = \{S_n = 2k - n\}$ est réalisé si et seulement si $\{T_n = k\}$ l'est, donc

$$p_{n,x} = \mathbb{P}(T_n = k) = \binom{n}{k}\frac{1}{2^n}.$$

18. Conformément à la représentation graphique donnée, on ordonne ici les entiers de $\mathbb{Z}$ verticalement plutôt qu'horizontalement ; un saut de $+1$ est donc « vers le haut » et non « vers la droite ».

19. La notation $\{S_n = x\}$ adoptée dans cet exercice diffère de celle utilisée dans le reste du chapitre (qui serait plutôt $(S_n = x)$), mais on adopte ici la convention de notation de l'énoncé qui est celle du sujet d'origine.

On a ainsi :

$$\forall x \in \mathbb{Z}, \quad p_{n,x} = \begin{cases} \dbinom{n}{k}\dfrac{1}{2^n} = \dbinom{n}{\frac{x+n}{2}}\dfrac{1}{2^n} & \text{si } x \in S_n(\Omega) \\ 0 & \text{sinon.} \end{cases}$$

4. Supposons que $n = 2k$ pour un $k \in \mathbb{N}^*$. Dans ce cas, on a $0 = 2k - n$ donc $0 \in S_n(\Omega)$, et d'après la question précédente

$$p_{2k,0} = \binom{n}{k}\frac{1}{2^n} = \binom{2k}{k}\frac{1}{2^{2k}} = q_{2k},$$

ce qu'il fallait démontrer.

5. En utilisant la formule de Stirling comme dans la question 7 de l'exercice 43 ci-dessus[20], on obtient

$$q_{2k} = \frac{(2k)!}{k!k!}\frac{1}{2^{2k}} \underset{k\to+\infty}{\sim} \frac{1}{\sqrt{\pi k}}.$$

6. L'événement $F_{n,x}$ s'écrit verbalement comme « la particule arrive en x en n étapes en commençant par un saut vers le haut, en étant repassée par 0 durant sa trajectoire ». L'événement $G_{n,x}$ s'écrit comme « la particule arrive en x en n étapes en commençant par un saut vers le bas ». Enfin, l'événement $H_{n,x}$ s'écrit comme « la particule arrive en x en n étapes en ne repassant pas par 0 durant sa trajectoire ».

   Comme $x > 0$, si le premier saut de la particule est vers le bas, alors elle ne peut arriver en x à l'étape n qu'en repassant par 0. Ainsi, si $H_{n,x}$ est réalisé alors nécessairement le premier saut de la particule est vers le haut. L'événement $F_{n,x} \cup H_{n,x}$ s'écrit donc « la particule effectue un premier saut vers le haut et arrive en x en n étapes », ce qui est exactement l'événement $E_{n,x} \cap \{S_1 = +1\}$.

7. Si la particule commence par effectuer un saut vers le bas, alors sa trajectoire à partir de l'étape suivante est une marche aléatoire symétrique de point de départ -1 ; en particulier, pour que la particule rejoigne le point x à l'étape n, elle doit effectuer en $n-1$ étapes une marche qui termine $x+1$ positions au-dessus de ce nouveau point de départ, ce qui se produit avec probabilité $\mathbb{P}(E_{n-1,x+1}) = p_{n-1,x+1}$. Plus formellement, on a $\mathbb{P}(E_{n,x} \mid S_1 = -1) = p_{n-1,x+1}$, d'où

$$\mathbb{P}(G_{n,x}) = \mathbb{P}(E_{n,x} \cap \{S_1 = -1\}) = \mathbb{P}(S_1 = -1)\,\mathbb{P}(E_{n,x} \,|\, S_1 = -1) = \frac{1}{2}p_{n-1,x+1},$$

   d'où le résultat attendu avec $n' = n - 1$ et $x' = x + 1$.

8. Il s'agit de trouver une bijection $\varphi : \Omega \to \Omega$ envoyant $F_{n,x}$ sur $G_{n,x}$. Pour cela, on utilise l'intuition donnée par le graphique fourni dans l'énoncé : un élément ω de $F_{n,x}$, qui correspond à une trajectoire qui commence par un saut

20. L'expression donnée pour $\mathbb{E}(V_n)$ dans cette question est égale à $(2n+1)q_{2n}$, d'où l'équivalent donné ici.

vers le haut et touche 0 avant la n-ième étape, peut être associé à un élément $\varphi(\omega) = \omega'$ qui correspond à la trajectoire *symétrisée jusqu'au premier retour en* 0, qui est par construction dans $G_{n,x}$.

Plus précisément [21], pour tout $\omega \in \Omega$, on définit le temps de retour en 0 de la trajectoire associée à ω comme

$$R(\omega) := \min\left(\{i \in [\![1, N]\!] \,:\, \omega_i = 0\} \cup \{N+1\}\right),$$

où l'ajout du singleton $N+1$ permet que l'ensemble dont on calcule le minimum soit non vide et revient à poser $R(\omega) = N+1$ si la trajectoire associée à ω ne revient pas en 0. On définit à présent $\varphi(\omega)$ comme le N-uplet $\omega' \in \Omega$ vérifiant

$$\forall i \in [\![1, N]\!], \quad \omega'_i := \begin{cases} -\omega_i & \text{si } i < R(\omega) \\ \omega_i & \text{si } i \geqslant R(\omega). \end{cases}$$

Ainsi, si la trajectoire associée à ω ne touche pas 0, alors l'intégralité de la trajectoire est symétrisée par φ, c'est-à-dire que $\varphi(\omega) = \omega' := -\omega$. L'application

$$\varphi : \begin{array}{ccc} \Omega & \longrightarrow & \Omega \\ \omega & \longmapsto & \varphi(\omega) = \omega' \end{array}$$

ainsi définie vérifie $(\varphi \circ \varphi)(\omega) = \omega$ pour tout $\omega \in \Omega$ (puisque le symétrique [22] du symétrique d'une trajectoire donnée est la trajectoire elle-même), c'est-à-dire qu'elle est *involutive*. Ainsi, elle est bijective et égale à sa propre bijection réciproque : en effet, tout $\omega \in \Omega$ admet $\varphi(\omega)$ pour unique antécédent par φ.

Enfin, on a $\varphi(F_{n,x}) = G_{n,x}$ puisque :

- La version symétrisée d'une trajectoire commençant par un saut vers le haut, atteignant 0 et arrivant en x à l'étape n est une trajectoire commençant par un saut vers le bas et arrivant en x à l'étape n (donc $\varphi(F_{n,x}) \subset G_{n,x}$).
- Réciproquement, toute trajectoire commençant par un saut vers le bas et arrivant en x à l'étape n passe nécessairement par 0 et est le symétrique de son propre symétrique, qui est une trajectoire commençant par un saut vers le haut, atteignant 0 et arrivant en x à l'étape n (donc $G_{n,x} \subset \varphi(F_{n,x})$).

Ainsi, comme φ est bijective et notamment injective, l'égalité $\varphi(F_{n,x}) = G_{n,x}$ assure que les ensembles finis $F_{n,x}$ et $G_{n,x}$ ont le même nombre d'éléments.

9. D'après la question précédente, on a $\mathbb{P}(F_{n,x}) = \mathbb{P}(G_{n,x})$ puisque $\mathbb{P}$ est l'équiprobabilité sur Ω, d'où $\mathbb{P}(F_{n,x}) = \frac{1}{2}p_{n-1,x+1}$ par la question 7. Or la question 6 donne, par incompatibilité de $F_{n,x}$ et $H_{n,x}$:

$$\mathbb{P}(F_{n,x}) + \mathbb{P}(H_{n,x}) = \mathbb{P}(E_{n,x} \cap \{S_1 = +1\})$$

21. Cette formalisation n'est pas strictement nécessaire dans une copie si l'action de φ est expliquée avec suffisamment de précision.

22. Parler de « symétrique » ne doit pas faire oublier que φ ne symétrise que le *début* de la trajectoire, c'est-à-dire ses premières valeurs jusqu'au retour de la particule en 0.

donc

$$\mathbb{P}(H_{n,x}) = \mathbb{P}(E_{n,x} \cap \{S_1 = +1\}) - \frac{1}{2} p_{n-1,x+1}.$$

Mais un raisonnement semblable [23] à celui développé dans la question 7 permet d'écrire que

$$\mathbb{P}(E_{n,x} \cap \{S_1 = +1\}) = \mathbb{P}(S_1 = +1)\, \mathbb{P}(E_{n,x} \,|\, S_1 = +1) = \frac{1}{2} p_{n-1,x-1}.$$

On a donc

$$\mathbb{P}(H_{n,x}) = \frac{p_{n-1,x-1} - p_{n-1,x+1}}{2}.$$

Calculons explicitement cette probabilité.

Si $x \notin S_n(\Omega)$, c'est-à-dire si x n'est pas accessible par la particule au terme de n étapes, alors $\mathbb{P}(H_{n,x}) = 0$, et on a de même $\mathbb{P}(H_{n,x} \,|\, E_{n,x}) = 0$.

Plaçons-nous désormais dans le cas où $x \in S_n(\Omega)$, et donnons-nous $k \in [\![0, n]\!]$ tel que $x = 2k - n$ (avec nécessairement $k > \frac{n}{2}$ puisque $x > 0$). On a alors, en utilisant l'expression trouvée dans la question 3 :

$$p_{n-1,x-1} = \binom{n-1}{\frac{n-1+x-1}{2}} \frac{1}{2^{n-1}} = \binom{n-1}{k-1} \frac{1}{2^{n-1}}$$

et

$$p_{n-1,x+1} = \binom{n-1}{\frac{n-1+x+1}{2}} \frac{1}{2^{n-1}} = \binom{n-1}{k} \frac{1}{2^{n-1}},$$

cette dernière probabilité pouvant être nulle si $k = n$. Ainsi, on a

$$\begin{aligned}
\mathbb{P}(H_{n,x}) &= \left(\binom{n-1}{k-1} - \binom{n-1}{k} \right) \frac{1}{2^n} \\
&= \left(\frac{(n-1)!}{(k-1)!(n-k)!} - \frac{(n-1)!}{k!(n-1-k)!} \right) \frac{1}{2^n} \\
&= \frac{k(n-1)! - (n-1)!(n-k)}{k!(n-k)!} \frac{1}{2^n} \\
&= \frac{(2k-n)(n-1)!}{k!(n-k)!} \frac{1}{2^n} \\
&= \frac{2k-n}{n} \cdot \frac{n!}{k!(n-k)!} \frac{1}{2^n} \\
&= \frac{2k-n}{n} \binom{n}{k} \frac{1}{2^n} \\
&= \frac{x}{n} p_{n,x}.
\end{aligned}$$

23. Après un premier saut vers le haut, la marche aléatoire est réinitialisée depuis le point de départ 1 et a une probabilité $p_{n-1,x-1}$ d'atteindre le point x à l'étape n.

On en déduit que[24]

$$\begin{aligned}
\mathbb{P}(H_{n,x} \mid E_{n,x}) &= \frac{\mathbb{P}(H_{n,x} \cap E_{n,x})}{\mathbb{P}(E_{n,x})} \\
&= \frac{\mathbb{P}(H_{n,x})}{\mathbb{P}(E_{n,x})} \quad \text{car } H_{n,x} \subset E_{n,x} \\
&= \frac{\frac{x}{n} p_{n,x}}{p_{n,x}} \\
&= \frac{x}{n}.
\end{aligned}$$

10. Les trajectoires ne touchant jamais 0 en n pas sont de deux types : celles qui visitent uniquement des valeurs positives et celles visitant uniquement des valeurs négatives. Or les vecteurs ω tels que $S_k(\omega) > 0$ pour tout $k \in [\![1, n]\!]$ sont aussi nombreux que les vecteurs vérifiant $S_k(\omega) < 0$ pour tout $k \in [\![1, n]\!]$ (on peut passer d'un ensemble à l'autre par la bijection $\omega \mapsto -\omega$, qui symétrise intégralement les trajectoires associées par rapport l'axe des abscisses) ; par équiprobabilité des trajectoires, on a donc

$$\mathbb{P}(S_1 > 0, \ldots, S_n > 0) = \mathbb{P}(S_1 < 0, \ldots, S_n < 0),$$

d'où

$$\begin{aligned}
\mathbb{P}(S_1 \neq 0, \ldots, S_n \neq 0) &= \mathbb{P}(S_1 > 0, \ldots, S_n > 0) + \mathbb{P}(S_1 < 0, \ldots, S_n < 0) \\
&= 2\,\mathbb{P}(S_1 > 0, \ldots, S_n > 0).
\end{aligned}$$

Or l'événement $\{S_1 > 0, \ldots, S_n > 0\}$ est réalisé si et seulement si l'un des $H_{n,x}$, pour x pair appartenant à $[\![1, n]\!]$, est réalisé, et les $H_{n,x}$ correspondant à des valeurs de x différentes sont deux à deux incompatibles, d'où

$$\mathbb{P}(S_1 > 0, \ldots, S_n > 0) = \mathbb{P}\left(\bigsqcup_{k=1}^{\frac{n}{2}} H_{n,2k}\right) = \sum_{k=1}^{\frac{n}{2}} \mathbb{P}(H_{n,2k}).$$

Or on a vu dans la question précédente que

$$\mathbb{P}(H_{n,2k}) = \frac{p_{n-1,2k-1} - p_{n-1,2k+1}}{2}$$

pour tout $k \in [\![1, \frac{n}{2}]\!]$, donc

$$\sum_{k=1}^{\frac{n}{2}} \mathbb{P}(H_{n,2k}) = \frac{1}{2} \sum_{k=1}^{\frac{n}{2}} (p_{n-1,2k-1} - p_{n-1,2k+1}) = \frac{1}{2}\,(p_{n-1,1} - p_{n-1,n+1})$$

24. Le résultat démontré ici est le *théorème du scrutin* : si lors d'une élection, chacune des n personnes d'une population vote pour un candidat choisi uniformément au hasard parmi deux candidats, et si un candidat arrive en tête de x voix au terme du scrutin, alors la probabilité que le candidat ait été en tête durant l'intégralité du dépouillement est égale à $\frac{x}{n}$.

en repérant une somme télescopique, soit

$$\begin{aligned}\sum_{k=1}^{\frac{n}{2}} \mathbb{P}(H_{n,2k}) &= \frac{1}{2}\left(p_{n-1,1} - 0\right) = \frac{1}{2} \cdot \frac{1}{2^{n-1}} \binom{n-1}{\frac{n}{2}} \\ &= \frac{1}{2^n} \frac{(n-1)!}{\left(\frac{n}{2}\right)!\left(\frac{n}{2}-1\right)!} = \frac{1}{2^n} \cdot \frac{\frac{n}{2}}{n} \cdot \frac{n!}{\left(\frac{n}{2}\right)!\left(\frac{n}{2}\right)!} = \frac{1}{2^{n+1}} \binom{n}{\frac{n}{2}},\end{aligned}$$

d'où enfin

$$\mathbb{P}(S_1 \neq 0, \ldots, S_n \neq 0) = 2 \cdot \frac{1}{2^{n+1}} \binom{n}{\frac{n}{2}} = \frac{1}{2^n} \binom{n}{\frac{n}{2}} = q_n.$$

11. Si n est impair, l'événement $\{S_n = 0\}$ ne peut être réalisé donc la probabilité à calculer est nulle. Supposons donc à nouveau que n est pair. Le cas $n = 2$ est facile et donne $\mathbb{P}(S_1 \neq 0, S_2 = 0) = \frac{1}{2}$. Supposons à présent que $n \geqslant 4$. Comme l'événement $\{S_{n-1} \neq 0\}$ est certain, on peut écrire

$$\begin{aligned}&\mathbb{P}(S_1 \neq 0, \ldots, S_{n-1} \neq 0, S_n = 0) \\ &\qquad = \mathbb{P}(S_1 \neq 0, \ldots, S_{n-2} \neq 0) - \mathbb{P}(S_1 \neq 0, \ldots, S_{n-1} \neq 0, S_n \neq 0),\end{aligned}$$

soit, en utilisant le résultat de la question précédente pour n et $n-2$ (ce qui est possible car $n \geqslant 4$) :

$$\begin{aligned}\mathbb{P}(S_1 \neq 0, \ldots, S_{n-1} \neq 0, S_n = 0) &= q_{n-2} - q_n \\ &= \frac{1}{2^{n-2}} \binom{n-2}{\frac{n}{2}-1} - \frac{1}{2^n} \binom{n}{\frac{n}{2}} \\ &= \frac{1}{2^{n-2}} \frac{(n-2)!}{\left(\frac{n}{2}-1\right)!\left(\frac{n}{2}-1\right)!} - \frac{1}{2^n} \binom{n}{\frac{n}{2}} \\ &= \frac{1}{2^{n-2}} \frac{\frac{n}{2} \cdot \frac{n}{2}}{n(n-1)} \frac{n!}{\left(\frac{n}{2}\right)!\left(\frac{n}{2}\right)!} - \frac{1}{2^n} \binom{n}{\frac{n}{2}} \\ &= \frac{1}{2^{n-2}} \frac{n}{4(n-1)} \binom{n}{\frac{n}{2}} - \frac{1}{2^n} \binom{n}{\frac{n}{2}} \\ &= \frac{1}{2^n(n-1)} \binom{n}{\frac{n}{2}}.\end{aligned}$$

On remarque que cette formule reste valable lorsque $n = 2$.

12. Considérons qu'au lieu de décrire le comportement d'une particule, on tienne le compte des piles et des faces obtenus lors de la réalisation de tirages successifs d'une pièce équilibrée, en partant de 0, en ajoutant 1 à chaque pile et en ajoutant -1 à chaque face. Dans ce cas, pour tout $n \in [\![1, N]\!]$, la variable aléatoire S_n donne la différence du nombre de piles et du nombre de faces obtenus lors des n premiers lancers. Notons Z la variable aléatoire indiquant le nombre de tirages nécessaires pour tirer exactement autant de piles que de faces ; pour tout $n \in [\![1, N]\!]$, on a alors

$$\mathbb{P}(Z = n) = \mathbb{P}(S_1 \neq 0, \ldots, S_{n-1} \neq 0, S_n = 0) = \frac{1}{2^n(n-1)} \binom{n}{\frac{n}{2}}.$$

Supposons à présent que l'on réalise une infinité de lancers ; dans ce cas, Z est une VARD à valeurs dans l'ensemble des entiers pairs supérieurs ou égaux à 2, dont la loi de probabilité est donnée[25] par

$$\mathbb{P}(Z = n) = \frac{1}{2^n(n-1)}\binom{n}{\frac{n}{2}} = \frac{q_n}{n-1}$$

pour $n \geqslant 2$ pair. Or

$$\sum_{n=2}^{+\infty} n\,\mathbb{P}(Z = n) = \sum_{k=1}^{+\infty} 2k\,\frac{q_{2k}}{2k-1}$$

et

$$2k\,\frac{q_{2k}}{2k-1} \underset{k\to+\infty}{\sim} q_{2k} \underset{k\to+\infty}{\sim} \frac{1}{\sqrt{\pi k}}$$

d'après la question 5, donc la série à termes positifs $\sum\limits_{k\geqslant 1} 2k\,\frac{q_{2k}}{2k-1}$ est de même nature que la série de terme général $\frac{1}{\sqrt{\pi k}}$, qui diverge en tant que série de Riemann de paramètre $\frac{1}{2} \leqslant 1$. Ainsi, la série de terme général $n\mathbb{P}(Z = n)$ diverge, ce qui signifie que Z n'admet pas d'espérance.

Le temps de retour en 0 d'une marche aléatoire symétrique sur $\mathbb{Z}$ est ainsi une variable aléatoire bien définie (le retour se fait avec probabilité 1) mais sans espérance (ce temps prend des valeurs élevées avec des probabilités trop grandes pour que le fait de parler de sa valeur moyenne ait un sens). La question posée par l'énoncé admet donc une réponse subtile : il n'y a pas, à strictement parler, de *nombre moyen* de parties à jouer avant d'avoir tiré exactement autant de piles que de faces – ou encore, en le disant autrement : ce nombre moyen existe mais est infini.

On pourra trouver davantage d'informations sur les propriétés des marches aléatoires symétriques simples en dimension quelconque dans l'exercice 60.

Correction de l'exercice 52.

1. La variable aléatoire $Z := \frac{X+1}{2}$ prend les valeurs 0 et 1, et

$$\mathbb{P}(Z = 1) = \mathbb{P}(X = 1) = \frac{1}{2}$$

donc $Z \sim \mathcal{B}(\frac{1}{2})$.

2. Comme $X = 2Z - 1$ et $Z \sim \mathcal{B}(\frac{1}{2})$, on a

$$\mathbb{E}(X) = 2\,\mathbb{E}(Z) - 1 = 2 \cdot \frac{1}{2} - 1 = 0$$

par linéarité de l'espérance, et

$$\mathbb{V}(X) = 4\mathbb{V}(Z) = 4 \cdot \frac{1}{2}\left(1 - \frac{1}{2}\right) = 1$$

25. Il est aisé de vérifier que la somme de ces probabilités vaut bien 1 en écrivant que pour tout $n \geqslant 2$ on a $\frac{1}{2^n(n-1)}\binom{n}{\frac{n}{2}} = q_{n-2} - q_n$. Cela justifie qu'il arrivera presque sûrement un rang auquel on aura obtenu le même nombre de piles que de faces, donc que Z est bien définie.

par quadraticité de la variance.

Alternativement, on pouvait bien sûr effectuer un calcul direct :

$$\mathbb{E}(X) = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 1 = 0$$

et d'après la formule de Koenig,

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \mathbb{E}(1) - 0^2 = 1.$$

3. (a) La variable aléatoire XY prend les valeurs -1 et 1, et on a

$$\begin{aligned}
\mathbb{P}(XY = 1) &= \mathbb{P}((X = 1) \cap (Y = 1)) + \mathbb{P}((X = -1) \cap (Y = -1)) \quad \text{par incompatibilité} \\
&= \mathbb{P}(X = 1)\, \mathbb{P}(Y = 1) + \mathbb{P}(X = -1)\, \mathbb{P}(Y = -1) \quad \text{par indépendance} \\
&= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{1}{2},
\end{aligned}$$

donc XY suit une loi de Rademacher.

(b) Les VARD X et Y sont indépendantes par hypothèse. Montrons que X et XY sont indépendantes. Il suffit pour cela de montrer que les événements $(X = 1)$ et $(XY = 1)$ sont indépendants, puisque la stabilité de l'indépendance par passage aux complémentaires assure que $(X = -1)$ et $(XY = 1)$ d'une part, $(X = 1)$ et $(XY = -1)$ d'autre part, et enfin $(X = -1)$ et $(XY = -1)$, le sont aussi. Or

$$\begin{aligned}
\mathbb{P}((X = 1) \cap (XY = 1)) &= \mathbb{P}((X = 1) \cap (Y = 1)) \\
&= \mathbb{P}(X = 1)\mathbb{P}(Y = 1) = \frac{1}{4}
\end{aligned}$$

par indépendance de X et Y, mais $\mathbb{P}(X = 1)\, \mathbb{P}(XY = 1) = \frac{1}{4}$ donc les événements $(X = 1)$ et $(XY = 1)$ sont indépendants. Ainsi, X et XY sont indépendantes. Un raisonnement symétrique montre que Y et XY sont indépendantes. On en déduit que X, Y et XY sont deux à deux indépendantes.

En revanche, on a

$$\mathbb{P}((X = 1) \cap (Y = -1) \cap (XY = 1)) = \mathbb{P}(\varnothing) = 0$$

et

$$\mathbb{P}(X = 1)\, \mathbb{P}(Y = -1)\, \mathbb{P}(XY = 1) = \frac{1}{8}$$

donc

$$\mathbb{P}((X = 1) \cap (Y = -1) \cap (XY = 1)) \neq \mathbb{P}(X = 1)\, \mathbb{P}(Y = -1)\, \mathbb{P}(XY = 1),$$

ce qui montre que X, Y et XY ne sont pas mutuellement indépendantes.

4. Comme XZ et Z prennent un nombre fini de valeurs, elles admettent toutes deux un moment d'ordre 2.

   Par indépendance de X et Z, on a $\mathbb{E}(XZ) = \mathbb{E}(X)\,\mathbb{E}(Z) = 0 \cdot \frac{1}{2} = 0$.

   On a à présent, en utilisant la formule de Koenig et le fait que $Z^2 = Z$:

$$\mathrm{Cov}(XZ, Z) = \mathbb{E}(XZZ) - \mathbb{E}(XZ)\,\mathbb{E}(Z) = \mathbb{E}(XZ) - \mathbb{E}(XZ)\,\mathbb{E}(Z) = 0.$$

   Ainsi, XZ et Z sont bien décorrélées.

   En revanche, XZ et Z ne sont pas indépendantes[26] puisque

$$\mathbb{P}(XZ = 1) = \mathbb{P}((X = 1) \cap (Z = 1)) = \mathbb{P}(X = 1)\,\mathbb{P}(Z = 1) = \frac{1}{4}$$

   et

$$\mathbb{P}((XZ = 1) \cap (Z = 0)) = \mathbb{P}(\varnothing) = 0 \neq \frac{1}{4} \cdot \frac{1}{2} = \mathbb{P}(XZ = 1)\,\mathbb{P}(Z = 0).$$

Correction de l'exercice 53.

1. Supposons que l'on réalise un tirage sans remise de n boules dans une urne contenant A boules dont pA boules rouges et $(1-p)A$ boules noires, et notons X le nombre de boules rouges tirées. On veut vérifier que X suit la loi annoncée.

   Le nombre de boules rouges tirées est un entier entre 0 et n puisque l'on a tiré n boules. Par ailleurs, comme le tirage a lieu sans remise, cet entier ne peut excéder le nombre pA de boules rouges dans l'urne : ainsi, X prend des valeurs inférieures ou égales à $\min(pA, n)$. De façon symétrique, le nombre de boules noires tirées ne peut excéder $(1-p)A$, donc si $(1-p)A \leqslant n$, alors le nombre de boules rouges tirées est au moins de $n-(1-p)A$ puisque n boules sont tirées en tout : ainsi, X prend des valeurs supérieures ou égales à $\max(0, n-(1-p)A)$. On a donc bien $X(\Omega) \subset [\![\max(0, n-(1-p)A), \min(pA, n)]\!]$.

   Pour établir l'inclusion réciproque et achever de déterminer la loi de X, fixons $k \in [\![\max(0, n-(1-p)A), \min(pA, n)]\!]$ et calculons $\mathbb{P}(X = k)$, dont on montrera qu'elle est strictement positive. Si l'on choisit de ne pas tenir compte de l'ordre dans lequel les boules sont tirées, le tirage de n boules sans remise dans l'urne s'apparente au tirage instantané et équiprobable d'une combinaison de n boules parmi les A boules que contient l'urne. Or il existe $\binom{A}{n}$ telles combinaisons, et parmi elles, il existe exactement $\binom{pA}{k}\binom{(1-p)A}{n-k}$ combinaisons contenant k boules rouges et $n-k$ boules noires (puisqu'une combinaison de ce type est construite en choisissant k boules rouges parmi les pA disponibles,

26. Ce phénomène de décorrélation sans indépendance s'explique en remarquant que la valeur prise par X n'a pas d'influence sur la valeur moyenne de XZ (qui est nulle), mais qu'elle a une influence sur les valeurs effectivement prises par XZ (puisque par exemple XZ prend nécessairement la valeur 0 lorsque X prend la valeur 0, et ne peut pas la prendre sinon).

puis $n-k$ boules noires parmi les $(1-p)A$ disponibles) : ainsi, la probabilité de tirer exactement k boules rouges vaut

$$\mathbb{P}(X=k) = \frac{\binom{pA}{k}\binom{(1-p)A}{n-k}}{\binom{A}{n}}$$

par équiprobabilité. Comme cette probabilité est strictement positive, on en déduit que $X(\Omega) = [\![\max(0, n-(1-p)A), \min(pA, n)]\!]$ et X suit donc bien la loi décrite dans l'énoncé.

2. La loi $\mathcal{H}(1, p, A)$ est la loi suivie par le nombre de boules rouges tirées lors d'un tirage unique réalisé dans l'urne. Comme la probabilité de tirer une boule rouge lors de ce tirage vaut p, il s'agit de la loi de Bernoulli de paramètre p : on a $\mathcal{H}(1, p, A) = \mathcal{B}(p)$.

3. D'après la formule de Vandermonde (voir l'exercice 25 du chapitre « Polynômes »), on a
$$\sum_{k=0}^{n}\binom{pA}{k}\binom{(1-p)A}{n-k} = \binom{A}{n}.$$
Le terme $\binom{pA}{k}\binom{(1-p)A}{n-k}$ étant nul si $k > pA$ ou $n-k > (1-p)A$ (c'est-à-dire $k < n-(1-p)A$), on en déduit que
$$\sum_{k=\max(0,n-(1-p)A)}^{\min(pA,n)}\binom{pA}{k}\binom{(1-p)A}{n-k} = \sum_{k=0}^{n}\binom{pA}{k}\binom{(1-p)A}{n-k} = \binom{A}{n},$$
ce qui permet bien d'affirmer que
$$\sum_{k=\max(0,n-(1-p)A)}^{\min(pA,n)}\frac{\binom{pA}{k}\binom{(1-p)A}{n-k}}{\binom{A}{n}} = 1,$$
et donc que la somme des $\mathbb{P}(X=k)$ vaut 1.

4. (a) Soit $k \in [\![\max(1, n-(1-p)A+1), \min(pA, n)]\!]$ (c'est-à-dire $k \in X(\Omega)$ tel que $k \neq 0$). La formule du capitaine donne alors
$$k\binom{pA}{k} = pA\binom{pA-1}{k-1},$$
et il suffit de multiplier cette égalité par $\binom{qA}{n-k} = \binom{qA}{(n-1)-(k-1)}$ pour obtenir la relation attendue.

(b) Comme X prend un nombre fini de valeurs, $\mathbb{E}(X)$ existe et on a
$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=\max(0,n-(1-p)A)}^{\min(pA,n)} k\,\frac{\binom{pA}{k}\binom{(1-p)A}{n-k}}{\binom{A}{n}}\\ &= \frac{1}{\binom{A}{n}}\sum_{k=\max(1,n-(1-p)A+1)}^{\min(pA,n)} k\binom{pA}{k}\binom{(1-p)A}{n-k} \quad \text{car le terme d'indice 0 est nul}\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\binom{A}{n}} \sum_{k=\max(1,n-(1-p)A+1)}^{\min(pA,n)} pA \binom{pA-1}{k-1}\binom{(1-p)A}{(n-1)-(k-1)} \quad \text{par la question 4(a)} \\
&= \frac{pA}{\binom{A}{n}} \sum_{j=\max(0,n-(1-p)A)}^{\min(pA-1,n-1)} \binom{pA-1}{j}\binom{(1-p)A}{(n-1)-j} \quad \text{en posant } j=k-1 \\
&= \frac{pA}{\binom{A}{n}} \binom{A-1}{n-1} \quad \text{par la formule de Vandermonde} \\
&= \frac{pA}{\frac{A}{n}} \quad \text{d'après la formule du capitaine} \\
&= np.
\end{aligned}
$$

(c) Par un raisonnement semblable à celui mené dans la question 4(a), on obtient la formule attendue grâce à une double application de la formule du capitaine.

(d) En raisonnant comme dans la question 4(b) et en utilisant la relation démontrée dans la question 4(c), on trouve que $\mathbb{E}(X(X-1))$ existe et vaut

$$
\begin{aligned}
\mathbb{E}(X(X-1)) &= \frac{pA(pA-1)}{\binom{A}{n}} \binom{A-2}{n-2} \\
&= \frac{pA(pA-1)}{\frac{A}{n}\cdot\frac{A-1}{n-1}} = \frac{n(n-1)pA(pA-1)}{A(A-1)}.
\end{aligned}
$$

Ainsi, par linéarité de l'espérance, $X^2 = X(X-1)+X$ admet une espérance égale à

$$
\mathbb{E}(X^2) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) = \frac{n(n-1)pA(pA-1)}{A(A-1)} + np,
$$

donc X admet une variance égale par la formule de Koenig à

$$
\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\
&= \frac{n(n-1)pA(pA-1)}{A(A-1)} + np - n^2p^2 \\
&= \frac{n(n-1)p(pA-1)}{A-1} + np(1-np) \\
&= \frac{np}{A-1}\Big((n-1)(pA-1) + (1-np)(A-1)\Big) \\
&= \frac{np}{A-1}\big(np - n + A - pA\big) \\
&= \frac{n(A-n)p(1-p)}{A-1}.
\end{aligned}
$$

5. (a) X étant le nombre de boules rouges tirées, c'est-à-dire la variable de comptage du nombre d'événements E_i réalisés, on a

$$
X = \sum_{i=1}^{pA} \mathbb{1}_{E_i}.
$$

(b) Le nombre de combinaisons de n boules tirées parmi A et contenant la boule rouge numéro 1 est $\binom{A-1}{n-1}$, tandis que le nombre de combinaisons total est $\binom{A}{n}$, d'où $\mathbb{P}(E_1) = \frac{\binom{A-1}{n-1}}{\binom{A}{n}} = \frac{n}{A}$ par équiprobabilité [27]. On a de même $\mathbb{P}(E_i) = \frac{n}{A}$ pour tout $i \in [\![1, pA]\!]$.

Par linéarité de l'espérance, on a donc

$$\mathbb{E}(X) = \sum_{i=1}^{pA} \mathbb{E}\left(\mathbb{1}_{E_i}\right) = \sum_{i=1}^{pA} \mathbb{P}(E_i) = pA \cdot \frac{n}{A} = np.$$

(c) Soit $(i,j) \in [\![1, n]\!]^2$. Si $i = j$, on a

$$\operatorname{Cov}(\mathbb{1}_{E_i}, \mathbb{1}_{E_j}) = \mathbb{V}(X_i) = \mathbb{P}(E_i)(1 - \mathbb{P}(E_i)) = \frac{n}{A}\left(1 - \frac{n}{A}\right) = \frac{n(A-n)}{A^2}$$

puisque $\mathbb{1}_{E_i} \sim \mathcal{B}(\frac{n}{A})$.

Supposons à présent que $i \neq j$. On a alors $\mathbb{1}_{E_i}\mathbb{1}_{E_j} = \mathbb{1}_{E_i \cap E_j}$, or l'événement $E_i \cap E_j$ est réalisé si les boules rouges numéro i et j figurent parmi les n boules tirées, et il existe $\binom{A-2}{n-2}$ combinaisons satisfaisant cette condition sur $\binom{A}{n}$ combinaisons en tout, donc

$$\mathbb{P}(E_i \cap E_j) = \frac{\binom{A-2}{n-2}}{\binom{A}{n}} = \frac{n(n-1)}{A(A-1)}$$

par équiprobabilité. Ainsi, $\mathbb{E}(\mathbb{1}_{E_i}\mathbb{1}_{E_j}) = \mathbb{P}(E_i \cap E_j) = \frac{n(n-1)}{A(A-1)}$. D'après la formule de Koenig, la covariance de $\mathbb{1}_{E_i}$ et $\mathbb{1}_{E_j}$ est donc

$$\begin{aligned}\operatorname{Cov}(\mathbb{1}_{E_i}, \mathbb{1}_{E_j}) &= \mathbb{E}(\mathbb{1}_{E_i}\mathbb{1}_{E_j}) - \mathbb{E}(\mathbb{1}_{E_i})\,\mathbb{E}(\mathbb{1}_{E_j}) \\ &= \frac{n(n-1)}{A(A-1)} - \frac{n^2}{A^2} = \frac{n(n-A)}{A^2(A-1)}.\end{aligned}$$

Ainsi, on a

$$\begin{aligned}\mathbb{V}(X) &= \sum_{i=1}^{pA} \mathbb{V}(\mathbb{1}_{E_i}) + \sum_{1 \leqslant i \neq j \leqslant pA} \operatorname{Cov}(\mathbb{1}_{E_i}, \mathbb{1}_{E_j}) \\ &= \sum_{i=1}^{pA} \frac{n(A-n)}{A^2} + \sum_{1 \leqslant i \neq j \leqslant pA} \frac{n(n-A)}{A^2(A-1)} \\ &= pA \cdot \frac{n(A-n)}{A^2} + pA(pA-1) \cdot \frac{n(n-A)}{A^2(A-1)} \\ &\underset{\cdots}{=} \frac{n(A-n)p(1-p)}{A-1},\end{aligned}$$

ce qui redonne le résultat de la question 4(c).

27. Une autre démonstration de ce fait est la suivante : on tire en tout une proportion $\frac{n}{A}$ des boules de l'urne. La boule rouge numéro 1 étant indiscernable des autres durant le tirage, elle a la même probabilité que chaque autre boule d'être tirée, donc une probabilité $\frac{n}{A}$. Ainsi, $\mathbb{P}(E_1) = \frac{n}{A}$.

Correction de l'exercice 54.

1. La VARD N_i compte le nombre d'expériences ayant amené le résultat i lors de la répétition de n expériences indépendantes amenant chacune ce résultat avec probabilité p_i : elle suit donc la loi binomiale $\mathcal{B}(n, p_i)$.
2. Toutes les configurations de résultats sont possibles pour peu que le nombre total de résultats obtenus soit égal à n, c'est-à-dire que $\sum_{i=1}^{m} n_i = n$.

   On a donc

   $$S_{m,n} = \{(n_1, \ldots, n_m) \in [\![0, n]\!]^m \ : \ n_1 + \ldots + n_m = n\}.$$

3. Soit $(n_1, \ldots, n_m) \in S_{m,n}$. Dénombrons tout d'abord les *listes* (ordonnées) de m résultats contenant n_1 fois le résultat 1, n_2 fois le résultat 2,... , n_m fois le résultat m. Pour créer une telle liste, il faut :
   - Choisir, parmi les m emplacements disponibles, les n_1 rangs auxquels on obtient le résultat 1 : on dispose pour cela de $\binom{m}{n_1}$ choix.
   - Choisir ensuite, parmi les $m - n_1$ emplacements restants, les n_2 rangs auxquels on obtient le résultat 2 : on dispose pour cela de $\binom{m-n_1}{n_2}$ choix.
   - $\ldots$
   - Choisir enfin, parmi les $m - n_1 - \ldots - n_{m-1} = n_m$ emplacements restants, les n_m rangs auxquels on obtient le résultat m : on dispose pour cela de $\binom{m-n_1-\ldots-n_{m-1}}{n_m}$ choix.

   On dispose donc en tout de

   $$\begin{aligned}\binom{m}{n_1}\binom{m-n_1}{n_2}\cdots\binom{m-n_1-\ldots-n_{m-1}}{n_m} &= \frac{m!}{n_1!\cancel{(m-n_1)}!} \cdot \frac{\cancel{(m-n_1)}!}{n_2!\cancel{(m-n_1-n_2)}!} \cdots \frac{\cancel{(m-n_1-\ldots-n_{m-1})}!}{n_m!\cancel{(m-n_1-\ldots-n_m)}!} \\ &= \frac{m!}{n_1!n_2!\cdots n_m!}\end{aligned}$$

   listes contenant n_1 fois le résultat 1, n_2 fois le résultat 2 et ainsi de suite[28].

   Or par indépendance des expériences effectuées, chacune de ces listes a une probabilité égale à $p_1^{n_1} \ldots p_m^{n_m}$ de se réaliser. Par incompatibilité des différentes listes, on a donc

   $$\mathbb{P}((N_1, N_2, \ldots, N_m) = (n_1, \ldots, n_m)) = \frac{n!}{n_1! \cdots n_m!} p_1^{n_1} \ldots p_m^{n_m},$$

   ce qu'il fallait démontrer.

28. Le *coefficient multinomial* $\frac{m!}{n_1!n_2!\ldots n_m!}$ ainsi obtenu, parfois noté $\binom{m}{n_1, n_2, \ldots, n_m}$, est la généralisation du coefficient binomial à des listes de résultats non binaires.

On peut retrouver cette formule en disant qu'une liste de longueur m contenant n_1 fois le résultat 1, n_2 fois le résultat 2 et ainsi de suite peut être permutée de $n_1!n_2!\cdots n_m!$ façons sans changer la valeur de cette liste (en permutant les 1 entre eux, les 2 entre eux et ainsi de suite), et qu'en considérant les $n_1!n_2!\ldots n_m!$ permutations de toutes les listes vérifiant ces conditions on obtient les $m!$ permutations possibles de la liste (unique) contenant n_1 fois le résultat 1 *puis* n_2 fois le résultat 2 et ainsi de suite.

4. Comme N_i et N_j prennent un nombre fini de valeurs, elles admettent un moment d'ordre 2 donc leur covariance existe.

Remarquons que si p_i ou p_j est nul ou égal à 1, alors N_i ou N_j est constante donc leur covariance est nulle. On suppose donc désormais que $p_i, p_j \in]0,1[$.

Si $i = j$, on a $\mathrm{Cov}(N_i, N_j) = \mathbb{V}(N_i) = np_i(1-p_i)$ puisque N_i suit la loi $\mathcal{B}(n, p_i)$.

Supposons à présent que $i \neq j$. La VARD $N_i + N_j$ compte le nombre d'expériences ayant amené le résultat i ou le résultat j parmi les n expériences indépendantes réalisées ; comme la probabilité pour une expérience donnée d'amener un tel résultat est égale à $p_i + p_j$, on a $N_i + N_j \sim \mathcal{B}(n, p_i + p_j)$ donc $\mathbb{V}(N_i + N_j) = n(p_i + p_j)(1 - p_i - p_j)$. La formule

$$\mathbb{V}(N_i + N_j) = \mathbb{V}(N_i) + \mathbb{V}(N_j) + 2\,\mathrm{Cov}(N_i, N_j)$$

donne alors

$$\begin{aligned}\mathrm{Cov}(N_i, N_j) &= \frac{\mathbb{V}(N_i + N_j) - \mathbb{V}(N_i) - \mathbb{V}(N_j)}{2} \\ &= \frac{n(p_i + p_j)(1 - p_i - p_j) - np_i(1 - p_i) - np_j(1 - p_j)}{2} \\ &= -np_ip_j.\end{aligned}$$

Ainsi, N_i et N_j sont négativement corrélées, ce qui n'est guère étonnant puisque le fait d'obtenir le résultat i à l'issue d'une expérience donnée empêche d'obtenir le résultat j à cette même expérience.

On pouvait aussi effectuer un calcul direct, que l'on détaille ci-dessous puisqu'il possède un intérêt intrinsèque. D'après la formule de Koenig, on a

$$\mathrm{Cov}(N_i, N_j) = \mathbb{E}(N_iN_j) - \mathbb{E}(N_i)\mathbb{E}(N_j).$$

Or le théorème de transfert pour la loi jointe (disponible en suivant ce lien) assure que

$$\mathbb{E}(N_iN_j) = \sum_{k=0}^{n}\sum_{\ell=0}^{n-k} k\ell\,\mathbb{P}((N_i = k) \cap (N_j = \ell)),$$

les bornes des sommes ayant été choisies de façon à prendre en compte tous les couples $(k, \ell) \in [\![0, n]\!]^2$ tels que $k + \ell \leqslant n$. Soit (k, ℓ) un tel couple ; on souhaite calculer $\mathbb{P}((N_i = k) \cap (N_j = \ell))$, c'est-à-dire la probabilité d'obtenir exactement k fois le résultat i et ℓ fois le résultat j. Or la probabilité d'obtenir une liste de résultats *commençant par k fois i puis ℓ fois j et se terminant par $n-k-\ell$ résultats autres que i et j* est égale, par indépendance des expériences, à $p_i^k p_j^\ell(1-p_i-p_j)^{n-k-\ell}$. Or il existe $\binom{n}{k}\binom{n-k}{\ell}$ façons de placer les k résultats égaux à i et les ℓ résultats égaux à j dans la liste de n résultats, et chacune de ces configurations a une probabilité d'occurrence [29] égale à $p_i^k p_j^\ell(1 - p_i - p_j)^{n-k-\ell}$. Ainsi, on a

$$\mathbb{P}((N_i = k) \cap (N_j = \ell)) = \binom{n}{k}\binom{n-k}{\ell} p_i^k p_j^\ell (1 - p_i - p_j)^{n-k-\ell}.$$

29. Cela ne revient pas à prétendre que les listes de résultats sont toutes équiprobables, mais que pour un placement donné, l'événement « On obtient des i et des j aux emplacements sélectionnées et des résultats différents de i et j ailleurs » est toujours de même probabilité $p_i^k p_j^\ell(1-p_i-p_j)^{n-k-\ell}$.

On en déduit que

$$\begin{aligned}
\mathbb{E}(N_iN_j) &= \sum_{k=0}^{n}\sum_{\ell=0}^{n-k} k\ell\binom{n}{k}\binom{n-k}{\ell}p_i^k p_j^\ell(1-p_i-p_j)^{n-k-\ell}\\
&= \sum_{k=0}^{n} k\binom{n}{k}p_i^k \sum_{\ell=0}^{n-k}\ell\binom{n-k}{\ell}p_j^\ell(1-p_i-p_j)^{n-k-\ell}.
\end{aligned}$$

Or pour tout $k \in [\![0, n-1]\!]$ on a d'après la formule du capitaine :

$$\begin{aligned}
\sum_{\ell=0}^{n-k}\ell\binom{n-k}{\ell}p_j^\ell(1-p_i-p_j)^{n-k-\ell} &= \sum_{\ell=1}^{n-k}(n-k)\binom{n-k-1}{\ell-1}p_j^\ell(1-p_i-p_j)^{n-k-\ell}\\
&= (n-k)\sum_{\ell'=0}^{n-k-1}\binom{n-k-1}{\ell'}p_j^{\ell'+1}(1-p_i-p_j)^{n-k-\ell'-1} \quad \text{en posant } \ell' = \ell-1\\
&= (n-k)p_j\sum_{\ell'=0}^{n-k-1}\binom{n-k-1}{\ell'}p_j^{\ell'}(1-p_i-p_j)^{n-k-1-\ell'}\\
&= (n-k)p_j(p_j+1-p_i-p_j)^{n-k-1} \quad \text{par la formule du binôme}\\
&= (n-k)p_j(1-p_i)^{n-k-1},
\end{aligned}$$

et on remarque que cette formule reste valide pour $k = n$. Ainsi,

$$\begin{aligned}
\mathbb{E}(N_iN_j) &= \sum_{k=0}^{n} k\binom{n}{k}p_i^k \sum_{\ell=0}^{n-k}\ell\binom{n-k}{\ell}p_j^\ell(1-p_i-p_j)^{n-k-\ell}\\
&= \sum_{k=0}^{n} k\binom{n}{k}(n-k)p_i^k p_j(1-p_i)^{n-k-1}\\
&= p_j\left(n\sum_{k=1}^{n} k\binom{n}{k}p_i^k(1-p_i)^{n-k-1} - \sum_{k=1}^{n} k^2\binom{n}{k}p_i^k(1-p_i)^{n-k-1}\right)\\
&= p_j(1-p_i)^{-1}\left(n\sum_{k=1}^{n} k\binom{n}{k}p_i^k(1-p_i)^{n-k} - \sum_{k=1}^{n} k^2\binom{n}{k}p_i^k(1-p_i)^{n-k}\right)\\
&= p_j(1-p_i)^{-1}\left(n^2p_i - np_i(1-p_i) - n^2p_i^2\right)\\
&= p_j(n^2p_i - np_i)\\
&= n(n-1)p_ip_j
\end{aligned}$$

en reconnaissant à la cinquième égalité l'espérance d'une loi $\mathcal{B}(n, p_i)$ (égale à np_i) et son moment d'ordre 2 (qui est égal, en utilisant la variance connue de la loi binomiale, à $np_i(1-p_i) + (np_i)^2$). On trouve donc bien le résultat attendu :

$$\begin{aligned}
\mathrm{Cov}(N_i, N_j) &= \mathbb{E}(N_iN_j) - \mathbb{E}(N_i)\mathbb{E}(N_j)\\
&= n(n-1)p_ip_j - np_inp_j\\
&= -np_ip_j.
\end{aligned}$$

Correction de l'exercice 55.

1. La fonction f est indéfiniment dérivable sur $]-1,1[$ en tant que fonction rationnelle bien définie sur $]-1,1[$. Pour tout $t \in [0,1[$, on a

$$f'(t) = (r+1)\frac{1}{(1-t)^{r+2}},$$

$$f''(t) = (r+1)(r+2)\frac{1}{(1-t)^{r+3}},$$

et on démontre plus généralement par une récurrence facile que pour tout $k \geqslant 1$ on a

$$f^{(k)}(t) = (r+1)(r+2)\cdots(r+k)\frac{1}{(1-t)^{r+k+1}} = \frac{(r+k)!}{r!}\frac{1}{(1-t)^{r+k+1}}.$$

2. On a
$$\binom{n+r}{n} = \frac{(n+r)!}{n!\,r!} = \frac{1}{r!}\prod_{i=1}^{r}(n+i) \underset{n\to+\infty}{\sim} \frac{n^r}{r!}$$
grâce au fait que le produit $\prod_{i=1}^{r}(n+i)$ est composé de r termes tous équivalents à n lorsque n tend vers l'infini.

3. Si $x = 0$, on a bien sûr $n^{r+1}x^n = 0 \xrightarrow[n\to+\infty]{} 0$. Si $x \in]-1,1[\setminus\{0\}$, alors $|n^{r+1}x^n| = n^{r+1}\exp\left(n\ln(|x|)\right) \xrightarrow[n\to+\infty]{} 0$ d'après le théorème des croissances comparées puisque $\ln(|x|) < 0$.

4. Fixons $x \in]-1,1[$. La proposition [30] est vraie pour $n = 0$ puisque l'on a

$$\sum_{k=0}^{0}\binom{k+r}{k}x^k + (r+1)\binom{r}{0}\int_0^x \frac{\mathrm{d}t}{(1-t)^{r+2}} = 1 + (r+1)\left[\frac{1}{(r+1)(1-t)^{r+1}}\right]_0^x$$
$$= 1 + \frac{1}{(1-x)^{r+1}} - 1 = \frac{1}{(1-x)^{r+1}} = f(x).$$

Soit à présent un entier $n \geqslant 0$. Supposons que la relation proposée est vraie au rang n. On réalise alors une intégration par parties, en dérivant la fonction $t \mapsto \frac{1}{(1-t)^{n+r+2}}$ de classe $\mathcal{C}^1$ et en primitivant la fonction continue $t \mapsto (x-t)^n$, pour obtenir :

$$\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\mathrm{d}t = \left[-\frac{1}{n+1}\frac{(x-t)^{n+1}}{(1-t)^{n+r+2}}\right]_0^x + \frac{n+r+2}{n+1}\int_0^x \frac{(x-t)^{n+1}}{(1-t)^{n+r+3}}\mathrm{d}t$$
$$= \frac{x^{n+1}}{n+1} + \frac{n+r+2}{n+1}\int_0^x \frac{(x-t)^{n+1}}{(1-t)^{n+r+3}}\mathrm{d}t,$$

30. La proposition à démontrer est la *formule de Taylor avec reste intégral* pour la fonction f (voir l'exercice **??** du chapitre **??**).

d'où, en utilisant l'hypothèse de récurrence :

$$\begin{aligned}
f(x) &= \sum_{k=0}^{n} \binom{k+r}{k} x^k + (n+r+1)\binom{n+r}{n} \int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}} \mathrm{d}t \\
&= \sum_{k=0}^{n} \binom{k+r}{k} x^k + (n+r+1)\binom{n+r}{n} \frac{x^{n+1}}{n+1} \\
&\qquad\qquad + (n+r+1)\binom{n+r}{n} \frac{n+r+2}{n+1} \int_0^x \frac{(x-t)^{n+1}}{(1-t)^{n+r+3}} \mathrm{d}t \\
&= \sum_{k=0}^{n} \binom{k+r}{k} x^k + \binom{n+r+1}{n+1} x^{n+1} + (n+r+2)\binom{n+r+1}{n+1} \int_0^x \frac{(x-t)^{n+1}}{(1-t)^{n+1+r+2}} \mathrm{d}t \\
&= \sum_{k=0}^{n+1} \binom{k+r}{k} x^k + (n+r+2)\binom{n+r+1}{n+1} \int_0^x \frac{(x-t)^{n+1}}{(1-t)^{n+r+3}} \mathrm{d}t,
\end{aligned}$$

ce qui établit la propriété attendue au rang $n+1$.

Ainsi, la propriété à démontrer est vraie pour tout $n \in \mathbb{N}$ d'après le principe de récurrence.

5. La relation à démontrer est évidente si $x = 0$.

   Soit à présent $x \in]0,1[$. On a

$$(n+r+1)\binom{n+r}{n} \underset{n\to+\infty}{\sim} (n+r+1)\frac{n^r}{r!} \underset{n\to+\infty}{\sim} \frac{n^{r+1}}{r!}$$

   d'après la question 2. Par ailleurs, pour tout $t \in [0,x]$ on a

$$\frac{x-t}{1-t} \leqslant \frac{x-xt}{1-t} = x$$

   puisque $1-t > 0$, donc pour tout $n \in \mathbb{N}$ on a

$$\begin{aligned}
0 &\leqslant \int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}} \, \mathrm{d}t \\
&= \int_0^x \left(\frac{x-t}{1-t}\right)^n \cdot \frac{1}{(1-t)^{r+2}} \, \mathrm{d}t \\
&\leqslant \int_0^x x^n \cdot \frac{1}{(1-t)^{r+2}} \, \mathrm{d}t \\
&= x^n \int_0^x \frac{1}{(1-t)^{r+2}} \, \mathrm{d}t.
\end{aligned}$$

   Il suffit pour la suite de remarquer que l'intégrale $I := \displaystyle\int_0^x \frac{1}{(1-t)^{r+2}} \, \mathrm{d}t$ ne dépend pas de n, mais on peut la calculer si l'on veut : on trouve dans ce cas $I = \frac{1}{r+1}\left(\frac{1}{(1-x)^{r+1}} - 1\right)$. On déduit de ce fait et de la question précédente que

$$\begin{aligned}
0 &\leqslant (n+r+1)\binom{n+r}{n} \int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}} \, \mathrm{d}t \\
&\leqslant (n+r+1)\binom{n+r}{n} x^n I \\
&\underset{n\to+\infty}{\sim} \frac{n^{r+1}}{r!} x^n I,
\end{aligned}$$

or d'après le résultat de la question 3 on a

$$\frac{n^{r+1}}{r!}x^n \xrightarrow[n\to+\infty]{} 0,$$

d'où

$$(n+r+1)\binom{n+r}{n}\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\,\mathrm{d}t \xrightarrow[n\to+\infty]{} 0$$

d'après le théorème des gendarmes. En passant à la limite dans la relation démontrée dans la question 4, on obtient donc

$$\sum_{k=0}^{n}\binom{k+r}{k}x^k = f(x) - (n+r+1)\binom{n+r}{n}\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\,\mathrm{d}t \xrightarrow[n\to+\infty]{} f(x),$$

d'où la convergence de la série de terme général $\binom{k+r}{k}x^k$ et la relation recherchée.

Soit à présent $x \in]-1,0[$. La majoration à effectuer est plus simple dans ce cas : en effet, pour tout $t \in [x,0]$ on a $1-t \geqslant 1$ donc, pour tout $n \in \mathbb{N}$:

$$\frac{|x-t|^n}{|1-t|^{n+r+2}} \leqslant |x-t|^n = (t-x)^n.$$

Ainsi, on obtient grâce à l'inégalité triangulaire que pour tout $n \in \mathbb{N}$,

$$\begin{aligned}\left|\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\mathrm{d}t\right| &\leqslant \int_x^0 \left|\frac{(x-t)^n}{(1-t)^{n+r+2}}\right|\mathrm{d}t \\ &\leqslant \int_x^0 (t-x)^n\mathrm{d}t = \left[\frac{(t-x)^{n+1}}{n+1}\right]_x^0 \\ &= \frac{(-x)^{n+1}}{n+1} = \frac{|x|^{n+1}}{n+1}.\end{aligned}$$

On en déduit que

$$\begin{aligned}&\left|(n+r+1)\binom{n+r}{r}\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\mathrm{d}t\right| \\ &\leqslant (n+r+1)\binom{n+r}{r}\frac{|x|^{n+1}}{n+1} \\ &\underset{n\to+\infty}{\sim} \frac{n^r|x|^{n+1}}{r!},\end{aligned}$$

or $n^r|x|^{n+1} \xrightarrow[n\to+\infty]{} 0$ d'après le résultat de la question 3 puisque $|x| \in [0,1[$, donc

$$(n+r+1)\binom{n+r}{r}\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\mathrm{d}t \xrightarrow[n\to+\infty]{} 0$$

par encadrement, d'où une fois encore

$$\sum_{k=0}^{n}\binom{k+r}{k}x^k = f(x) - (n+r+1)\binom{n+r}{n}\int_0^x \frac{(x-t)^n}{(1-t)^{n+r+2}}\,\mathrm{d}t \xrightarrow[n\to+\infty]{} f(x),$$

ce qui permet de conclure.

6. Soit $z \in]-1,1[$ et soit $n \in \mathbb{N}$. La série de terme général

$$\begin{aligned}\binom{-n}{j} z^j &= \frac{(-n)(-n-1)\cdots(-n-j+1)}{j!} z^j \\ &= (-1)^j \frac{n(n+1)\cdots(n+j-1)}{j!} z^j = \binom{n+j-1}{j} (-z)^j\end{aligned}$$

converge d'après la question 5 (appliquée à $r = n-1$ et $x = -z$), et on a

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

7. Si $n = 1$, $X_n = X_1$ est la loi du temps d'attente du premier succès lors de la répétition d'épreuves de Bernoulli indépendantes de probabilité de succès p ; elle suit donc une loi géométrique de paramètre p. D'après le cours, son espérance vaut $\frac{1}{p}$ et sa variance $\frac{1-p}{p^2}$.

8. (a) Compte tenu de sa définition, X_n prend nécessairement des valeurs entières positives. Dans le cas où n succès sont obtenus, elle représente le temps d'attente du n-ème succès et ne peut donc prendre de valeurs strictement inférieures à n. Par contre, toute valeur entière $j \geqslant n$ peut être prise par X_n, par exemple dans le cas où la suite des résultats obtenus est composée de $j-n$ échecs suivis de n succès consécutifs. On a donc $X_n(\Omega) = \{0\} \cup \{k \in \mathbb{N} : k \geqslant n\}$.

   (b) Soit $k \in \mathbb{N}$. Il existe $\binom{n+k-1}{n-1}$ façons de placer exactement $n-1$ succès (et donc k échecs) dans la liste des $n+k-1$ premiers résultats obtenus. Par indépendance des expériences de Bernoulli réalisées, chaque liste de résultats ainsi décrite a une probabilité égale à $p^{n-1}(1-p)^k$ d'être obtenue lors des $n-1+k$ premières expériences. Par incompatibilité des différentes listes entre elles, la probabilité recherchée vaut donc $\binom{n+k-1}{n-1} p^{n-1}(1-p)^k$.

   (c) Si $k \in \mathbb{N}$, $(X_n = n+k)$ est réalisé si et seulement si les $n+k-1$ premières expériences amènent exactement $n-1$ succès et si la $(n+k)$-ième est un succès. Comme cette dernière expérience est indépendante des précédentes, on a

$$\mathbb{P}(X_n = n+k) = \binom{n+k-1}{n-1} p^{n-1}(1-p)^k \times p = \binom{n+k-1}{n-1} p^n (1-p)^k$$

   grâce la question précédente.

   (d) Si $N \in \mathbb{N}$, on a

$$\begin{aligned}\sum_{k=0}^{n+N} \mathbb{P}(X_n = n+k) &= \sum_{k=0}^{n+N} \binom{n+k-1}{n-1} p^n (1-p)^k \\ &= p^n \sum_{k=0}^{n+N} \binom{n+k-1}{k} (1-p)^k.\end{aligned}$$

En appliquant le résultat de la question 5 avec $r = n-1$ et $x = 1-p$, on obtient alors

$$\lim_{N\to+\infty} \sum_{k=0}^{n+N} \mathbb{P}(X_n = n+k) = p^n \sum_{k=0}^{+\infty} \binom{n+k-1}{k} (1-p)^k = \frac{p^n}{p^n} = 1$$

d'où la convergence de la série de terme général $\mathbb{P}(X_n = n+k)$, le fait que sa somme vaut 1 et l'égalité

$$\mathbb{P}(X_n = 0) = 1 - \sum_{k=0}^{+\infty} \mathbb{P}(X_n = n+k) = 0.$$

Ainsi, l'événement $(X_n = 0)$ est quasi-impossible, c'est-à-dire que l'on observe au moins n succès avec probabilité 1 au cours de la suite infinie d'épreuves réalisées.

9. (a) Il s'agit de la formule du capitaine : pour tout $k \in \mathbb{N}$, on a

$$\begin{aligned} \frac{n+k}{n}\binom{n+k-1}{n-1} &= \frac{n+k}{n} \cdot \frac{(n+k-1)!}{(n-1)!k!} \\ &= \frac{(n+k)!}{n!k!} = \binom{n+k}{k}, \end{aligned}$$

d'où le premier résultat attendu.

La VARD X_n admet une espérance si et seulement si la série

$$\sum_{k\geqslant 0} (n+k)\mathbb{P}(X_n = n+k) = \sum_{k\geqslant 0} (n+k)\binom{n+k-1}{n-1} p^n (1-p)^k$$

converge, or cette série peut s'écrire sous la forme

$$\frac{p^n}{n} \sum_{k\geqslant 0} \binom{n+k}{n-1} (1-p)^k = \frac{p^n}{n} \sum_{k\geqslant 0} \binom{n+k}{k} (1-p)^k$$

d'après la question précédente. Le résultat de la question 5 (avec $r = n$ et $x = 1-p$) nous assure alors que X_n admet une espérance et que celle-ci vaut

$$\mathbb{E}(X_n) = \frac{p^n}{n} \cdot \frac{1}{p^{n+1}} = \frac{n}{p}.$$

(b) Le temps d'attente X_1 du premier succès suit la loi $\mathcal{G}(p)$. Une fois ce premier succès atteint, l'indépendance des épreuves de Bernoulli implique que le temps d'attente du deuxième succès depuis le premier succès, c'est-à-dire $X_2 - X_1$, suit lui aussi la loi $\mathcal{G}(p)$ et est indépendant de X_1. Plus généralement, les *temps inter-succès* $X_{i+1} - X_i$ (en posant $X_0 := 0$) sont indépendants les uns des autres et de même loi $\mathcal{G}(p)$. Ainsi, la VARD

$$X_n = \sum_{k=0}^{n-1} (X_{i+1} - X_i)$$

s'écrit comme une somme de n variables aléatoires indépendantes de même loi $\mathcal{G}(p)$.

On retrouve alors, par linéarité de l'espérance, le fait que X_n admet une espérance et que

$$\mathbb{E}(X_n) = \sum_{k=0}^{n-1} \mathbb{E}(X_{i+1} - X_i) = \sum_{k=0}^{n-1} \frac{1}{p} = \frac{n}{p}.$$

On retrouve aussi qu'en tant que somme de VARD admettant une variance, X_n en admet aussi une, et que l'on a

$$\begin{aligned}\mathbb{V}(X_n) &= \sum_{k=0}^{n-1} \mathbb{V}(X_{i+1} - X_i) \quad \text{par indépendance des } X_{i+1} - X_i \\ &= \sum_{k=0}^{n-1} \frac{1-p}{p^2} \\ &= \frac{n(1-p)}{p^2}.\end{aligned}$$

Bien entendu, un calcul direct mais fastidieux en utilisant l'expression de la loi de X_n comme dans la question 9(a) aurait aussi permis de trouver ce résultat.

(c) Soit $m \in \mathbb{N}$. On va montrer que X_n admet un moment d'ordre m. Soit $q \in]1-p, 1[$; on a

$$(n+k)^m p^n (1-p)^k \underset{k \to +\infty}{=} o\left(q^{k-1}\right)$$

par croissance comparée, donc

$$\begin{aligned}(n+k)^m \, \mathbb{P}(X_n = n+k) &= (n+k)^m \binom{n+k-1}{n-1} p^n (1-p)^k \\ &= (n+k)^m \binom{n+k-1}{k} p^n (1-p)^k \underset{k \to +\infty}{=} o\left(\binom{n+k-1}{k} q^{k-1}\right),\end{aligned}$$

mais $\binom{n+k-1}{k} q^{n-1}$ est le terme général d'une série convergente d'après la question 5 (avec $r = n-1$ et $x = q$) puisque $q \in [0, 1[$, donc la série de terme général $(n+k)^m \, \mathbb{P}(X_n = n+k)$ converge par comparaison des séries à termes positifs. Ainsi, X_n^m admet une espérance d'après le théorème de transfert, c'est-à-dire que X_n admet un moment d'ordre m.

Comme m est quelconque dans $\mathbb{N}$, on en déduit que X_n admet des moments de tous ordres.

Une autre justification pourrait être donnée en remarquant comme dans la question 9(c) que X_n a la même loi qu'une somme[31] de n VARD indépendantes de loi $\mathcal{G}(p)$.

31. L'existence d'un moment d'ordre m est stable par somme : si X et Y sont deux VARD positives admettant un moment d'ordre $m \in \mathbb{N}$, alors c'est aussi le cas de $X + Y$ en vertu de l'inégalité $0 \leqslant (X+Y)^m \leqslant 2^m (X^m + Y^m)$ dont la démonstration est laissée en exercice.

10. (a) Pour tout $k \in \mathbb{N}$, on a

$$\frac{n-1}{n+k-1}\binom{n+k-1}{n-1} = \frac{n-1}{n+k-1} \cdot \frac{(n+k-1)!}{(n-1)!k!}$$
$$= \frac{(n+k-2)!}{(n-2)!k!} = \binom{n+k-2}{k}$$

ce qui est le résultat attendu.

(b) La série $\sum_{k \geqslant 0} \frac{n-1}{n+k-1}\mathbb{P}(X_n = n+k)$ peut se réécrire sous la forme

$$\sum_{k \geqslant 0} \frac{n-1}{n+k-1}\mathbb{P}(X_n = n+k) = \sum_{k \geqslant 0} \frac{n-1}{n+k-1}\binom{n+k-1}{n-1}p^n(1-p)^k$$

ou encore, grâce à la question précédente, sous la forme

$$\sum_{k \geqslant 0} \binom{n+k-2}{k} p^n (1-p)^k = p^n \sum_{k \geqslant 0} \binom{n+k-2}{k}(1-p)^k$$

Or le résultat de la question 5 (avec $r = n-2$ et $x = 1-p$) indique que cette série est convergente et de somme $\frac{p^n}{p^{n-1}}$. D'après le théorème de transfert, $\frac{n-1}{X_n - 1}$ admet donc une espérance égale à p.

11. (a) Au moment de l'obtention du n-ième succès, on a obtenu Y_n échecs et n succès ; ainsi, on a $X_n = Y_n + n$.

(b) Comme [32] $X_n(\Omega) = \{k \in \mathbb{N} : k \geqslant n\}$ et $Y_n = X_n - n$, on a $Y_n(\Omega) = \mathbb{N}$, et pour tout $k \in \mathbb{N}$,

$$\mathbb{P}(Y_n = k) = \mathbb{P}(X_n = n+k) = \binom{n+k-1}{n-1}p^n(1-p)^k$$

d'après la question 8(c).

12. On a vu dans la question 9(c) que X_n admet des moments de tous ordres ; c'est donc aussi le cas de Y_n puisque $0 \leqslant Y_n \leqslant X_n$. En écrivant $Y_n = X_n - n$, on obtient par ailleurs

$$\mathbb{E}(Y_n) = \mathbb{E}(X_n - n) = \mathbb{E}(X_n) - n = \frac{n}{p} - n = \frac{n(1-p)}{p}$$

et

$$\mathbb{V}(Y_n) = \mathbb{V}(X_n - n) = \mathbb{V}(X_n) = \frac{n(1-p)}{p^2}.$$

32. On a retiré la valeur 0 de l'expression du support $X_n(\Omega)$ après avoir remarqué dans la question 8(d) que $\mathbb{P}(X_n = 0) = 0$. Cette opération, qui pourra choquer les lecteurs les plus rigoureux, est rendue intellectuellement supportable par le fait que nous n'explicitons pas la VARD X_n en tant qu'application et nous intéressons uniquement à son interprétation probabiliste.

Correction de l'exercice 56.

1. (a) On démontre l'égalité attendue [33] par récurrence.

La proposition est vraie pour $n=0$ puisque pour tout $x \in]-1,1[$ on a

$$-\sum_{k=1}^{0} \frac{(-x)^k}{k} + \int_0^x \frac{(t-x)^0}{(1+t)^1}\,\mathrm{d}t = 0 + \int_0^x \frac{\mathrm{d}t}{1+t} = \ln(1+x).$$

Soit à présent $n \in \mathbb{N}$. Supposons la propriété vraie au rang n, c'est-à-dire que pour tout $x \in]-1,1[$ on a

$$\ln(1+x) = -\sum_{k=1}^{n} \frac{(-x)^k}{k} + \int_0^x \frac{(t-x)^n}{(1+t)^{n+1}}\,\mathrm{d}t.$$

Fixons $x \in]-1,1[$. On peut réécrire l'intégrale ci-dessus à l'aide d'une intégration par parties, en dérivant la fonction $t \mapsto \frac{1}{(1+t)^{n+1}}$ de classe $\mathcal{C}^1$ et en primitivant la fonction continue $t \mapsto (t-x)^n$, pour obtenir

$$\begin{aligned}\int_0^x \frac{(t-x)^n}{(1+t)^{n+1}}\,\mathrm{d}t &= \left[\frac{(t-x)^{n+1}}{(n+1)(1+t)^{n+1}}\right]_0^x + \int_0^x \frac{(t-x)^{n+1}}{(1+t)^{n+2}}\,\mathrm{d}t \\ &= \frac{-(-x)^{n+1}}{n+1} + \int_0^x \frac{(t-x)^{n+1}}{(1+t)^{n+2}}\,\mathrm{d}t\end{aligned}$$

d'où, d'après l'hypothèse de récurrence :

$$\begin{aligned}\ln(1+x) &= -\sum_{k=1}^{n} \frac{(-x)^k}{k} + \frac{-(-x)^{n+1}}{n+1} + \int_0^x \frac{(t-x)^{n+1}}{(1+t)^{n+2}}\,\mathrm{d}t \\ &= -\sum_{k=1}^{n+1} \frac{(-x)^k}{k} + \int_0^x \frac{(t-x)^{n+1}}{(1+t)^{n+2}}\,\mathrm{d}t,\end{aligned}$$

ce qui est la proposition à démontrer au rang $n+1$.

Ainsi, la proposition est vraie pour tout $n \in \mathbb{N}$ d'après le principe de récurrence.

(b) En appliquant le résultat de la question précédente à $x = -p \in]-1,1[$, on obtient :

$$\forall n \in \mathbb{N}, \quad -\ln(1-p) = \sum_{k=1}^{n} \frac{p^k}{k} - \int_0^{-p} \frac{(t+p)^n}{(1+t)^{n+1}}\,\mathrm{d}t. \tag{1}$$

Or on a

$$\begin{aligned}\left|\int_0^{-p} \frac{(t+p)^n}{(1+t)^{n+1}}\,\mathrm{d}t\right| &\leqslant \int_{-p}^{0} \frac{|t+p|^n}{|1+t|^{n+1}}\,\mathrm{d}t \quad \text{d'après l'inégalité triangulaire} \\ &\leqslant \int_{-p}^{0} \frac{1}{1+t}\left(\frac{t+p}{1+t}\right)^n \mathrm{d}t \quad \text{car } 1+t \geqslant 0 \text{ et } t+p \geqslant 0.\end{aligned}$$

33. On a ici encore affaire à la *formule de Taylor avec reste intégral*, cette fois pour la fonction $x \mapsto \ln(1+x)$ (voir l'exercice **??** du chapitre **??**).

De plus, sur l'intervalle $[-p, 0]$, la fonction $t \mapsto \frac{t+p}{1+t}$ est positive, et elle est croissante[34], donc majorée sur cet intervalle par sa valeur en 0, c'est-à-dire par p. Ainsi, pour tout $t \in [-p, 0]$ on a

$$0 \leqslant \frac{1}{1+t} \cdot \left(\frac{t+p}{1+t}\right)^n = \frac{1}{1+t} \cdot p^n \leqslant \frac{p^n}{1-p},$$

d'où

$$\begin{aligned}\left|\int_0^{-p} \frac{(t+p)^n}{(1+t)^{n+1}} \,\mathrm{d}t\right| &\leqslant \int_{-p}^{0} \frac{1}{1+t} \cdot \left(\frac{t+p}{1+t}\right)^n \mathrm{d}t \\ &\leqslant \int_{-p}^{0} \frac{p^n}{1-p} \mathrm{d}t = \frac{p^{n+1}}{1-p} \xrightarrow[n\to+\infty]{} 0\end{aligned}$$

puisque $|p| < 1$, d'où $\displaystyle\int_0^{-p} \frac{(t+p)^n}{(1+t)^{n+1}} \,\mathrm{d}t \xrightarrow[n\to+\infty]{} 0$. La formule (1) montre alors que la série de terme général $\frac{p^k}{k}$ converge et que l'on a

$$\sum_{k=1}^{+\infty} \frac{p^k}{k} = -\ln(1-p),$$

ce qu'il fallait démontrer.

2. (a) Les valeurs proposées pour $\mathbb{P}(X = k)$ sont bien positives puisque $\ln(1-p) < 0$, et leur somme est égale à 1 d'après la question 1(b).

   (b) La série de terme général $k\,\mathbb{P}(X = k) = \frac{-1}{\ln(1-p)} p^k$ converge en tant que série géométrique de paramètre $p \in]-1, 1[$, donc X admet une espérance égale à

$$\mathbb{E}(X) = \sum_{k=1}^{+\infty} k\,\mathbb{P}(X = k) = \frac{-1}{\ln(1-p)} \sum_{k=1}^{+\infty} p^k = \frac{-1}{(1-p)\ln(1-p)}.$$

   (c) Pour tout $n \geqslant 0$, on a

$$k^n\,\mathbb{P}(X = k) = \frac{-1}{\ln(1-p)} k^{n-1} p^k \underset{n\to+\infty}{=} o\left(\frac{1}{n^2}\right)$$

   par croissance comparée, or la série de terme général $\frac{1}{n^2}$ converge en tant que série de Riemann de paramètre $2 > 1$, donc la série de terme général $k^n\,\mathbb{P}(X = k)$ converge elle aussi par comparaison des séries à termes positifs ; ainsi, X^n admet une espérance, c'est-à-dire que X^n admet un moment d'ordre n. On en déduit que X possède des moments de tous ordres.

   (d) Soit $s \in [0, 1]$. La VARD s^X est à valeurs dans $[0, 1]$, donc elle admet une espérance, et on a d'après le théorème de transfert

$$\begin{aligned}\mathbb{E}(s^X) &= \sum_{k=1}^{+\infty} s^k\,\mathbb{P}(X = k) \\ &= \frac{-1}{\ln(1-p)} \sum_{k=1}^{+\infty} \frac{(ps)^k}{k} = \frac{\ln(1-ps)}{\ln(1-p)},\end{aligned}$$

34. On le démontre en écrivant $\frac{t+p}{1+t} = 1 - \frac{1-p}{1+t}$ ou par dérivation.

où la dernière égalité provient de la question 1(b) appliquée à $ps \in]0,1[$ si $s>0$ et est triviale si $s=0$.

3. Pour tout $\omega \in \Omega$, le réel $Y(\omega)$ est une somme de $N(\omega) \geqslant 0$ entiers strictement positifs (les $X_k(\omega)$) ; c'est donc un entier positif, donc $Y(\Omega) \subset \mathbb{N}$. On se convainc par ailleurs du fait que tout $n \in \mathbb{N}$ est une valeur possible pour Y en considérant une situation dans laquelle N prend la valeur n et les VARD $X_1, \ldots, X_n$ prennent toutes la valeur 1. Ainsi, $Y(\Omega) = \mathbb{N}$.

   On va à présent utiliser l'indication et calculer la quantité $\mathbb{E}(s^Y)$ pour tout $s \in [0,1]$ (l'existence d'une telle espérance étant assurée par le fait que s^Y est à valeurs dans $[0,1]$). Soit $s \in [0,1]$. D'après le théorème de transfert et la formule des probabilités totales appliquée au système complet d'événements de probabilités non nulles $((N=n))_{n\in\mathbb{N}}$, on a :

$$\begin{aligned}\mathbb{E}(s^Y) &= \sum_{k=0}^{+\infty} s^k\,\mathbb{P}(Y=k)\\ &= \sum_{k=0}^{+\infty} s^k \sum_{n=0}^{+\infty} \mathbb{P}(N=n)\,\mathbb{P}(Y=k\,|\,N=n)\\ &= \sum_{k=0}^{+\infty} s^k \sum_{n=0}^{+\infty} \mathbb{P}(N=n)\,\mathbb{P}\left(\sum_{i=1}^{n} X_i = k\right).\end{aligned}$$

   On peut alors appliquer le théorème de Fubini puisque les termes sommés sont tous positifs, ce qui donne [35] :

$$\begin{aligned}\mathbb{E}(s^Y) &= \sum_{n=0}^{+\infty} \mathbb{P}(N=n) \sum_{k=0}^{+\infty} s^k\,\mathbb{P}\left(\sum_{i=1}^{n} X_i = k\right)\\ &= \sum_{n=0}^{+\infty} \mathbb{P}(N=n)\,\mathbb{E}(s^{X_1+\cdots+X_n})\end{aligned}$$

   où la dernière égalité provient du théorème de transfert. Or les X_i sont indépendantes, donc les s^{X_i} le sont aussi d'après le lemme des coalitions ; ainsi, pour tout $n \in \mathbb{N}$ on a

$$\mathbb{E}(s^{X_1+\cdots+X_n}) = \mathbb{E}(s^{X_1}\cdots s^{X_n}) = \mathbb{E}(s^{X_1})\cdots\mathbb{E}(s^{X_n}),$$

   d'où

$$\mathbb{E}(s^{X_1+\cdots+X_n}) = \left(\frac{\ln(1-ps)}{\ln(1-p)}\right)^n$$

   d'après la question précédente puisque toutes les X_k suivent la même loi logarithmique de paramètre p. Ainsi, on a

$$\mathbb{E}(s^Y) = \sum_{n=0}^{+\infty} e^{-\lambda}\frac{\lambda^n}{n!}\left(\frac{\ln(1-ps)}{\ln(1-p)}\right)^n = e^{-\lambda}e^{\lambda\frac{\ln(1-ps)}{\ln(1-p)}}.$$

35. Il s'agit d'une nouvelle occurrence de la *formule de l'espérance totale* évoquée dans la correction de l'exercice 47 : l'espérance de s^Y est égale à la valeur moyenne (sur toutes les valeurs possibles de N) de l'espérance de s^Y *conditionnellement à N*.

Considérons à présent une VARD B suivant une loi binomiale négative de paramètres $m := \frac{-\lambda}{\ln(1-p)}$ et $1-p$. D'après la question 11(b) de l'exercice 55, on a donc $B(\Omega) = \mathbb{N}$ et :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(B = k) = \binom{m+k-1}{k}(1-p)^m p^k.$$

D'après le théorème de transfert, on a donc

$$\mathbb{E}(s^B) = \sum_{k=0}^{+\infty} s^k \,\mathbb{P}(B = k) = (1-p)^m \sum_{k=0}^{+\infty} \binom{m+k-1}{k} p^k s^k$$

d'où, en utilisant le résultat de la question 5 de l'exercice 55 (ce qui est licite puisque $ps \in [0, 1[$) :

$$\mathbb{E}(s^B) = \frac{(1-p)^m}{(1-s(1-p))^m} = \left(\frac{1-p}{1-ps}\right)^m.$$

Il nous reste à comparer les quantités $\mathbb{E}(s^Y)$ et $\mathbb{E}(s^B)$. Comme $\lambda = -m\ln(1-p)$, on peut écrire

$$\mathbb{E}(s^Y) = e^{-\lambda} e^{\lambda \frac{\ln(1-ps)}{\ln(1-p)}} = e^{m\ln(1-p)} e^{-m\ln(1-ps)} = \left(\frac{1-p}{1-ps}\right)^m = \mathbb{E}(s^B).$$

Ainsi, $\mathbb{E}(s^Y) = \mathbb{E}(s^B)$ quel que soit $s \in [0, 1]$; d'après le résultat admis, les VARD à valeurs entières Y et B ont donc même loi, ce qui clôt la preuve.

Correction de l'exercice 57.

1. En adaptant la preuve donnée dans l'exercice 35 du chapitre « Matrices et systèmes linéaires I », on obtient l'inversibilité de ${}^tA$ et donc de A. Une preuve plus rapide peut être trouvée dans l'exercice XXX du chapitre « Matrices et systèmes linéaires II » ; on la reproduit ci-dessous dans le cas de la matrice étudiée.

   On verra dans le chapitre **??** qu'il suffit pour établir que ${}^tA$ est inversible de montrer que l'équation ${}^tAX = 0$ d'inconnue $X \in \mathcal{M}_{n+1,1}(\mathbb{R})$ admet $X = 0$ pour unique solution (c'est-à-dire, pour anticiper sur la terminologie introduite dans ce chapitre, que *le noyau de* ${}^tA$ *est réduit à* $\{0\}$). Soit donc un vecteur $X = {}^t(\lambda_0, \ldots, \lambda_n) \in \mathcal{M}_{n+1,1}$ tel que ${}^tAX = 0$. On a alors

$${}^tA \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad i.e. \quad \begin{cases} \displaystyle\sum_{k=0}^{n} \lambda_k 0^k = 0 \\ \displaystyle\sum_{k=0}^{n} \lambda_k 1^k = 0 \\ \qquad \vdots \\ \displaystyle\sum_{k=0}^{n} \lambda_k n^k = 0, \end{cases}$$

c'est-à-dire que le polynôme $P(X) = \sum_{k=0}^{n} \lambda_k X^k$ admet $0, 1, \ldots, n$ pour racines.

Or P est de degré au plus n, donc le fait qu'il admette au moins $n+1$ racines distinctes permet de conclure que $P = 0$, c'est-à-dire que ses coefficients λ_k sont tous nuls. On en déduit que $X = 0$.

Comme il est clair que réciproquement, si $X = 0$ alors ${}^tAX = 0$, on a bien ${}^tAX = 0$ si et seulement si $X = 0$. Ainsi, ${}^tA$ est inversible, donc A l'est aussi.

2. Pour tout $k \in [\![0, n]\!]$, la $(k+1)$-ième coordonnée du vecteur AU est

$$\sum_{j=1}^{n+1} A_{k+1,j} U_j = \sum_{j=0}^{n} j^k \mathbb{P}(X = j) = \mathbb{E}(X^k)$$

d'après le théorème de transfert, et la $(k+1)$-ième coordonnée du vecteur AV est $\mathbb{E}(Y^k)$ par un calcul similaire. Or on a $\mathbb{E}(X^k) = \mathbb{E}(Y^k)$ pour tout $k \in [\![1, n]\!]$ par hypothèse, et il est évident que cette relation est aussi vraie pour $k = 0$, donc $AU = AV$. Comme A est inversible, on en déduit que $U = V$, c'est-à-dire que pour tout $i \in [\![0, n]\!]$ on a $\mathbb{P}(X = i) = \mathbb{P}(Y = i)$: ainsi, X et Y ont même loi, ce que nous voulions démontrer.

L'égalité des moments de X et Y implique donc l'égalité de leurs lois : on dit que *les moments d'une variable aléatoire prenant un nombre fini de valeurs réelles caractérisent sa loi.*

Correction de l'exercice 58.

1. Supposons qu'il existe une fonction g vérifiant la condition attendue. Soit $x \in X(\Omega)$; en appliquant l'hypothèse à la fonction $f_x = \mathbb{1}_{\{x\}}$ et en remarquant que $g(X)\mathbb{1}_{\{x\}}(X) = g(x)\mathbb{1}_{\{x\}}(X)$, on obtient

$$\mathbb{E}(Y\mathbb{1}_{\{x\}}(X)) = \mathbb{E}(g(X)\mathbb{1}_{\{x\}}(X)) = \mathbb{E}(g(x)\mathbb{1}_{\{x\}}(X)) = g(x)\,\mathbb{P}(X = x).$$

Ainsi, pour tout $x \in \mathbb{R}$ tel que $\mathbb{P}(X = x) > 0$ on a

$$g(x) = \frac{\mathbb{E}(Y\mathbb{1}_{\{x\}}(X))}{\mathbb{P}(X = x)}.$$

Réciproquement, soit g la fonction définie sur $X(\Omega)$ de la façon suivante :

$$\forall x \in X(\Omega), \quad \begin{cases} \dfrac{\mathbb{E}(Y\mathbb{1}_{\{x\}}(X))}{\mathbb{P}(X = x)} & \text{si } \mathbb{P}(X = x) > 0 \\ 0 & \text{sinon.} \end{cases}$$

Remarquons que cette définition est licite car pour tout $x \in \mathbb{R}$ on a $|Y\mathbb{1}_{\{x\}}(X)| \leqslant |Y|$, ce qui implique que la VARD $Y\mathbb{1}_{\{x\}}(X)$ admet bien une espérance puisque c'est aussi le cas de $|Y|$.

Montrons que g vérifie la condition attendue. Soit une fonction $f : X(\Omega) \to \mathbb{R}$. Alors $g(X)f(X)$ admet une espérance puisqu'elle prend un nombre fini de valeurs, et on a d'après le théorème de transfert et la définition de g :

$$\mathbb{E}(g(X)f(X)) = \sum_{x \in X(\Omega)} g(x)f(x)\,\mathbb{P}(X = x) = \sum_{x \in X(\Omega)} f(x)\,\mathbb{E}(Y\mathbb{1}_{\{x\}}(X)).$$

Par linéarité de l'espérance (rappelons que $X(\Omega)$ est un ensemble fini !), on a alors
$$\mathbb{E}(g(X)f(X)) = \mathbb{E}\left(\sum_{x\in X(\Omega)} f(x)Y\mathbb{1}_{\{x\}}(X)\right) = \mathbb{E}\left(Yf(X)\right)$$
puisque $f(X) = \sum_{x\in X(\Omega)} f(x)\mathbb{1}_{\{x\}}(X)$. Ainsi, g vérifie bien la condition attendue.

2. Supposons que h soit une fonction vérifiant la même propriété que g. On a vu dans la question précédente (plus précisément dans la phase d'analyse) que l'on a nécessairement
$$h(x) = \frac{\mathbb{E}(Y\mathbb{1}_{\{x\}}(X))}{\mathbb{P}(X = x)} = g(x)$$
pour tout $x \in \mathbb{R}$ tel que $\mathbb{P}(X = x) > 0$. Ainsi,
$$\begin{aligned}
\mathbb{P}(h(X) = g(X)) &= \mathbb{P}\left(\bigsqcup_{x\in X(\Omega)} (X = x) \cap (h(X) = g(X))\right) \\
&= \mathbb{P}\left(\bigsqcup_{x\in X(\Omega)} \Big((X = x) \cap (h(x) = g(x))\Big)\right) \\
&= \sum_{x\in X(\Omega)} \mathbb{P}\Big((X = x) \cap (h(x) = g(x))\Big) \quad \text{par incompatibilité} \\
&= \sum_{\substack{x\in X(\Omega) \\ h(x)=g(x)}} \mathbb{P}(X = x) \\
&\geqslant \sum_{\substack{x\in X(\Omega) \\ \mathbb{P}(X=x)>0}} \mathbb{P}(X = x) = \sum_{x\in X(\Omega)} \mathbb{P}(X = x) = 1,
\end{aligned}$$
d'où $\mathbb{P}(h(X) = g(X)) = 1$, ce qu'il fallait démontrer.

3. Soit $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) > 0$. On a d'après le théorème de transfert pour la loi jointe[36] de (X, Y) :
$$\begin{aligned}
g(x) = \frac{\mathbb{E}(Y\mathbb{1}_{\{x\}}(X))}{\mathbb{P}(X = x)} &= \sum_{\substack{x'\in X(\Omega) \\ y\in Y(\Omega)}} y\mathbb{1}_{\{x\}}(x') \frac{\mathbb{P}((X = x') \cap (Y = y))}{\mathbb{P}(X = x)} \\
&= \sum_{y\in Y(\Omega)} y \frac{\mathbb{P}((X = x) \cap (Y = y))}{\mathbb{P}(X = x)} \\
&= \sum_{y\in Y(\Omega)} y\,\mathbb{P}(Y = y \,|\, X = x).
\end{aligned}$$
Ainsi, $g(x)$ est la valeur de l'espérance de Y calculée non pas par rapport à la probabilité $\mathbb{P}$ mais par rapport à la mesure de probabilité conditionnelle

36. Ce théorème, on le rappelle, est disponible en suivant ce lien.

$\mathbb{P}(\cdot \mid X = x)$, c'est-à-dire l'espérance de la loi suivie par la variable aléatoire Y conditionnellement à l'événement $(X = x)$.

On peut montrer que l'existence de g est assurée dès lors que X admet une espérance, même si $X(\Omega)$ est infini, et il est coutume de noter $g(X)$ sous la forme $\mathbb{E}(Y \mid X)$, notation que l'on prononce[37] « espérance de Y sachant X ».

Correction de l'exercice 59.

1. Comme X est à valeurs dans I et comme I est un intervalle, l'espérance $\mathbb{E}(X)$, qui est une moyenne pondérée des valeurs prises par X (et donc d'éléments de I), est elle aussi dans I.

   Démontrons cette propriété plus rigoureusement. On peut écrire I sous l'une des formes $[a, b]$, $[a, b[$, $]a, b[$ ou $]a, b]$, avec a et b éventuellement infinis. Établissons la propriété dans le cas où a et b sont finis et où $I = [a, b[$ (les autres cas étant similaires). On a alors $a \leqslant x < b$ pour tout $x \in X(\Omega)$, donc $a\,\mathbb{P}(X = x) \leqslant x\,\mathbb{P}(X = x) < b\,\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) > 0$, et donc, en sommant[38] cette double inégalité :

$$\sum_{x \in X(\Omega)} a\,\mathbb{P}(X = x) \leqslant \sum_{x \in X(\Omega)} x\,\mathbb{P}(X = x) < \sum_{x \in X(\Omega)} b\,\mathbb{P}(X = x),$$

   soit $a \leqslant \mathbb{E}(X) < b$, et donc $\mathbb{E}(X) \in I$.

2. Comme X est à valeurs dans I et $\mathbb{E}(X) \in I$, pour tout $\omega \in \Omega$ on a, en appliquant l'inégalité de convexité vérifiée par f aux points $x = \mathbb{E}(X)$ et $y = X(\omega)$:

$$f(X(\omega)) \geqslant f(\mathbb{E}(X)) + (X(\omega) - \mathbb{E}(X))f'(\mathbb{E}(X)).$$

   Cette relation étant vérifiée pour tout $\omega \in \Omega$, on a alors[39] :

$$f(X) \geqslant f(\mathbb{E}(X)) + (X - \mathbb{E}(X))f'(\mathbb{E}(X)).$$

   En passant à l'espérance dans l'inégalité ci-dessus (c'est-à-dire en utilisant la croissance de l'espérance) on obtient, par linéarité de l'espérance :

$$\begin{aligned}\mathbb{E}(f(X)) &\geqslant \mathbb{E}\Big(f(\mathbb{E}(X)) + (X - \mathbb{E}(X))f'(\mathbb{E}(X))\Big)\\ &= f(\mathbb{E}(X)) + \mathbb{E}(X - \mathbb{E}(X))f'(\mathbb{E}(X)).\end{aligned}$$

   Mais $\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0$, à nouveau par linéarité de l'espérance, donc on a $\mathbb{E}(f(X)) \geqslant f(\mathbb{E}(X))$, ce qu'il fallait démontrer.

---

37. Avec cet outil conceptuel, on peut par exemple écrire que dans le cadre de l'exercice 40, comme le nombre d'œufs éclos Y suit une loi $\mathcal{B}(n, p)$ conditionnellement à l'événement $(X = x)$, on a $\mathbb{E}(Y \mid X) = nX$. Dans le cadre de la question 3 de l'exercice 56, on a cette fois $\mathbb{E}(Y \mid N) = \lambda N$.

38. On aurait pu vouloir invoquer la croissance de l'espérance pour passer directement de l'encadrement $a \leqslant X < b$ à $a \leqslant \mathbb{E}(X) < b$, mais il aurait fallu utiliser une propriété de *stricte* croissance qui stipule que si Y est une VARD admettant une espérance et si $X < Y$ alors $\mathbb{E}(X) < \mathbb{E}(Y)$; cette propriété est vraie (on la démontre en considérant la VARD strictement positive $Y - X$) mais elle ne figure pas dans le cours, d'où la démonstration plus « pédestre » donnée ici.

39. Rappelons que $f(X)$ est une notation commode utilisée pour l'application $f \circ X$ lorsque X est une variable aléatoire.

3. Supposons que f est concave et que X et $f(X)$ admettent une espérance. Alors la fonction $-f$ est convexe, et $-f(X)$ admet une espérance, donc en appliquant l'inégalité que nous venons de démontrer à $-f$ on obtient que $\mathbb{E}(-f(X)) \geqslant -f(\mathbb{E}(X))$. En utilisant la linéarité de l'espérance et en passant à l'opposé dans les deux termes de l'inégalité, on obtient alors la relation $\mathbb{E}(f(X)) \leqslant f(\mathbb{E}(X))$, qui est la version concave de l'inégalité de Jensen.
4. La quantité $\mathbb{E}(u(X))$ représente l'*utilité espérée* du jeu, c'est-à-dire le niveau moyen de satisfaction [40] que le joueur retire de sa participation au jeu. La quantité $u(\mathbb{E}(X))$, quant à elle, représente l'utilité associée à un gain certain égal à $\mathbb{E}(X)$, c'est-à-dire le niveau de satisfaction du joueur si on lui proposait, sans incertitude, d'empocher le gain moyen du jeu.

   Si u est convexe, l'inégalité de Jensen stipule que $\mathbb{E}(u(X)) \geqslant u(\mathbb{E}(X))$, c'est-à-dire que le joueur préfère participer au jeu (avec incertitude) plutôt que de toucher un gain certain d'un même montant moyen : il a donc une propension à préférer le risque. Si u est concave, c'est l'inverse : le joueur a une aversion au risque.

   Ainsi, le caractère concave ou convexe de la fonction d'utilité permet de modéliser les préférences du joueur en termes de prise de risque. On peut bien sûr envisager des cas « mixtes » plus subtils, par exemple celui d'une fonction d'utilité convexe en-deçà d'un certain point et concave ensuite, pour modéliser le fait que le joueur est attiré par le risque lorsqu'il joue des montants faibles mais averse [41] au risque lorsque les sommes en jeu deviennent trop importantes.

Correction de l'exercice 60.

1. La particule revient en 0 au cours de sa trajectoire avec probabilité α. Si elle atteint 0, elle suit à partir de ce moment une trajectoire indépendante de ses mouvements précédents et selon la même loi de probabilité que sa trajectoire initiale [42] ; cette nouvelle trajectoire revient donc en 0 avec probabilité α. Si ce nouveau retour en 0 a lieu, la particule suit à partir de ce moment une nouvelle trajectoire indépendante du passé qui revient en 0 avec probabilité α, et ainsi de suite.

   Ainsi, si $\alpha = 1$, la particule revient en 0 avec probabilité $\alpha = 1$, puis une nouvelle fois avec probabilité $\alpha = 1$ indépendamment du passé, puis à nouveau avec probabilité $\alpha = 1$, et ainsi de suite : elle revient donc en 0 une infinité de fois avec probabilité 1. Cette dernière déduction, qui répond à la demande de justification informelle de l'énoncé, repose sur l'idée selon laquelle l'intersection d'une famille dénombrable d'événements presque certains est un événement presque certain [43].

40. Le postulat selon lequel un agent rationnel participant à un jeu cherche à maximiser la quantité $\mathbb{E}(u(X))$ est au fondement de la *théorie de l'utilité espérée*, une théorie de la décision en environnement risqué développée par John von Neumann et Oskar Morgenstern dans leur ouvrage *Theory of Games and Economic Behavior* (1944). Elle a notamment été discutée lors de la décennie suivante par Maurice Allais (voir l'article Wikipédia sur le paradoxe d'Allais).

41. Dire qu'un agent est « averse au risque » est un anglicisme courant en économie.

42. Ce point, que nous ne serions pas en mesure de démontrer rigoureusement sans introduire un lourd formalisme hors-programme, est appelé *propriété de Markov forte* de la marche aléatoire.

43. Elle peut être formalisée en introduisant pour tout $n \in \mathbb{N}^*$ l'événement A_n : « la particule

Supposons à présent que $\alpha < 1$. On va interpréter les retours successifs en 0 de la particule comme des *échecs* dans un schéma de Bernoulli infini. En effet, considérons que la particule entreprend au début de la marche une expérience de Bernoulli qui est un succès si et seulement si elle ne repasse pas par 0 au cours de sa trajectoire, ce qui advient avec probabilité $1-\alpha$. Comme on l'a déjà évoqué, si l'expérience est un échec, c'est-à-dire si la particule atteint 0 au cours de sa trajectoire, une nouvelle expérience est relancée indépendamment du passé, avec une probabilité de succès $1-\alpha$, et ainsi de suite jusqu'à obtention d'un premier succès. Ainsi, si l'on note N le nombre de retours en 0 de la particule au cours de la trajectoire (c'est-à-dire le nombre d'échecs rencontrés lors du schéma de Bernoulli), alors $N+1$ représente le rang d'obtention du premier succès dans le schéma ; on a donc $N+1 \sim \mathcal{G}(1-\alpha)$, ce qu'il fallait démontrer.

2. Soit $n \in \mathbb{N}^*$. L'événement $(X_n = 0)$ est réalisé si et seulement si la particule a effectué autant de sauts vers la droite que vers la gauche au cours des n premiers sauts. C'est évidemment impossible si n est impair, et si n est pair, cela revient à dire qu'exactement $\frac{n}{2}$ sauts parmi n sont des sauts vers la droite, ce qui se produit avec probabilité

$$\binom{n}{\frac{n}{2}}\left(\frac{1}{2}\right)^{\frac{n}{2}}\left(\frac{1}{2}\right)^{\frac{n}{2}} = \binom{n}{\frac{n}{2}}\frac{1}{2^n}$$

puisque les sauts sont indépendants les uns des autres.

Ainsi, pour tout $k \in \mathbb{N}^*$ on a

$$\mathbb{P}(X_{2k-1} = 0) = 0 \quad \text{et} \quad \mathbb{P}(X_{2k} = 0) = \binom{2k}{k}\frac{1}{2^{2k}}.$$

On a donc

$$\sum_{n=1}^{+\infty} \mathbb{P}(X_n = 0) = \sum_{k=1}^{+\infty} \binom{2k}{k}\frac{1}{2^{2k}}.$$

Or la formule de Stirling (voir la question 5 de l'exercice 51) permet d'écrire

$$\binom{2k}{k}\frac{1}{2^{2k}} \underset{k\to+\infty}{\sim} \frac{1}{\sqrt{\pi k}},$$

---

revient au moins n fois en 0 lors de sa trajectoire ». Notre discussion préliminaire montre alors que pour tout $k \in \mathbb{N}^*$ on a $\mathbb{P}(A_{k+1} \mid A_k) = 1$. On en déduit grâce à la formule des probabilités composées que pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1) \cdots \mathbb{P}(A_n \mid A_{n-1}) = 1.$$

La probabilité que la particule revienne en 0 une infinité de fois est donc

$$\mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n) = 1$$

d'après le théorème de la limite monotone puisque $(A_n)_{n\in\mathbb{N}^*}$ est une suite décroissante d'événements, ce qui permet de conclure.

mais la série de terme général $\frac{1}{\sqrt{\pi k}}$ diverge en tant que série de Riemann de paramètre $\frac{1}{2} \leqslant 1$, donc la série de terme général $\binom{2k}{k}\frac{1}{2^{2k}}$ diverge par comparaison des séries à termes positifs. Ainsi, la série de terme général $\mathbb{P}(X_n = 0)$ diverge aussi.

On veut à présent montrer que $\alpha = 1$. Raisonnons par l'absurde et supposons que $\alpha < 1$. On a alors $N + 1 \sim \mathcal{G}(1 - \alpha)$ d'après la question précédente, donc $\mathbb{E}(N) = \frac{1}{1-\alpha} - 1 = \frac{\alpha}{1-\alpha}$. Pour aboutir à une contradiction, on pourrait utiliser le calcul suivant :

$$\mathbb{E}(N) = \mathbb{E}\left(\sum_{n=1}^{+\infty} \mathbb{1}_{(X_n=0)}\right) = \sum_{n=1}^{+\infty} \mathbb{E}(\mathbb{1}_{(X_n=0)}) = \sum_{n=1}^{+\infty} \mathbb{P}(X_n = 0) = +\infty,$$

qui entre bien en contradiction avec le fait que l'espérance de N est finie, mais l'interversion des symboles $\mathbb{E}$ et $\sum$ dans la deuxième égalité ci-dessus n'est justifiable que par un résultat hors-programme (le théorème de Fubini pour l'espérance et les sommes de séries, qui spécifie qu'une telle opération est licite lorsque les VARD étudiées sont positives). On va donc recourir à un argument qui fait appel à une somme finie. Pour ce faire, introduisons un entier $A \in \mathbb{N}^*$ et notons N_A la variable aléatoire égale au nombre de retours en 0 de la particule *avant l'étape* A (incluse). On a donc $0 \leqslant N_A \leqslant N$, ce qui implique que N_A admet une espérance et que $0 \leqslant \mathbb{E}(N_A) \leqslant \mathbb{E}(N) = \frac{\alpha}{1-\alpha}$. Or on a aussi

$$N_A = \sum_{n=1}^{A} \mathbb{1}_{(X_n=0)},$$

donc par linéarité de l'espérance,

$$\mathbb{E}(N_A) = \mathbb{E}\left(\sum_{n=1}^{A} \mathbb{1}_{(X_n=0)}\right) = \sum_{n=1}^{A} \mathbb{E}(\mathbb{1}_{(X_n=0)}) = \sum_{n=1}^{A} \mathbb{P}(X_n = 0),$$

ce qui montre que

$$\sum_{n=1}^{A} \mathbb{P}(X_n = 0) \leqslant \frac{\alpha}{1 - \alpha}.$$

Cette relation étant vérifiée quel que soit $A \in \mathbb{N}^*$, on en déduit que la série de terme général $\mathbb{P}(X_n = 0)$ est majorée par $\frac{\alpha}{1-\alpha}$; or cela contredit le fait que la série diverge vers $+\infty$.

Ainsi, on ne peut avoir $\alpha < 1$: on a donc $\alpha = 1$.

3. Soit $n \in \mathbb{Z}$. Montrons tout d'abord que la particule atteint n presque sûrement. On a vu que la trajectoire de la particule consiste presque sûrement en une infinité d'excursions hors de 0 indépendantes les unes des autres et de même loi. Or la probabilité que la particule atteigne n avant son premier retour en 0 est strictement positive : en effet, elle est minorée par la probabilité que la particule rejoigne directement n en $|n|$ sauts puis retourne en 0 en $|n|$ nouveaux sauts, qui vaut $\frac{1}{2^{2|n|}}$. Notons $p > 0$ cette probabilité. Les excursions hors de 0 étant indépendantes et de même loi, pour tout $k \in \mathbb{N}^*$ la particule a une probabilité $1 - (1 - p)^k$ d'avoir atteint n au moins une fois lors des k premières

excursions ; d'après le théorème de la limite monotone [44], la probabilité qu'elle atteigne n au moins une fois au cours de sa trajectoire dans son ensemble est égale à $\lim\limits_{k\to+\infty}(1-(1-p)^k)=1$, ce qui signifie que la particule atteint n presque sûrement [45].

Justifions à présent que la particule atteint presque sûrement n une infinité de fois. Après que la particule a atteint n pour la première fois, sa trajectoire est une marche aléatoire simple décalée au départ de n ; or cette marche repasse presque sûrement une infinité de fois par son point de départ comme on l'a vu dans les questions 1 et 2, donc la trajectoire passe presque sûrement une infinité de fois par n, ce qu'il fallait démontrer.

4. Plaçons-nous dans le cas $d=2$. Le même raisonnement que dans les questions 1 et 2 montre que si la série de terme général $\mathbb{P}(X_n=0)$ diverge, alors la particule revient presque sûrement une infinité de fois en 0 (notons que l'on parle à présent de l'élément neutre $0=(0,0)\in\mathbb{R}^2$). Or pour tout $k\in\mathbb{N}^*$, la particule revient en 0 en n étapes si et seulement si elle a effectué pendant ces n étapes autant de sauts vers la droite que vers la gauche (c'est-à-dire autant de sauts dirigés par e_1 que de sauts dirigés par $-e_1$) et autant de sauts vers le haut que vers le bas (c'est-à-dire autant de sauts dirigés par e_2 que de sauts dirigés par $-e_2$). Ainsi, on a $\mathbb{P}(X_n=0)=0$ si $n\in\mathbb{N}^*$ est impair. Supposons à présent que n s'écrit $2k$ avec $k\in\mathbb{N}^*$: on a alors

$$\mathbb{P}(X_{2k}=0)=\sum_{i=0}^{k}\binom{2k}{i}\binom{2k-i}{i}\binom{2k-2i}{k-i}\binom{k-i}{k-i}\left(\frac{1}{4}\right)^i\left(\frac{1}{4}\right)^i\left(\frac{1}{4}\right)^{k-i}\left(\frac{1}{4}\right)^{k-i},$$

où i représente le nombre de sauts effectués vers la droite et $k-i$ le nombre de sauts effectués vers le haut, donc, après simplification des coefficients binomiaux :

$$\mathbb{P}(X_{2k}=0)=\sum_{i=0}^{k}\frac{(2k)!}{i!^2(k-i)!^2}\left(\frac{1}{4}\right)^{2k}=\left(\frac{1}{4}\right)^{2k}\frac{(2k)!}{k!^2}\sum_{i=0}^{k}\binom{k}{k-i}\binom{k}{k-i}.$$

Or

$$\sum_{i=0}^{k}\binom{k}{k-i}\binom{k}{k-i}=\sum_{i=0}^{k}\binom{k}{k-i}\binom{k}{i}=\binom{2k}{k}$$

d'après la formule de Vandermonde, donc

$$\mathbb{P}(X_{2k}=0)=\left(\frac{1}{4}\right)^{2k}\frac{(2k)!}{k!^2}\binom{2k}{k}=\left(\binom{2k}{k}\left(\frac{1}{4}\right)^k\right)^2.$$

En utilisant l'équivalent obtenu dans la question 2, on obtient alors

$$\mathbb{P}(X_{2k}=0)=\left(\binom{2k}{k}\left(\frac{1}{4}\right)^k\right)^2\underset{k\to+\infty}{\sim}\left(\frac{1}{\sqrt{\pi k}}\right)^2=\frac{1}{\pi k}.$$

44. En effet, les événements A_k : « la particule atteint n au moins une fois lors des k premières excursions » forment une suite croissante.

45. Plus généralement, on montre par un raisonnement similaire que si $(E_i)_{i\in\mathbb{N}^*}$ est une suite d'événements indépendants de même probabilité strictement positive, alors presque sûrement l'un des E_i au moins est réalisé.

Or la série de terme général $\frac{1}{\pi k}$ diverge en tant que série harmonique, donc on a, par comparaison des séries à termes positifs :

$$\sum_{n=1}^{+\infty} \mathbb{P}(X_n = 0) = \sum_{k=1}^{+\infty} \mathbb{P}(X_{2k} = 0) = +\infty,$$

ce qui, comme annoncé plus haut, montre que la particule revient presque sûrement une infinité de fois en 0.

5. Plaçons-nous à présent dans le cas $d \geqslant 3$. On admet que la série des $\mathbb{P}(X_n = 0)$ converge. On va montrer que ce fait implique que le retour à l'origine de la particule ne se produit presque sûrement qu'un nombre fini de fois.

   On propose pour cela deux démonstrations différentes.

   La première consiste à remarquer qu'il s'agit simplement d'une occurrence du *lemme de Borel-Cantelli* démontré dans un exercice du chapitre « Probabilités élémentaires » du tome de première année, et qui s'énonce comme suit : si $(E_n)_{n\in\mathbb{N}}$ est une suite d'événements telle que la série des $\mathbb{P}(E_n)$ converge, alors la probabilité qu'une infinité d'événements E_n soit réalisée[46] est nulle. Le résultat recherché est alors le lemme de Borel-Cantelli appliqué aux événements $E_n := (X_n = 0)$.

   Une autre idée de preuve, plus autocontenue, serait d'écrire comme dans la question 2 que

$$\mathbb{E}(N) = \mathbb{E}\left(\sum_{n=1}^{+\infty} \mathbb{1}_{(X_n=0)}\right) = \sum_{n=1}^{+\infty} \mathbb{E}(\mathbb{1}_{(X_n=0)}) = \sum_{n=1}^{+\infty} \mathbb{P}(X_n = 0) < +\infty,$$

   ce qui montrerait que le nombre N de retours en 0 est d'espérance finie donc est presque sûrement fini. Cependant, l'interversion des symboles $\mathbb{E}$ et $\sum$ n'est toujours pas licite dans le cadre de notre programme, ce qui oblige à effectuer un détour de raisonnement (que le lecteur pourra sauter sans dommage).

   Supposons que la conclusion à laquelle on souhaite arriver soit fausse, et donc que l'événement E : « la particule effectue une infinité de retours en 0 » soit de probabilité p strictement positive.

   Considérons à présent un entier $r \in \mathbb{N}^*$ tel que

$$r > \frac{1}{p} \sum_{n=1}^{+\infty} \mathbb{P}(X_n = 0),$$

46. Rappelons l'idée de la preuve de ce lemme. Notons S l'événement « une infinité de E_n sont réalisés ». Fixons à présent $\varepsilon > 0$ et considérons n_0 tel que $\sum\limits_{k=n_0}^{+\infty} \mathbb{P}(A_k) \leqslant \varepsilon$. On montre alors que

$$\mathbb{P}(S) \leqslant \mathbb{P}\left(\bigcup_{k=n_0}^{+\infty} E_k\right) \leqslant \sum_{k=n_0}^{+\infty} \mathbb{P}(E_k) \leqslant \varepsilon,$$

d'où l'on déduit que $\mathbb{P}(S) = 0$ puisque $\varepsilon > 0$ est quelconque.

ce qui est rendu possible par la convergence de la série de terme général $\mathbb{P}(X_n = 0)$. Pour tout $n \in \mathbb{N}^*$, on considère l'événement

C_n : « la particule revient en 0 au moins r fois avant l'étape n (incluse) »,

soit, avec les notations de la réponse à la question 2 : $C_n = (N_n \geqslant r)$. La suite d'événements $(C_n)_{n\in\mathbb{N}^*}$ étant croissante, on a alors, grâce au théorème de la limite monotone :

$$\begin{aligned}\mathbb{P}(C_n) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\bigcup_{k\in\mathbb{N}} C_k\right) & \\ &= \mathbb{P}(\text{la particule revient au moins } r \text{ fois en 0 durant sa trajectoire}) \\ &\geqslant \mathbb{P}(\text{la particule revient une infinité de fois en 0 durant sa trajectoire}) \\ &= \mathbb{P}(E) = p.\end{aligned}$$

Ainsi,

$$r\mathbb{P}(C_n) \xrightarrow[n\to+\infty]{} r\mathbb{P}\left(\bigcup_{k\in\mathbb{N}} C_k\right) \geqslant rp > \sum_{k=1}^{+\infty} \mathbb{P}(X_k = 0),$$

ce qui implique que

$$r\mathbb{P}(C_n) > \sum_{k=1}^{+\infty} \mathbb{P}(X_k = 0) \tag{2}$$

à partir d'un certain rang. Or pour tout $n \in \mathbb{N}^*$ on a $N_n \geqslant r\mathbb{1}_{(N_n\geqslant r)}$ puisque N_n est une VARD positive[47], donc

$$\mathbb{E}(N_n) \geqslant \mathbb{E}(r\mathbb{1}_{(N_n\geqslant r)}) = r\mathbb{P}(N_n \geqslant r) = r\mathbb{P}(C_n),$$

et par ailleurs

$$\mathbb{E}(N_n) = \mathbb{E}\left(\sum_{k=1}^{n} \mathbb{1}_{(X_k=0)}\right) = \sum_{k=1}^{n} \mathbb{P}(X_k = 0)$$

par linéarité de l'espérance, donc

$$r\mathbb{P}(C_n) \leqslant \mathbb{E}(N_n) = \sum_{k=1}^{n} \mathbb{P}(X_k = 0) \leqslant \sum_{k=1}^{+\infty} \mathbb{P}(X_k = 0).$$

Or cela contredit la relation (2) censée être vérifiée à partir d'un certain rang. On a donc démontré par l'absurde que l'événement E n'est pas de probabilité strictement positive, c'est-à-dire qu'il est quasi-impossible, donc que la particule effectue presque sûrement un nombre fini de retours en 0.

En utilisant le résultat de la question 1 extrapolé sans difficulté à la dimension d, on a plus précisément que la probabilité de retour en 0 de la particule est un réel $\alpha < 1$ et que le nombre N de retours en 0 au cours de la trajectoire vérifie $N + 1 \sim \mathcal{G}(1 - \alpha)$.

47. En effet, pour tout $\omega \in \Omega$ on a d'une part $N_n(\omega) \geqslant r = r\mathbb{1}_{(N_n(\omega)\geqslant r)}(\omega)$ si $N_n(\omega) \geqslant r$, et d'autre part $N_n(\omega) \geqslant 0 = r\mathbb{1}_{(N_n(\omega)\geqslant r)}(\omega)$ si $N(\omega) < r$. On retrouvera cette relation au fondement de la démonstration de l'*inégalité de Markov* que nous étudierons dans le chapitre sur la convergence des variables aléatoires.

Si $d = 1$ ou $d = 2$, on peut montrer en utilisant la question 3 que toute partie non vide de $\mathbb{Z}$ (ou $\mathbb{Z}^2$) est visitée presque sûrement une infinité de fois par la particule. En revanche, si $d \geqslant 3$, on montre facilement que quel que soit le domaine borné $\mathcal{D} \subset \mathbb{Z}^d$ considéré, alors presque sûrement il existe un moment où la particule quitte $\mathcal{D}$ pour ne plus jamais y pénétrer.

La marche aléatoire symétrique simple étudiée dans cet exercice est souvent appelée marche de l'homme ivre *puisqu'elle rappelle la démarche titubante d'une personne en état d'ébriété. Le théorème de Pólya stipule qu'un homme ivre se déplaçant le long d'une ligne droite ou sur un damier rentrera presque sûrement chez lui au bout d'un certain temps, mais qu'il n'est pas certain d'y arriver s'il se déplace dans l'espace ou dans un hyperespace (c'est-à-dire dans un espace de dimension supérieure à 3).*

Correction de l'exercice 61.

1. Pour tout $i \in [\![1, n]\!]$, la VARD X_i prend la valeur 1 avec probabilité p et la valeur 0 avec probabilité $1 - p$: on a donc $X_i \sim \mathcal{B}(p)$.
2. Soit $x = (x_1, \dots, x_n) \in \{0, 1\}^n$. Notons

$$s(x) = \sum_{i=1}^{n} x_i$$

la somme des coordonnées de x. On remarque que pour tout $i \in [\![1, n]\!]$, on a $\mathbb{P}(X_i = x_i) = p^{x_i}(1 - p)^{1-x_i}$ puisque cette probabilité vaut p si $x_i = 1$ et $1 - p$ si $x_i = 0$. On a donc

$$\begin{aligned} L(x, p) &= \mathbb{P}((X_1 = x_1) \cap \dots \cap (X_n = x_n)) \\ &= \prod_{i=1}^{n} \mathbb{P}(X_i = x_i) \quad \text{par indépendance des } X_i \\ &= \prod_{i=1}^{n} p^{x_i}(1 - p)^{1-x_i} \\ &= p^{s(x)}(1 - p)^{n-s(x)}. \end{aligned}$$

La quantité $L(x, p)$ est la probabilité d'obtenir le vecteur d'observations x lorsque l'on collecte une occurrence du vecteur aléatoire $(X_1, \dots, X_n)$; il s'agit donc bien du degré de *vraisemblance* (au sens courant du terme) de ce vecteur d'observations.

L'idée de l'estimation par maximum de vraisemblance est de se fonder sur un vecteur d'observations donné et d'estimer le paramètre inconnu p par la valeur qui maximise la vraisemblance des observations réalisées. On considère donc toujours que x est fixé, et on cherche cette fois le paramètre $p \in [0, 1]$ qui maximise la quantité $L(x, p)$.

On distingue plusieurs cas :

- Si $s(x) = 0$, on a $L(x, p) = (1 - p)^n$ pour tout $p \in [0, 1]$, donc $L(x, p)$ est maximale pour $p = 0$: l'estimateur du maximum de vraisemblance pour p est 0. Ce résultat est assez intuitif : si aucune personne sondée ne s'est exprimée en faveur du changement de programme, il est raisonnable d'estimer par $p = 0$ la proportion de personnes favorables au changement dans la population française.

- Si $s(x) = n$, la situation est inverse : on a alors $L(x,p) = p^n$ pour tout $p \in [0,1]$, donc $L(x,p)$ est maximale pour $p = 1$: l'estimateur du maximum de vraisemblance pour p est alors 1.
- Supposons à présent que $s(x) \in [\![1, n-1]\!]$, c'est-à-dire que l'échantillon d'observations récolté contient au moins une personne favorable au changement de programme et une personne défavorable à ce changement. La fonction $f : p \mapsto L(x,p) = p^{s(x)}(1-p)^{n-s(x)}$ est dérivable sur $[0,1]$ en tant que fonction polynomiale, et pour tout $p \in [0,1]$ on a

$$\begin{aligned} f'(p) &= s(x)p^{s(x)-1}(1-p)^{n-s(x)} + p^{s(x)}(n-s(x))(1-p)^{n-s(x)-1} \\ &= p^{s(x)-1}(1-p)^{n-s(x)-1}\left(s(x)(1-p) - (n-s(x))p\right) \\ &= p^{s(x)-1}(1-p)^{n-s(x)-1}\left(s(x) - np\right), \end{aligned}$$

or $p^{s(x)-1}(1-p)^{n-s(x)-1} > 0$ donc

$$f'(p) \geqslant 0 \iff s(x) - np \geqslant 0 \iff p \leqslant \frac{s(x)}{n}.$$

Ainsi, f admet son maximum sur $[0,1]$ en le point $p = \frac{s(x)}{n}$.

L'estimateur du maximum de vraisemblance pour p est donc égal, dans chacun des trois cas envisagés, à

$$\frac{s(x)}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

c'est-à-dire à la moyenne empirique des observations x_i (égale à la fréquence empirique de réponses favorables).

3. La variable aléatoire T représente le temps d'attente du premier succès lors de la répétition d'expériences de Bernoulli indépendantes de probabilité de succès p ; elle suit donc une loi géométrique $\mathcal{G}(p)$.
4. Pour tout $k \in \mathbb{N}^*$, on a $L'(k,p) = (1-p)^{k-1}p$.

   Si $k = 1$, alors $L'(k,p) = p$ donc la vraisemblance est maximisée pour $p = 1$.

   Supposons à présent que $k \geqslant 2$. La fonction $f : p \mapsto (1-p)^{k-1}p$ est dérivable sur $[0,1]$ en tant que fonction polynomiale, et pour tout $p \in [0,1]$ on a

$$\begin{aligned} f'(p) &= -(k-1)(1-p)^{k-2}p + (1-p)^{k-1} \\ &= (-(k-1)p + (1-p))(1-p)^{k-2} \\ &= (1-kp)(1-p)^{k-2}, \end{aligned}$$

   or $(1-p)^{k-2} > 0$ donc $f'(p) > 0$ si $0 < p < \frac{1}{k}$ et $f'(p) < 0$ si $\frac{1}{k} < p < 1$; ainsi, f admet un maximum en $\frac{1}{k}$.

   Dans les deux cas envisagés, l'estimateur du maximum de vraisemblance pour p est $\frac{1}{k}$.

   Ce résultat est conforme à l'intuition dans la mesure où la valeur moyenne prise par T est $\mathbb{E}(T) = \frac{1}{p}$, ce qui conduit à estimer p par l'inverse de la valeur observée de T.

L'estimation par maximum de vraisemblance constitue un pan important des statistiques théoriques. Cette technique d'estimation, qui possède ses avantages et inconvénients intrinsèques, s'intéresse entre autres à la recherche de statistiques exhaustives *qui permettent de concentrer en un minimum de bits d'informations les données nécessaires pour une estimation paramétrique efficace, ainsi qu'au caractère optimal (en termes d'absence de biais et de précision) des estimateurs obtenus.*

Correction de l'exercice 62.

1. (a) Si $X \sim \mathcal{B}(p)$, alors pour tout $s \in [0,1]$ on a
$$g_X(s) = (1-p)s^0 + ps^1 = 1 - p + ps.$$
   (b) Si $X \sim \mathcal{B}(N,p)$, alors pour tout $s \in [0,1]$ on a
$$g_X(s) = \sum_{k=0}^{N} s^k \binom{N}{k} p^k (1-p)^{N-k} = \sum_{k=0}^{N} \binom{N}{k} (sp)^k (1-p)^{N-k} = (sp+1-p)^N$$
   d'après la formule du binôme. On remarque que l'on retrouve bien le résultat de la question précédente dans le cas où $N = 1$.
   (c) Si $X \sim \mathcal{U}(\llbracket 1, N \rrbracket)$, alors pour tout $s \in [0,1]$ on a
$$g_X(s) = \frac{1}{N} \sum_{k=1}^{N} s^k,$$
   ce qui se réécrit
$$g_X(s) = \frac{s - s^{N+1}}{N(1-s)}$$
   lorsque $s \neq 1$.
2. (a) On a $g_X(0) = \mathbb{E}(0^X) = \mathbb{P}(X = 0)$ puisque 0^X prend la valeur 1 si X prend la valeur 0, et la valeur 0 sinon. On a par ailleurs $g_X(1) = \mathbb{E}(1) = 1$.
   (b) i. En tant que fonction polynomiale, g_X est indéfiniment dérivable sur $[0,1]$. On a [48] pour tout $s \in [0,1]$:
$$g'_X(s) = \sum_{k=1}^{N} k s^{k-1} \mathbb{P}(X = k),$$
$$g''_X(s) = \sum_{k=2}^{N} k(k-1) s^{k-2} \mathbb{P}(X = k)$$
   et
$$g'''_X(s) = \sum_{k=3}^{N} k(k-1)(k-2) s^{k-3} \mathbb{P}(X = k).$$

48. Remarquons que ces trois relations se réécrivent grâce au théorème de transfert sous la forme $g'_X(s) = \mathbb{E}(Xs^{X-1})$, $g''_X(s) = \mathbb{E}(X(X-1)s^{X-2})$ et $g'''_X(s) = \mathbb{E}(X(X-1)(X-2)s^{X-3})$, ce qui signifie que l'on peut « dériver sous l'espérance » le terme s^X.

On trouve alors

$$g'_X(1) = \sum_{k=1}^{N} k\,\mathbb{P}(X=k) = \mathbb{E}(X),$$

$$g''_X(1) = \sum_{k=2}^{N} k(k-1)\mathbb{P}(X=k) = \mathbb{E}(X(X-1)) = \mathbb{E}(X^2) - \mathbb{E}(X)$$

et

$$\begin{aligned} g'''_X(1) &= \sum_{k=3}^{N} k(k-1)(k-2)\mathbb{P}(X=k) \\ &= \mathbb{E}(X(X-1)(X-2)) = \mathbb{E}(X^3) - 3\,\mathbb{E}(X^2) + 2\mathbb{E}(X) \end{aligned}$$

d'après le théorème de transfert et par linéarité de l'espérance.

ii. D'après la question précédente, on a

$$\begin{aligned} \mathbb{E}(X^3) &= g'''_X(1) + 3\,\mathbb{E}(X^2) - 2\,\mathbb{E}(X) \\ &= g'''_X(1) + 3\left(g''_X(1) + \mathbb{E}(X)\right) - 2\,\mathbb{E}(X) \\ &= g'''_X(1) + 3\,g''_X(1) + \mathbb{E}(X) \\ &= g'''_X(1) + 3\,g''_X(1) + g'_X(1), \end{aligned}$$

or pour tout $s \in [0,1]$ on a $g_X(s) = (1-p+sp)^N$ donc

$$g'_X(s) = Np\,(1-p+sp)^{N-1},$$

$$g''_X(s) = N(N-1)p^2(1-p+sp)^{N-2}$$

et

$$g'''_X(s) = N(N-1)(N-2)p^3(1-p+sp)^{N-3},$$

d'où $g'_X(1) = Np$, $g''_X(1) = N(N-1)p^2$ et $g'''_X(1) = N(N-1)(N-2)p^3$, et enfin

$$\mathbb{E}(X^3) = N(N-1)(N-2)p^3 + 3N(N-1)p^2 + Np.$$

(c) D'après la formule de Taylor exprimant les coefficients d'une fonction polynomiale en fonction de ses dérivées en 0, pour tout $k \in \mathbb{N}$ on a

$$g_X^{(k)}(0) = k!\,\mathbb{P}(X=k), \quad \text{donc} \quad \mathbb{P}(X=k) = \frac{g_X^{(k)}(0)}{k!}.$$

3. Si Y est une variable aléatoire à valeurs dans $[\![0,N]\!]$ telle que $g_X = g_Y$, alors $g_X^{(k)}(0) = g_Y^{(k)}(0)$ pour tout $k \in \mathbb{N}$, ce qui implique d'après la question précédente que l'on a $\mathbb{P}(X=k) = \mathbb{P}(Y=k)$ quel que soit $k \in [\![0,N]\!]$, c'est-à-dire que X et Y ont la même loi.

4. Signalons tout d'abord que ce *problème des moments* a été traité dans l'exercice 57 par des techniques d'algèbre linéaire.

   Soit Y une VARD à valeurs dans $[\![0, N]\!]$. Supposons que X et Y ont les mêmes moments d'ordre k pour tout $k \in [\![0, N]\!]$. Pour tout $n \in [\![0, N]\!]$, un calcul similaire à celui effectué dans la question 2(b)i montre que

$$g_X^{(n)}(1) = \mathbb{E}(X(X-1)\dots(X-n+1)),$$

   d'où l'on déduit par linéarité de l'espérance que $g_X^{(n)}(1)$ s'exprime comme une combinaison linéaire des moments d'ordre $1, 2, \dots, n$ de X (puisque le produit $X(X-1)\dots(X-n+1)$ contient n termes). Ces moments étant les mêmes que ceux de Y, on en déduit que

$$\begin{aligned} g_Y^{(n)}(1) &= \mathbb{E}(Y(Y-1)\dots(Y-n+1)) \\ &= \mathbb{E}(X(X-1)\dots(X-n+1)) = g_X^{(n)}(1). \end{aligned}$$

   Ainsi, les dérivées d'ordre $1, \dots, N$ de g_X et g_Y en 1 coïncident, et il en va de même pour les dérivées d'ordre 0 puisque $g_X(1) = 1 = g_Y(1)$. Comme g_X et g_Y sont des fonctions polynomiales de degré au plus N, la formule de Taylor en 1 donne alors :

$$\forall s \in [0,1], \quad g_X(s) = \sum_{k=0}^{N} \frac{g_X^{(k)}(1)}{k!}(s-1)^k = \sum_{k=0}^{N} \frac{g_Y^{(k)}(1)}{k!}(s-1)^k = g_Y(s),$$

   donc $g_X = g_Y$, ce qui implique d'après la question précédente que X et Y ont même loi.

Correction de l'exercice 63.

1. Si $s \in [0,1]$, la variable aléatoire s^X est à valeurs dans $[0,1]$ puisque X est à valeurs dans $\mathbb{N}$; en tant que variable aléatoire bornée, elle admet donc une espérance. L'expression de $g_X(s)$ donnée dans l'énoncé est alors justifiée par le théorème de transfert.

2. (a) Si $X \sim \mathcal{P}(\lambda)$, alors pour tout $s \in [0,1]$ on a

$$g_X(s) = \sum_{k=0}^{+\infty} s^k p_k = \sum_{k=0}^{+\infty} s^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{(\lambda s)^k}{k!} = e^{-\lambda}e^{\lambda s} = e^{\lambda(s-1)}.$$

   (b) Si $g_X \in \mathcal{G}(p)$, alors pour tout $s \in [0,1]$ on a

$$\begin{aligned} g_X(s) &= \sum_{k=1}^{+\infty} s^k p(1-p)^{k-1} = ps\sum_{k=1}^{+\infty}(s(1-p))^{k-1} \\ &= ps\sum_{i=0}^{+\infty}(s(1-p))^i = \frac{ps}{1-s(1-p)} \end{aligned}$$

   puisque $s(1-p) \in [0,1[$.

3. Pour les mêmes raisons que dans le cas fini étudié dans l'exercice précédent, on a $g_X(0) = p_0$ et $g_X(1) = 1$.
4. (a) Soient $a \in [0, 1[$ et $s \in [0, 1]$. On a

$$|g_X(s) - g_X(a)| = \left| \sum_{k=0}^{+\infty} p_k (s^k - a^k) \right|$$

par linéarité des sommes de séries convergentes, d'où

$$|g_X(s) - g_X(a)| \leqslant \sum_{k=0}^{+\infty} p_k |s^k - a^k| = \sum_{k=1}^{+\infty} p_k |s^k - a^k|$$

d'après l'inégalité triangulaire. Or pour tout $k \geqslant 1$, on a

$$s^k - a^k = (s - a)(s^{k-1} + s^{k-2}a + \cdots + sa^{k-2} + a^{k-1})$$

d'après la formule de Bernoulli, d'où, en utilisant le fait que $s \in [0, 1]$ et $a \in [0, 1[$:

$$|s^k - a^k| \leqslant |s - a|(1 + a + \cdots + a^{k-1}) \leqslant |s - a| \sum_{i=0}^{+\infty} a^i = \frac{|s - a|}{1 - a}.$$

Ainsi, on a

$$|g_X(s) - g_X(a)| \leqslant \sum_{k=1}^{+\infty} p_k \frac{|s - a|}{1 - a} \leqslant \frac{|s - a|}{1 - a} \sum_{k=0}^{+\infty} p_k = \frac{|s - a|}{1 - a}.$$

(b) Soit $a \in [0, 1[$. Le résultat de la question précédente permet alors d'écrire

$$|g_X(s) - g_X(a)| \leqslant \frac{|s - a|}{1 - a} \xrightarrow[s \to a]{} 0,$$

donc $g_X(s) \xrightarrow[s \to a]{} g_X(a)$ par encadrement, la limite étant à comprendre au sens d'une limite à droite si $a = 0$. Ainsi, g_X est continue sur $[0, 1[$.

(c) i. L'existence de l'entier N est justifiée par le fait que le reste de la série convergente des p_k tend vers 0.

ii. La fonction polynomiale $p : s \mapsto \sum_{k=0}^{N} p_k s^k$ étant continue en 1, il existe $\eta \in]0, 1]$ tel que pour tout $s \in [1 - \eta, 1]$ on ait $p(s) \geqslant p(1) - \frac{\varepsilon}{2}$, ce qui fournit l'inégalité recherchée.

iii. Soit $s \in [1 - \eta, 1]$. On a $g_X(s) \leqslant g_X(1) = 1$ par croissance de g_X (qui résulte de la positivité des p_k). On a par ailleurs

$$g_X(s) = \sum_{k=0}^{+\infty} p_k s^k = \sum_{k=0}^{N} p_k s^k + \sum_{k=N+1}^{+\infty} p_k s^k \geqslant \sum_{k=0}^{N} p_k s^k,$$

or on a, par propriété de η, que

$$\sum_{k=0}^{N} s^k p_k \geqslant \sum_{k=0}^{N} p_k - \frac{\varepsilon}{2},$$

donc

$$g_X(s) \geqslant \sum_{k=0}^{N} p_k - \frac{\varepsilon}{2},$$

et

$$\sum_{k=0}^{N} p_k = \underbrace{\sum_{k=0}^{+\infty} p_k}_{=1} - \sum_{k=N+1}^{+\infty} p_k \geqslant 1 - \frac{\varepsilon}{2}$$

par définition de N, d'où

$$g_X(s) \geqslant 1 - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = 1 - \varepsilon,$$

ce qui fournit l'encadrement

$$1 - \varepsilon \leqslant g_X(s) \leqslant 1.$$

On a donc montré que quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que si $s \in [1-\eta, 1]$ alors $1 - \varepsilon \leqslant g_X(s) \leqslant 1$: ainsi, on a $g_X(s) \xrightarrow[s\to 1^-]{} 1$, c'est-à-dire que g_X est continue à gauche en 1.

5. (a) Pour tout $s \in [0,1[$, on a $kp_k s^{k-1} \underset{k\to+\infty}{=} o\left(\frac{1}{k^2}\right)$ par croissance comparée [49] puisque les p_k sont à valeurs dans $[0,1]$; or la série de terme général $\frac{1}{k^2}$ converge en tant que série de Riemann de paramètre $2 > 1$, donc la série des $kp_k s^{k-1}$ converge elle aussi par comparaison des séries à termes positifs.

(b) Soit $a \in [0,1[$, et soit $s \in [0,1]$ tel que $s \neq a$. On a alors

$$\frac{g_X(s) - g_X(a)}{s-a} - \sum_{k=1}^{+\infty} kp_k a^{k-1} = \sum_{k=1}^{+\infty} p_k \left(\frac{s^k - a^k}{s-a} - ka^{k-1} \right)$$

par linéarité des sommes de séries convergentes. Or pour tout $k \geqslant 1$, on a d'après la formule de Bernoull :

$$\begin{aligned} \frac{s^k - a^k}{s-a} - ka^{k-1} &= s^{k-1} + s^{k-2}a + \ldots + sa^{k-2} + a^{k-1} - ka^{k-1} \\ &= \sum_{i=0}^{k-1} (s^i - a^i)a^{k-1-i} = \sum_{i=1}^{k-1} (s^i - a^i)a^{k-1-i}, \end{aligned}$$

donc, en utilisant l'inégalité triangulaire et l'inégalité $|s^i - a^i| \leqslant \frac{|s-a|}{1-a}$ prouvée dans la question 4(a) pour tout $i \geqslant 1$:

$$\begin{aligned} \left| \frac{s^k - a^k}{s-a} - ka^{k-1} \right| &\leqslant \frac{|s-a|}{1-a} \sum_{i=1}^{k-1} a^{k-1-i} \\ &= \frac{|s-a|}{1-a} \sum_{j=0}^{k-2} a^j \quad \text{en posant } j = k-1-i \\ &\leqslant \frac{|s-a|}{1-a} \sum_{j=0}^{+\infty} a^j = \frac{|s-a|}{(1-a)^2}. \end{aligned}$$

49. On aurait aussi pu majorer plus finement $kp_k s^{k-1}$ par t^{k-1} pour n'importe quel $t \in]s,1[$, ce qui aurait permis de conclure par comparaison à une série géométrique convergente.

Ainsi, on a par une nouvelle application de l'inégalité triangulaire :

$$\begin{aligned}\left|\frac{g_X(s)-g_X(a)}{s-a}-\sum_{k=1}^{+\infty} kp_k a^{k-1}\right| &\leqslant \sum_{k=1}^{+\infty} p_k \left|\frac{s^k-a^k}{s-a}-ka^{k-1}\right| \\ &\leqslant \sum_{k=1}^{+\infty} p_k \frac{|s-a|}{(1-a)^2} = \frac{|s-a|}{(1-a)^2}\sum_{k=1}^{+\infty} p_k \\ &\leqslant \frac{|s-a|}{(1-a)^2}\sum_{k=0}^{+\infty} p_k = \frac{|s-a|}{(1-a)^2}.\end{aligned}$$

Cette majoration est valable pour tout $s \in [0,1]$ tel que $s \neq a$; comme $\frac{|s-a|}{(1-a)^2} \xrightarrow[s\to a]{} 0$, on en déduit par encadrement que

$$\frac{g_X(s)-g_X(a)}{s-a} \xrightarrow[s\to a]{} \sum_{k=1}^{+\infty} kp_k a^{k-1},$$

c'est-à-dire que g_X est dérivable en a et que $g_X'(a) = \sum_{k=1}^{+\infty} kp_k a^{k-1}$.

(c) Soit $\varepsilon > 0$. Comme la série des kp_k converge (puisque X admet une espérance), il existe $N \in \mathbb{N}$ tel que $\sum_{k=N+1}^{+\infty} kp_k \leqslant \frac{\varepsilon}{2}$, donc $\sum_{k=0}^{N} kp_k \geqslant \mathbb{E}(X) - \frac{\varepsilon}{2}$.

Considérons la fonction polynomiale $h : s \mapsto \sum_{k=0}^{N} p_k s^k$. On sait que h est dérivable en 1 et que $h'(1) = \sum_{k=0}^{N} kp_k$, d'où $h'(1) \geqslant \mathbb{E}(X) - \frac{\varepsilon}{2}$.

Par ailleurs, pour tout $s \in [0,1[$, on a

$$g_X(1) - g_X(s) = \sum_{k=0}^{+\infty}(1-s^k)p_k \geqslant \sum_{k=0}^{N}(1-s^k)p_k = h(1) - h(s),$$

puisque les $1-s^k$ et les p_k sont positifs, donc, en divisant par $1-s>0$:

$$\frac{g_X(1)-g_X(s)}{1-s} \geqslant \frac{h(1)-h(s)}{1-s}.$$

Or comme $\frac{h(1)-h(s)}{1-s)} \xrightarrow[s\to 1^-]{} h'(1)$ (puisque h est dérivable en 1), il existe $\eta \in]0,1]$ tel que pour tout $s \in [1-\eta, 1[$ on ait

$$\frac{h(1)-h(s)}{1-s)} \geqslant h'(1) - \frac{\varepsilon}{2}.$$

On a alors, pour tout $s \in [1-\eta, 1[$:

$$\frac{g_X(1)-g_X(s)}{1-s} \geqslant \frac{h(1)-h(s)}{1-s} \geqslant h'(1) - \frac{\varepsilon}{2} \geqslant \mathbb{E}(X) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = \mathbb{E}(X) - \varepsilon.$$

Mais pour un tel s, en appliquant le théorème des accroissements finis à g_X entre s et 1 (ce qui est loisible puisque g_X est continue sur $[s,1]$ et dérivable sur $]s,1[$), on obtient l'existence d'un $t \in]s,1[$ tel que

$$\frac{g_X(1)-g_X(s)}{1-s} = g_X'(t) = \sum_{k=1}^{+\infty} kp_k t^{k-1}$$

où la dernière égalité provient du résultat de la question précédente, d'où, comme $t \in [0,1]$ et comme les kp_k sont positifs :

$$\frac{g_X(1)-g_X(s)}{1-s} \leqslant \sum_{k=1}^{+\infty} kp_k = \mathbb{E}(X).$$

Ainsi, dès que $s \in [1-\eta, 1[$, on a

$$\mathbb{E}(X) - \varepsilon \leqslant \frac{g_X(1)-g_X(s)}{1-s} \leqslant \mathbb{E}(X).$$

Comme $\varepsilon > 0$ a été choisi arbitrairement, on en déduit que

$$\frac{g_X(1)-g_X(s)}{1-s} \xrightarrow[s\to 1^-]{} \mathbb{E}(X),$$

c'est-à-dire que g_X est dérivable (à gauche) en 1 et vérifie $g_X'(1) = \mathbb{E}(X)$.

(d) i. Pour tout $s \in [0,1[$ on a

$$\frac{g_X(s)-g_X(1)}{s-1} = \sum_{k=0}^{+\infty} p_k \frac{s^k-1}{s-1} = \sum_{k=0}^{+\infty} p_k \sum_{i=0}^{k-1} s^i$$

par linéarité de la somme des séries convergentes. Comme g_X est dérivable à gauche en 1, on a donc

$$\sum_{k=0}^{+\infty} p_k \sum_{i=0}^{k-1} s^i = \frac{g_X(s)-g_X(1)}{s-1} \xrightarrow[s\to 1^-]{} g_X'(1).$$

Or la fonction $s \mapsto \displaystyle\sum_{k=0}^{+\infty} p_k \sum_{i=0}^{k-1} s^i$ est croissante en tant que somme infinie de fonctions croissantes, donc pour tout $s \in [0,1[$ on a

$$\sum_{k=0}^{+\infty} p_k \sum_{i=0}^{k-1} s^i \leqslant g_X'(1).$$

Les p_k étant positifs, on peut alors affirmer que pour tout $n \in \mathbb{N}^*$ et tout $s \in [0,1[$, on a

$$\sum_{k=1}^{n} p_k \sum_{i=0}^{k-1} s^i \leqslant \sum_{k=1}^{+\infty} p_k \sum_{i=0}^{k-1} s^i \leqslant g_X'(1),$$

ce qu'il fallait démontrer.

ii. Soit $n \in \mathbb{N}^*$. On a montré dans la question précédente que

$$\sum_{k=1}^{n} p_k \sum_{i=0}^{k-1} s^i \leqslant g_X'(1)$$

pour tout $s \in [0,1[$. En passant à la limite lorsque $s \to 1^-$ (ce qui est possible car la somme considérée est finie), on obtient alors

$$\sum_{k=1}^{n} p_k \sum_{i=0}^{k-1} 1 \leqslant g_X'(1), \quad \text{soit} \quad \sum_{k=1}^{n} kp_k \leqslant g_X'(1).$$

Il en résulte que la série de terme général positif kp_k est majorée, donc convergente : ainsi, X admet une espérance.

On peut alors appliquer la question 5(c) pour conclure que $g_X'(1)$ est égal à $\mathbb{E}(X)$.

6. (a) Soit $n \in \mathbb{N}$. Pour tout $s \in [0,1[$, on a

$$g_X(s) = \sum_{k=0}^{n} p_k s^k + \sum_{k=n+1}^{+\infty} p_k s^k.$$

Les p_k étant à valeurs dans $[0,1[$, on peut écrire que quel que soit $s \in [0,1[$, on a

$$0 \leqslant \sum_{k=n+1}^{+\infty} p_k s^k \leqslant \sum_{k=n+1}^{+\infty} s_k = \frac{s^{n+1}}{1-s},$$

d'où $\displaystyle\sum_{k=n+1}^{+\infty} p_k s^k \underset{s\to 0}{=} o(s^n)$. On en déduit que

$$g_X(s) \underset{s\to 0}{=} \sum_{k=0}^{n} p_k s^k + o(s^n),$$

ce qu'il fallait démontrer.

(b) Soit Y une variable aléatoire à valeurs dans $\mathbb{N}$ telle que $g_X = g_Y$. Pour tout $k \in \mathbb{N}$, on pose $q_k := \mathbb{P}(Y = k)$. Supposons que X et Y ne suivent pas la même loi et considérons le plus petit entier k_0 tel que $p_{k_0} \neq q_{k_0}$. En appliquant le résultat de la question précédente à $g_X = g_Y$ et à $n = k_0$ on obtient alors

$$\sum_{k=0}^{k_0-1} p_k s^k + p_{k_0} s^{k_0} + o(s^{k_0}) \underset{s\to 0}{=} \sum_{k=0}^{k_0-1} q_k s^k + q_{k_0} s^{k_0} + o(s^{k_0})$$

soit, en soustrayant $\displaystyle\sum_{k=0}^{k_0-1} p_k s^k = \sum_{k=0}^{k_0-1} q_k s^k$ aux deux membres de l'égalité :

$$(p_{k_0} - q_{k_0}) s^{k_0} \underset{s\to 0}{=} o(s^{k_0}).$$

Ainsi, on a nécessairement $p_{k_0} - q_{k_0} = 0$, ce qui contredit l'hypothèse faite sur k_0. Les variables aléatoires X et Y suivent donc la même loi, ce qui clôt la preuve. [50]

Correction de l'exercice 64.

1. La fonction génératrice g_{X+Y} existe puisque $X + Y$ est elle aussi à valeurs dans $\mathbb{N}$. Soit $s \in [0,1]$. On a alors

$$g_{X+Y}(s) = \mathbb{E}(s^{X+Y}) = \mathbb{E}(s^X s^Y),$$

or s^X et s^Y sont indépendantes d'après le lemme des coalitions, donc

$$g_{X+Y}(s) = \mathbb{E}(s^X)\mathbb{E}(s^Y) = g_X(s)g_Y(s),$$

ce qui permet de conclure que $g_{X+Y} = g_X g_Y$.

2. Pour tout $s \in [0,1]$, en utilisant le résultat de la question précédente et celui de la question 1(b) de l'exercice 62 on obtient

$$g_{X+Y}(s) = g_X(s)g_Y(s) = (sp + 1 - p)^m (sp + 1 - p)^n = (sp + 1 - p)^{m+n}.$$

Ainsi, la fonction génératrice de $X+Y$ est celle d'une variable aléatoire suivant une loi $\mathcal{B}(m+n, p)$. La fonction génératrice caractérisant la loi des variables aléatoires à valeurs dans $\mathbb{N}$ d'après la question 6(b) de l'exercice 63, on a $X + Y \sim \mathcal{B}(m+n, p)$.

3. Pour tout $s \in [0,1]$, en utilisant le résultat de la question 1 du présent exercice et celui de la question 2(a) de l'exercice 63 on obtient

$$g_{X+Y}(s) = g_X(s)g_Y(s) = e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)}.$$

Ainsi, la fonction génératrice de $X+Y$ est celle d'une variable aléatoire suivant une loi $\mathcal{P}(\lambda + \mu)$. La fonction génératrice caractérisant la loi des variables aléatoires à valeurs dans $\mathbb{N}$ d'après la question 6(b) de l'exercice 63, on a donc $X + Y \sim \mathcal{P}(\lambda + \mu)$.

4. Supposons que X et Y soient à support dans $[\![1, 6]\!]$ et que $X+Y \sim \mathcal{U}([\![2, 12]\!])$.

Les fonctions g_X et g_Y sont alors des fonctions polynomiales de degré au plus 6 admettant 0 pour racine (puisque les termes constants de ces fonctions,

50. Remarquons, en nous avançant sur les résultats de cours sur les développements limités, que l'on aurait pu déduire l'égalité des lois de la relation

$$\sum_{k=0}^{n} p_k s^k + o(s^n) \underset{s\to 0}{=} \sum_{k=0}^{n} q_k s^k + o(s^n),$$

valable pour tout $n \in \mathbb{N}$ d'après la question précédente, en invoquant l'unicité du développement limité d'une fonction en un point – un résultat qui repose sur le même argument que celui développé ici.

$\mathbb{P}(X=0)$ et $\mathbb{P}(Y=0)$, sont nuls) ; par conséquent, il existe $P, Q \in \mathbb{R}_5[x]$ telles que $g_X(x) = xP(x)$ et $g_Y(x) = xQ(x)$ pour tout $x \in [0,1]$.

La fonction $g_{X+Y} = g_X g_Y$, qui est la fonction génératrice d'une variable aléatoire de loi $\mathcal{U}([\![2,12]\!])$, vérifie donc :

$$\forall x \in [0,1], \quad g_{X+Y}(x) = \frac{1}{11}\sum_{k=2}^{12} x^k = x^2 \cdot \frac{1}{11}\sum_{k=0}^{10} x^k,$$

mais aussi, d'après la question 1 :

$$\forall x \in [0,1], \quad g_{X+Y}(x) = g_X(x)g_Y(x) = x^2 P(x)Q(x).$$

Ainsi, pour tout $x \in]0,1]$, on a

$$\frac{1}{11}\sum_{k=0}^{10} x^k = P(x)Q(x).$$

Cette égalité entre expressions polynomiales étant valide sur une infinité de points, elle l'est nécessairement sur $\mathbb{C}$ tout entier : on a alors

$$\forall z \in \mathbb{C}, \quad \frac{1}{11}\sum_{k=0}^{10} z^k = P(z)Q(z).$$

Les degrés de P et Q étant au plus égaux à 5, cette égalité montre qu'ils sont en fait exactement de degré 5 puisque le terme de gauche est de degré 10.

Or les racines de la fonction polynomiale $z \mapsto \frac{1}{11}\sum_{k=0}^{10} z^k$ sur $\mathbb{C}$ sont les racines onzièmes de l'unité différentes de 1, puisque 1 n'est pas racine et que l'on a :

$$\forall z \neq 1, \quad \sum_{k=0}^{10} z^k = \frac{1-z^{11}}{1-z}.$$

Aucune de ces racines n'est réelle. Ainsi, la fonction polynomiale $z \mapsto P(z)Q(z)$ n'admet pas de racine sur $\mathbb{R}$, ce qui contredit le fait que P et Q sont de degré 5, donc admettent chacun au moins une racine réelle.

Ainsi, il n'est pas possible de choisir les lois de X et de Y à support dans $[\![1,6]\!]$ de sorte à avoir $X+Y \sim \mathcal{U}([\![2,12]\!])$. Ce résultat signifie qu'il n'est pas possible de piper deux dés de sorte que la somme de leurs résultats amène une valeur tirée uniformément dans $[\![2,12]\!]$.

Correction de l'exercice 65.

1. On a $f(1) = \sum_{n=1}^{+\infty} p_n = 1 - p_0$, donc $p_0 = 1 - f(1)$.

   On considère qu'un nombre infini de lancers a lieu[51] (même après la ruine éventuelle du joueur), et pour tout $k \in \mathbb{N}^*$, on note P_k l'événement « le joueur obtient pile au k-ième lancer ».

   La ruine du joueur advient au terme du premier lancer si et seulement si ce lancer amène face ; on a alors $p_1 = \mathbb{P}(\overline{P_1}) = q$.

   La ruine ne peut advenir au terme de deux lancers puisqu'un premier lancer amenant pile fait passer le capital du joueur à 2 €, et ce capital ne peut être perdu en un seul nouveau lancer : on a $p_2 = 0$. Plus généralement, on remarque que la ruine du joueur ne peut survenir au terme d'un nombre pair de lancers puisqu'elle suppose que le joueur a obtenu face exactement une fois de plus que pile, et donc qu'un nombre impair de lancers ont eu lieu : pour tout $k \in \mathbb{N}^*$, on a donc $p_{2k} = 0$.

   La ruine du joueur advient au terme du troisième lancer si et seulement si le joueur a obtenu une fois pile puis deux fois face au cours des trois premiers lancers ; on a donc $p_3 = \mathbb{P}(P_1 \cap \overline{P_2} \cap \overline{P_3}) = \mathbb{P}(P_1)\mathbb{P}(\overline{P_2})\mathbb{P}(\overline{P_3}) = pq^2$ par indépendance des lancers.

2. (a) i. Soit $j \in [\![1, n-1]\!]$. La trajectoire (c'est-à-dire la suite de résultats des lancers successifs) menant d'un capital de 2 € à un capital de 1 € en retombant pour la première fois à 1 € à l'issue de j lancers est exactement une trajectoire de ruine en j lancers, c'est-à-dire une trajectoire qui mènerait d'un capital de 1 € à la ruine du joueur en j lancers. Ainsi, une fois le premier pile obtenu, la probabilité que l'événement E_j se réalise est p_j : on a donc

   $$\mathbb{P}(E_j) = \mathbb{P}(P_1)\mathbb{P}(E_j \mid P_1) = pp_j.$$

   ii. Soit $j \in [\![1, n-1]\!]$. La réalisation de G_j suppose que E_j soit réalisé ; or conditionnellement à E_j, l'événement G_j est réalisé si et seulement si les $n-j$ lancers entre $j+2$ et $n+1$ (inclus) forment une trajectoire de ruine, ce qui advient avec probabilité p_{n-j}. Ainsi, on a

   $$\mathbb{P}(G_j) = \mathbb{P}(E_j)\mathbb{P}(G_j \mid E_j) = pp_jp_{n-j}.$$

   iii. Étudions la probabilité que le joueur soit ruiné en exactement $n+1$ lancers. Par définition, cette probabilité vaut p_{n+1}. Or une trajectoire amenant la ruine du joueur en $n+1$ lancers commence nécessairement par un pile (puisque $n+1 \geqslant 3$), puis par un retour à un capital de 1 € (en un nombre de lancers que l'on note $j \in [\![1, n-1]\!]$), et par une ruine (à l'issue des $n-j$ lancers suivants). En d'autres termes,

51. Cette précision permet de ne pas avoir à discuter du fait qu'un lancer est bel et bien réalisé et de modéliser la situation à partir d'expériences binaires, indépendantes et de même probabilité de réussite.

l'événement « le joueur est ruiné à l'issue du lancer $n+1$ » peut se réécrire comme l'union disjointe $\bigsqcup_{j=1}^{n-1} G_j$. On a donc

$$p_{n+1} = \mathbb{P}\left(\bigsqcup_{j=1}^{n-1} G_j\right) = \sum_{j=1}^{n-1} \mathbb{P}(G_j),$$

soit

$$p_{n+1} = p \sum_{j=1}^{n-1} p_j p_{n-j}$$

d'après le résultat de la question précédente.

(b) Soit $s \in [0,1]$. En utilisant la formule de sommation admise et le fait que $p_n s^n$ est le terme général positif d'une série convergente, on obtient

$$\begin{aligned}
ps(f(s))^2 &= ps\left(\sum_{n=1}^{+\infty} p_n s^n\right)^2 \\
&= ps \sum_{n=2}^{+\infty} \sum_{j=1}^{n-1} p_j s^j p_{n-j} s^{n-j} \\
&= \sum_{n=2}^{+\infty} \sum_{j=1}^{n-1} p p_j p_{n-j} s^{n+1} \\
&= p p_1 p_1 s^3 + \sum_{n=3}^{+\infty} \sum_{j=1}^{n-1} p p_j p_{n-j} s^{n+1} \\
&= pq^2 s^3 + \sum_{n=3}^{+\infty} p_{n+1} s^{n+1},
\end{aligned}$$

où la dernière égalité provient du résultat de la question 2(a)iii, valable pour tout $n \geqslant 3$. Ainsi, on a

$$ps(f(s))^2 = pq^2 s^3 + \sum_{k=4}^{+\infty} p_k s^k = \sum_{k=3}^{+\infty} p_k s^k$$

puisque $p_3 = pq^2$, d'où, en utilisant le fait que $p_2 = 0$:

$$ps(f(s))^2 = \sum_{k=2}^{+\infty} p_k s^k = \sum_{k=1}^{+\infty} p_k s^k - p_1 s = f(s) - p_1 s = f(s) - qs.$$

3. (a) La fonction f est croissante en tant que somme infinie de fonctions croissantes. Une démonstration plus pédestre de ce résultat est la suivante : si $s_1, s_2 \in [0,1]$ sont tels que $s_1 \leqslant s_2$, alors pour tout $k \geqslant 1$ on a $s_1^k \leqslant s_2^k$ (puisque $x \mapsto x^k$ est croissante sur $[0,1]$), donc $p_k s_1^k \leqslant p_k s_2^k$ puisque $p_k \geqslant 0$, et donc, en sommant sur k les inégalités ainsi obtenues :

$$\sum_{k=1}^{+\infty} p_k s_1^k \leqslant \sum_{k=1}^{+\infty} p_k s_2^k \quad \text{soit} \quad f(s_1) \leqslant f(s_2),$$

d'où la croissance de f.

(b) On a $f(1) = g(1) - p_0 = 1 - p_0$ donc $f(1) \leqslant 1$; la croissance de f implique donc que $f(s) \leqslant 1$ pour tout $s \in [0,1]$.

(c) Soit $s \in]0,1[$. On a

$$\begin{aligned} 1 - 4pqs^2 &= 1 - 4p(1-p)s^2 \\ &= 1 - 4ps^2 + 4p^2s^2 > 1 - 4ps + 4p^2s^2 = (1-2ps)^2, \end{aligned}$$

donc $\sqrt{1 - 4pqs^2}$ existe et

$$\sqrt{1 - 4pqs^2} > \sqrt{(1-2ps)^2} = |1 - 2ps|$$

puisque $ps > ps^2$ et par stricte croissance de la fonction racine carrée.

L'inégalité à démontrer est équivalente à $\sqrt{1 - 4pqs^2} > 2ps - 1$. Elle est évidente si $2ps - 1 \leqslant 0$ puisque $\sqrt{1 - 4pqs^2} > |1 - 2ps| \geqslant 0$. Si $2ps - 1 > 0$, à présent, on a

$$\sqrt{1 - 4pqs^2} > |1 - 2ps| = 2ps - 1,$$

donc l'inégalité est vraie une fois encore. Ainsi, on a bien

$$\frac{1 + \sqrt{1 - 4pqs^2}}{2ps} > 1$$

dans les deux cas.

(d) En utilisant l'équation fonctionnelle (2), on trouve pour tout $s \in [0,1]$:

$$f(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} \quad \text{ou} \quad f(s) = \frac{1 + \sqrt{1 - 4pqs^2}}{2ps}.$$

Les résultats des deux questions précédentes assurent alors que quel que soit $s \in]0,1[$, on a

$$f(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps}.$$

Il est par ailleurs clair que $f(0) = 0$.

4. (a) On sait que g est continue en 1 d'après la question 4(c)iii de l'exercice 63 ; c'est donc le cas de $f = g - p_0$. Ainsi, on a

$$\begin{aligned} f(1) &= \lim_{s \to 1^-} \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} = \frac{1 - \sqrt{1 - 4pq}}{2p} = \frac{1 - |1 - 2p|}{2p} \\ &= \begin{cases} 1 & \text{si } \ 1 - 2p \geqslant 0 \\ \frac{1-p}{p} & \text{si } \ 1 - 2p < 0. \end{cases} \end{aligned}$$

Or $p_0 = 1 - f(1)$ d'après la question 1, donc

$$p_0 = \begin{cases} 0 & \text{si } \ 1 - 2p \geqslant 0 \\ \frac{2p-1}{p} & \text{si } \ 1 - 2p < 0. \end{cases}$$

(b) La probabilité de ruine du joueur (qui vaut $1 - p_0 = f(1)$) est donc égale à 1 si $p < \frac{1}{2}$, ce qui ne choque pas l'intuition : on a dans ce cas $q > p$, donc on s'attend par la loi des grands nombres à obtenir strictement plus de faces que de piles à partir d'un certain nombre de lancers (et donc la ruine du joueur), comme on le verra dans la question suivante.

La probabilité de ruine est aussi égale à 1 si $p = q = \frac{1}{2}$, ce qui est moins évident mais traduit le fait [52] qu'une marche aléatoire symétrique sur $\mathbb{Z}$ visite chaque point de $\mathbb{Z}$ avec probabilité 1.

En revanche, si $p > \frac{1}{2}$ (c'est-à-dire si la pièce est biaisée en faveur des piles), la probabilité que la ruine advienne est strictement inférieure à 1.

Ainsi, le jeu s'arrête presque sûrement si et seulement si $p \leqslant \frac{1}{2}$.

(c) Considérons toujours que le joueur réalise une infinité de lancers (en continuant à gagner et perdre même après que sa fortune a atteint 0) et pour tout $n \geqslant 1$, notons X_n le capital dont il dispose à l'issue du n-ième lancer. Pour tout $k \in \mathbb{N}^*$, on note $Y_k := \mathbb{1}_{P_k}$ la variable aléatoire prenant la valeur 1 si le joueur obtient pile au k-ième lancer et 0 sinon. Les Y_k sont donc des variables aléatoires indépendantes de même loi de Bernoulli de paramètre p, et pour tout $n \geqslant 1$, on peut écrire

$$X_n = 1 + \sum_{k=1}^{n} (2Y_k - 1).$$

Supposons à présent que $q > p$, c'est-à-dire que $p < \frac{1}{2}$. D'après [53] la loi faible des grands nombres appliquée aux variables aléatoires indépendantes $2Y_k - 1$, qui admettent pour espérance $2p - 1 < 0$ et admettent un moment d'ordre 2 en tant que VARD prenant un nombre fini de valeurs, on a :

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}(2Y_k - 1) - (2p-1)\right| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0,$$

donc en prenant $\varepsilon = |2p - 1| > 0$:

$$\mathbb{P}\left(\left|\frac{X_n - 1}{n} - (2p-1)\right| \geqslant |2p-1|\right) \xrightarrow[n\to+\infty]{} 0.$$

Or un réel x positif vérifie automatiquement $|x - (2p - 1)| \geqslant |2p - 1|$ puisque $2p - 1 < 0$, donc

$$0 \leqslant \mathbb{P}\left(\frac{X_n - 1}{n} \geqslant 0\right) \leqslant \mathbb{P}\left(\left|\frac{X_n - 1}{n} - (2p-1)\right| \geqslant |2p-1|\right) \xrightarrow[n\to+\infty]{} 0,$$

52. Ce résultat est encore vrai pour une marche aléatoire symétrique sur $\mathbb{Z}^2$ (c'est-à-dire un processus décrivant la trajectoire d'une particule qui se meut sur $\mathbb{Z}^2$ en choisissant à chaque étape de se déplacer vers la gauche, la droite, le haut ou le bas avec probabilité $\frac{1}{4}$ et indépendamment du passé), mais faux en dimension 3 et au-delà ; c'est le *théorème de Pólya* étudié dans l'exercice 60).

53. La preuve plutôt technique qui suit formalise une idée très simple : si $q > p$, la loi des grands nombres assure que le capital du joueur prend des valeurs négatives avec une probabilité arbitrairement proche de 1, donc que la ruine du joueur advient presque sûrement.

d'où, par encadrement,

$$\mathbb{P}\left(\frac{X_n - 1}{n} \geqslant 0\right) \xrightarrow[n\to+\infty]{} 0,$$

soit $\mathbb{P}\left(X_n - 1 \geqslant 0\right) \xrightarrow[n\to+\infty]{} 0$, donc $\mathbb{P}\left(X_n > 0\right) = \mathbb{P}(X_n \geqslant 1) \xrightarrow[n\to+\infty]{} 0$.

Or la ruine du joueur advient dès lors que $X_n \leqslant 0$ pour un certain n, donc pour tout $n \geqslant 1$ on a $(X_n \leqslant 0) \subset (T > 0)$, d'où l'on déduit que $\mathbb{P}(T > 0) \geqslant \mathbb{P}(X_n \leqslant 0)$. Comme $\mathbb{P}(X_n \leqslant 0) \xrightarrow[n\to+\infty]{} 1$, on en conclut que $\mathbb{P}(T > 0) = 1$, c'est-à-dire que la ruine du joueur advient avec probabilité 1.

Un raisonnement similaire permettrait de montrer que si $p > q$, alors le capital du joueur tend presque sûrement vers $+\infty$ dans les cas où sa ruine n'advient pas.

5. (a) Il suffit de prendre $a = \frac{1}{2}$ dans l'expression rappelée. On obtient, lorsque $t \to 0$:

$$\begin{aligned}\sqrt{1+t} &= 1 + \sum_{k=1}^{n} \frac{\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(\frac{1}{2} - k + 1\right)}{k!} t^k + o(t^n) \\ &= 1 + \frac{t}{2} + \sum_{k=2}^{n} \frac{1(-1)(-3)\cdots(3-2k)}{2^k k!} t^k + o(t^n) \\ &= 1 + \frac{t}{2} + \sum_{k=2}^{n} (-1)^{k-1} \frac{1 \cdot 3 \cdots (2k-3)}{2^k k!} t^k + o(t^n),\end{aligned}$$

or pour tout $k \geqslant 2$ on a

$$\begin{aligned}1 \cdot 3 \cdots (2k-3) &= \frac{1 \cdot 2 \cdot 3 \cdots (2k-3)}{2 \cdot 4 \cdots (2k-4)} \\ &= \frac{(2k-3)!}{2^{k-2} \cdot 1 \cdot 2 \cdots (k-2)} = \frac{(2k-3)!}{2^{k-2}(k-2)!}\end{aligned}$$

donc lorsque $t \to 0$:

$$\begin{aligned}\sqrt{1+t} &= 1 + \frac{t}{2} + \sum_{k=2}^{n} (-1)^{k-1} \frac{(2k-3)!}{2^{2k-2}(k-2)!k!} t^k + o(t^n) \\ &= 1 + \frac{t}{2} + \sum_{k=2}^{n} \left(-\frac{1}{4}\right)^{k-1} \frac{(2k-3)!}{(k-2)!k!} t^k + o(t^n).\end{aligned}$$

(b) Soit $n \geqslant 2$. En appliquant le développement limité que nous venons de donner à l'expression de f déterminée dans la question 4(d), avec

$t = -4pqs^2$, on peut écrire que lorsque $s \to 0$:

$$\begin{aligned}
f(s) &= \frac{1-\sqrt{1-4pqs^2}}{2ps} \\
&= \frac{1}{2ps}\left(1-\left(1-\frac{4pqs^2}{2}+\sum_{k=2}^{n}\left(-\frac{1}{4}\right)^{k-1}\frac{(2k-3)!}{(k-2)!k!}(-4pqs^2)^k+o((-4pqs^2)^n)\right)\right) \\
&= \frac{1}{2ps}\left(2pqs^2+\sum_{k=2}^{n}\frac{4(2k-3)!}{(k-2)!k!}p^kq^ks^{2k}+o(s^{2n})\right) \\
&= qs+\sum_{k=2}^{n}\frac{2(2k-3)!}{(k-2)!k!}p^{k-1}q^ks^{2k-1}+o(s^{2n-1}).
\end{aligned}$$

Or $f(s) \underset{s\to 0}{=} \sum_{i=1}^{2n-1} p_i s^i + o(s^{2n-1})$ d'après la question 6(a) de l'exercice 63.

Par unicité du développement limité d'une fonction en un point (qui peut se redémontrer de la même façon que le résultat de la question 6(b) de l'exercice 63), on obtient donc $p_1 = q$ (ce que l'on savait déjà), et pour tout $k \in [\![2, n]\!]$:

$$p_{2k} = 0 \quad \text{et} \quad p_{2k-1} = \frac{2(2k-3)!}{(k-2)!k!}p^{k-1}q^k = \frac{1}{k-1}\binom{2k-2}{k}p^{k-1}q^k.$$

Remarquons que la nullité des p_{2k} est conforme à la remarque que nous avions faite à la question 2.

L'entier $n \geqslant 2$ étant quelconque, on peut écrire que pour tout $k \geqslant 2$:

$$p_{2k} = 0 \quad \text{et} \quad p_{2k-1} = \frac{1}{k-1}\binom{2k-2}{k}p^{k-1}q^k.$$

et donc que pour tout $i \geqslant 1$, on a

$$\mathbb{P}(T = 2i) = 0 \quad \text{et} \quad \mathbb{P}(T = 2i+1) = \frac{1}{i}\binom{2i}{i+1}p^iq^{i+1}.$$

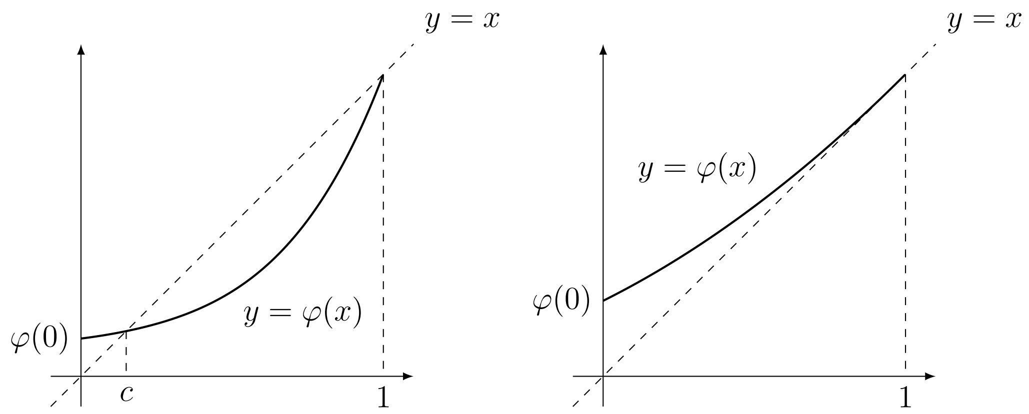
Pour achever de décrire la loi de T, il faut ajouter à cette définition, outre le fait que $\mathbb{P}(T = 1) = p_1 = 0$, la relation

$$p_0 = \begin{cases} 0 & \text{si} \quad p \leqslant \frac{1}{2} \\ 2-\frac{1}{p} & \text{si} \quad p > \frac{1}{2} \end{cases}$$

obtenue dans la question 4(a).

Correction de l'exercice 66. On remarque que dans tous les cas, 1 est un point fixe de φ par hypothèse.

Il est utile de représenter graphiquement φ avant de commencer :



Allure du graphe de φ. À gauche, cas $\varphi'(1) > 1$. À droite, cas $\varphi'(1) = 1$.

1. Supposons que $\varphi'(1) > 1$.

   Le taux d'accroissement $\frac{\varphi(1)-\varphi(x)}{1-x} = \frac{1-\varphi(x)}{1-x}$ admet dans ce cas une limite strictement supérieure à 1 lorsque $x < 1$ tend vers 1 (puisque cette limite vaut $\varphi'(1)$), ce qui assure qu'il existe $x' \in]0,1[$ tel que $\frac{1-\varphi(x')}{1-x'} > 1$ et donc tel que $1 - \varphi(x') > 1 - x'$, c'est-à-dire $\varphi(x') < x'$. La fonction $\psi : x \mapsto \varphi(x) - x$ vérifie donc

   $$\psi(0) = \varphi(0) > 0 \quad \text{et} \quad \psi(x') = \varphi(x') - x' < 0,$$

   or ψ est continue comme différence de deux fonctions continues, donc elle s'annule en un point $c \in]0, x'[$ d'après le théorème des valeurs intermédiaires. Or un tel point c est un point fixe de φ, strictement inférieur à 1 ; ainsi, φ admet au moins deux points fixes sur $[0,1]$ (c et 1). Montrons à présent que ce sont les seuls.

   La fonction ψ est dérivable deux fois sur $]0,1[$ en tant que différence de fonctions deux fois dérivables, et on a $\psi'' = \varphi'' > 0$, donc ψ' est strictement croissante. Supposons qu'il existe un troisième point fixe de φ sur $[0,1]$, donc un troisième point d'annulation[54] de ψ sur $[0,1]$. On note a, b et 1 les trois points fixes obtenus, avec $a < b < 1$ (c étant donc égal à a ou à b). En appliquant le théorème de Rolle entre a et b (ce qui est possible puisque ψ est continue sur $[a,b]$, dérivable sur $]a,b[$ et vérifie $\psi(a) = \psi(b)$), on obtient un point d'annulation a' de ψ' dans $]a,b[$. En l'appliquant entre b et 1, on obtient un point d'annulation b' de ψ' dans $]b',1[$. On a alors $a' < b'$ et $\psi'(a') = \psi'(b') = 0$; mais cela est contraire au fait que ψ' est strictement croissante. Ainsi, on a montré par l'absurde que φ n'admet que deux points fixes sur $[0,1]$, ce qui clôt la preuve.

2. Supposons à présent que $\varphi'(1) \leqslant 1$.

   Comme φ' est strictement croissante sur $[0,1]$, on a pour tout $x \in]0,1[$ la majoration $\varphi'(x) < \varphi'(1) \leqslant 1$, donc $\varphi'(x) < 1$. Si φ admettait deux points

54. Prenez le temps d'essayer de vous représenter la situation : ψ est une fonction s'annulant trois fois sur $[0,1]$ et telle que $\psi'' > 0$... l'impossibilité d'une telle configuration est graphiquement évidente !

fixes $c \in]0,1[$ et 1, le théorème des accroissements finis[55] appliqué à φ entre c et 1 (ce qui est possible car φ est continue sur $[c,1]$ et dérivable sur $]c,1[$) donnerait

$$|c-1| = |\varphi(c) - \varphi(1)| = |\varphi'(x)| \cdot |c-1| < |c-1|$$

pour un certain $x \in]c,1[$, ce qui est absurde. La fonction φ admet donc 1 pour unique point fixe sur $[0,1]$.

3. Considérons à nouveau la fonction $\psi : x \mapsto \varphi(x) - x$, qui vérifie $\psi'' = \varphi''$ et dont les points d'annulation sont les points fixes de φ.

   Supposons que $\varphi'(1) > 1$. La fonction ψ s'annule en deux points de $[0,1]$ d'après la question 1 : $c < 1$ et 1. Sa dérivée s'annule donc en un point $d \in]c,1[$ d'après le théorème de Rolle. Comme ψ' est croissante, elle prend des valeurs négatives avant ce point d'annulation et positives après, c'est-à-dire que ψ est décroissante sur $[0,d]$ et croissante sur $[d,1]$. Comme ψ s'annule en c et en 1, on a $\psi(x) \geqslant 0$ si $x \in [0,c]$ et $\psi(x) \leqslant 0$ si $x \in [c,1]$. En d'autres termes :

   $$\varphi(x) - x \geqslant 0 \quad \text{si } x \in [0,c] \quad \text{et} \quad \varphi(x) - x \leqslant 0 \quad \text{si } x \in [c,1].$$

   Considérons à présent le cas où $\varphi'(1) \leqslant 1$. La fonction ψ s'annule alors uniquement en 1 d'après la question 2 ; comme elle est continue et vérifie $\psi(0) > 0$, elle est donc positive sur $[0,1]$ d'après le théorème des valeurs intermédiaires. Ainsi, on a $\varphi(x) - x \geqslant 0$ pour tout $x \in [0,1]$.

   Remarquons que ces résultats sont conformes aux graphes donnés au début de la question 1.

4. Notons $c \in [0,1]$ le plus petit point fixe de φ sur $[0,1]$ (qui est donc égal à 1 si $\varphi'(1) \leqslant 1$).

   La fonction φ est croissante et $0 \leqslant \varphi(0) \leqslant \varphi(c) = c$, donc l'intervalle $[0,c]$ est stable par φ. Comme $u_0 = 0 \in [0,c]$, on en conclut par une récurrence immédiate que $(u_n)_{n\in\mathbb{N}}$ est à valeurs dans $[0,c]$.

   D'après la question précédente, on a $\varphi(x) \geqslant x$ pour tout $x \in [0,c]$; ainsi, pour tout $n \in \mathbb{N}$ on a $\varphi(u_n) \geqslant u_n$, soit $u_{n+1} \geqslant u_n$, ce qui montre que $(u_n)_{n\in\mathbb{N}}$ est croissante. Comme elle est par ailleurs majorée par c, elle converge d'après le théorème de la limite monotone vers un réel $\ell \in [0,c]$. En passant à la limite dans la relation de récurrence $u_{n+1} = \varphi(u_n)$ grâce à la continuité de φ, on obtient $\ell = \varphi(\ell)$. Ainsi, ℓ est un point de fixe de φ dans $[0,c]$: on a donc $\ell = c$, ce qui clôt la preuve.

5. Si $p_0 = 0$, chaque individu a au moins un enfant avec probabilité 1, donc on a $\mathbb{P}(\forall n \in \mathbb{N}, X_n \geqslant 1) = 1$. L'extinction est donc de probabilité nulle.

6. Si $p_0 + p_1 = 1$ et $p_0 = 0$, on a vu dans la question précédente que l'extinction est de probabilité nulle.

   Supposons à présent que $p_0 + p_1 = 1$ et $p_0 > 0$ (donc $p_1 < 1$). Chaque individu a un enfant avec probabilité p_1 et n'en a aucun avec probabilité p_0. Si $n \in \mathbb{N}$,

55. Ici encore, une (tentative de) représentation graphique est très utile pour comprendre pourquoi φ ne peut pas admettre deux points fixes.

l'événement $(X_n = 1)$ est réalisé si et seulement si l'unique individu de chaque génération $0, 1, \dots, n-1$ engendre un enfant. Par indépendance des $\xi_{i,j}$, on a donc :

$$\begin{aligned}\forall n \in \mathbb{N}^*, \quad \mathbb{P}(X_n = 0) &= 1 - \mathbb{P}\left((\xi_{1,1} = 1) \cap (\xi_{2,1} = 1) \cap \cdots \cap (\xi_{n,1} = 1)\right) \\ &= 1 - p_1^n.\end{aligned}$$

La suite d'événements $((X_n = 0))_{n \in \mathbb{N}}$ étant croissante (puisque l'extinction de la population à une génération donnée implique que les générations suivantes sont vides), on a alors, d'après le théorème de la limite monotone :

$$\mathbb{P}(\exists n \in \mathbb{N} \, : \, X_n = 0) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} (X_n = 0)\right) = \lim_{n \to +\infty} \mathbb{P}(X_n = 0) = 1$$

puisque $p_1^n \xrightarrow[n \to +\infty]{} 0$ (rappelons que $p_1 \in [0, 1[$).

L'extinction de la population est donc presque sûre.

7. Rappelons que l'on a démontré dans l'exercice 63 que g est dérivable sur $[0, 1]$ et que pour tout $s \in [0, 1]$ on a

$$g'(s) = \sum_{k=1}^{+\infty} k p_k s^{k-1} = \sum_{k=0}^{+\infty} (k+1) p_{k+1} s^k.$$

La démarche utilisée dans la question 5(b) de l'exercice 63 permet plus généralement de démontrer le résultat suivant : si $(a_k)_{n \in \mathbb{N}}$ est une suite de réels positifs telle que la série de terme général a_k converge (sans nécessairement être de somme 1), alors la fonction $f : x \mapsto \sum_{k=0}^{+\infty} a_k s^k$ est bien définie et dérivable sur $[0, 1[$, et elle vérifie

$$\forall s \in [0, 1[, \quad f'(s) = \sum_{k=1}^{+\infty} k a_k s^{k-1}.$$

On dit qu'une *série entière de rayon de convergence* 1 *est dérivable terme à terme sur* $[0, 1[$.

On peut à présent appliquer ce résultat aux $a_k := (k+1) p_{k+1}$, qui forment le terme général d'une série convergente : la fonction g' est donc dérivable sur $[0, 1[$, et pour tout $s \in [0, 1[$ on a

$$g''(s) = \sum_{k=1}^{+\infty} k a_k s^{k-1} = \sum_{k=1}^{+\infty} k(k+1) p_{k+1} s^{k-1} = \sum_{k=2}^{+\infty} (k-1) k p_k s^{k-2}.$$

La stricte positivité de g'' sur $]0, 1[$ provient alors du fait que les p_k sont tous positifs et que l'un au moins des p_k, avec $k \geqslant 2$, l'est strictement (puisque $p_0 + p_1 < 1$).

8. *Présentons l'esprit général du calcul que nous allons effectuer. Pour exprimer la fonction génératrice de X_{n+1} en fonction de celle de X_n, on doit tenir*

compte du fait que l'expression de X_{n+1} donnée par l'énoncé dépend de deux types d'aléas : la valeur de X_n, c'est-à-dire le nombre d'individus présents à la génération n, et le nombre d'enfants de chacun de ces individus. Si l'on connaît la valeur prise par X_n, il n'est pas difficile de conclure en utilisant le fait que tous les individus se reproduisent indépendamment et suivant la même loi : on choisit donc de conditionner par la valeur de X_n, *c'est-à-dire de supprimer temporairement une source d'aléa pour se concentrer sur l'autre.*
Par bonheur, X_n et les $\xi_{n+1,\,j}$ sont indépendantes, ce qui permet d'écrire l'égalité cruciale

$$\mathbb{P}((X_{n+1}=i)\cap(X_n=k)) = \mathbb{P}\left(\sum_{j=1}^{k}\xi_{n,\,j}=i\right)\mathbb{P}(X_n=k)$$

qui donne, via *la formule des probabilités totales, un lien entre la loi de X_{n+1} et celle de X_n ne dépendant que de la loi des $\xi_{i,j}$.*

Soit $n\in\mathbb{N}$ et soit $s\in[0,1]$. D'après le théorème de transfert et la formule des probabilités totales appliquée au système complet d'événements $(X_n=k)_{k\in\mathbb{N}}$, on peut écrire :

$$\begin{aligned}
g_{n+1}(s) = \mathbb{E}\left(s^{X_{n+1}}\right) &= \sum_{i=0}^{+\infty} s^i\,\mathbb{P}(X_{n+1}=i)\\
&= \sum_{i=0}^{+\infty} s^i \sum_{k=0}^{+\infty}\mathbb{P}((X_{n+1}=i)\cap(X_n=k))\\
&= \sum_{i=0}^{+\infty} s^i \sum_{k=0}^{+\infty}\mathbb{P}\left(\left(\sum_{j=1}^{X_n}\xi_{n+1,j}=i\right)\cap(X_n=k)\right)\\
&= \sum_{i=0}^{+\infty} s^i \sum_{k=0}^{+\infty}\mathbb{P}\left(\left(\sum_{j=1}^{k}\xi_{n+1,j}=i\right)\cap(X_n=k)\right),
\end{aligned}$$

or X_n et les $\xi_{n+1,j}$ sont indépendantes d'après le lemme des coalitions puisque X_n ne dépend que des $\xi_{m,j}$ pour $m\in[\![1,n]\!]$, donc pour tous $i,k\in\mathbb{N}$ on a

$$\mathbb{P}\left(\left(\sum_{j=1}^{k}\xi_{n+1,j}=i\right)\cap(X_n=k)\right) = \mathbb{P}\left(\sum_{j=1}^{k}\xi_{n+1,\,j}=i\right)\mathbb{P}(X_n=k).$$

Ainsi,

$$\begin{aligned}
g_{n+1}(s) &= \sum_{i=0}^{+\infty} s^i \sum_{k=0}^{+\infty}\mathbb{P}\left(\sum_{j=1}^{k}\xi_{n+1,\,j}=i\right)\mathbb{P}(X_n=k)\\
&= \sum_{k=0}^{+\infty}\mathbb{P}(X_n=k)\sum_{i=0}^{+\infty} s^i\,\mathbb{P}\left(\sum_{j=1}^{k}\xi_{n+1,\,j}=i\right)\\
&= \sum_{k=0}^{+\infty}\mathbb{P}(X_n=k)\mathbb{E}\left(s^{\sum_{j=1}^{k}\xi_{n+1,\,j}}\right)
\end{aligned}$$

où la deuxième égalité est assurée par le théorème de Fubini pour les séries à termes positifs et la dernière par le théorème de transfert. Les $s^{\xi_{n+1,\,j}}$ étant indépendantes (d'après le lemme des coalitions) et de même loi que $s^{\xi_{1,\,1}}$, on a donc :

$$\begin{aligned}
g_{n+1}(s) &= \sum_{k=0}^{+\infty} \mathbb{P}(X_n = k)\, \mathbb{E}\left(s^{\xi_{n+1,1}} \cdots s^{\xi_{n+1,k}} \right) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}(X_n = k)\, \mathbb{E}(s^{\xi_{n+1,1}}) \cdots \mathbb{E}(s^{\xi_{n+1,k}}) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}(X_n = k) \left(\mathbb{E}\left(s^{\xi_{1,1}} \right) \right)^k \\
&= \sum_{k=0}^{+\infty} \mathbb{P}(X_n = k) g(s)^k \\
&= \mathbb{E}\left(g(s)^{X_n} \right) \\
&= g_n(g(s))
\end{aligned}$$

grâce au théorème de transfert et au fait que $g(s) \in [0,1]$ si $s \in [0,1]$, ce qui clôt la preuve.

9. (a) Pour tout $n \in \mathbb{N}$, on a d'après la question précédente (et une récurrence évidente) que $u_n = g_n(0)$, donc $u_n = \mathbb{P}(X_n = 0)$ puisque g_n est la fonction génératrice de X_n.

   (b) Les événements $(X_n = 0)$ formant une suite croissante, on a d'après le théorème de la limite monotone :

   $$\mathbb{P}(\exists n \in \mathbb{N} \,:\, X_n = 0) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} (X_n = 0) \right) = \lim_{n \to +\infty} \mathbb{P}(X_n = 0) = \lim_{n \to +\infty} u_n.$$

   (c) La fonction g est définie sur $[0,1]$ et à valeurs dans $[0,1]$. Elle est continue sur $[0,1]$ (par la question 4 de l'exercice 63), croissante sur $[0,1]$ (comme somme de fonctions croissantes) et dérivable deux fois sur $[0,1[$ (par la question 7 du présent exercice). On a par ailleurs $g'(s) > 0$ pour tout $s \in]0,1[$ d'après la question 7, ainsi que $g(0) = p_0 > 0$ et $g(1) = 1$, donc le résultat du préliminaire analytique (question 4) s'applique et montre que $(u_n)_{n \in \mathbb{N}}$ converge vers le plus petit point fixe de g sur $[0,1]$.

   D'après la question précédente, la probabilité d'extinction de la population est donc le plus petit point fixe de g sur $[0,1]$.

10. On sait grâce à la question 5(c) de l'exercice 63 que g est dérivable à gauche en 1 puisque X_1 (qui est de même loi que les $\xi_{i,j}$) admet une espérance, et que l'on a $g'(1) = \mathbb{E}(X_1)$.

    Ainsi, si $\mathbb{E}(X_1) > 1$ alors $g'(1) > 1$; or on a vu dans la question 1 que dans ce cas, g admet un point fixe $c \in [0,1[$: d'après la question précédente, la probabilité d'extinction de la population est alors égale à c.

    On a de même vu dans la question 2 que si $\mathbb{E}(X_1) \leqslant 1$, alors g admet 1 pour unique point fixe : dans ce cas, la probabilité d'extinction de la population est donc égale à 1.

On a donc démontré le résultat suivant : si le nombre d'enfants moyen par individu est inférieur ou égal à 1, alors la population s'éteint presque sûrement (sauf dans le cas déterministe où $p_1 = 1$, que nous avons exclu de notre discussion en supposant que $0 < p_0 < p_0 + p_1 < 1$), et elle survit avec une probabilité non nulle dans le cas contraire.

11. Dans ce cas, la probabilité d'extinction de la population est le plus petit point fixe de la fonction $g : s \mapsto p_2 s^2 + p_1 s + p_0$, donc la plus petite solution de l'équation polynomiale $g(x) - x = \frac{3}{5}x^2 - \frac{4}{5}x + \frac{1}{5} = 0$ dans $[0,1]$. Or les deux solutions de cette équation sont $\frac{1}{3}$ et 1, si bien que la population s'éteint avec probabilité $\frac{1}{3}$ (on retrouve bien une probabilité d'extinction différente de 1, conformément au fait que $\mathbb{E}(X_1) = \frac{7}{5} > 1$).

12. Pour tout $n \geqslant 0$, on rappelle que la fonction génératrice g_n de X_n est la composée de g n fois par elle-même (avec $g_0 = \mathrm{Id}_{[0,1]}$).

    Pour tout $n \in \mathbb{N}$, on considère la propriété $\mathcal{P}_n$: « g_n est dérivable en 1 et $g_n'(1) = g(1)^n$ ».

    La proposition $\mathcal{P}_0$ est trivialement vraie puisque $g_0 = \mathrm{Id}_{[0,1]}$ est de dérivée constante égale à 1. À présent, si $n \in \mathbb{N}$ est tel que $\mathcal{P}_n$ soit vraie, alors la fonction $g_{n+1} = g_n \circ g$ est dérivable en 1 en tant que composée de fonctions dérivables, et on a grâce à l'hypothèse de récurrence

$$g_{n+1}'(1) = g_n'(g(1))g'(1) = g_n'(1)g'(1) = g'(1)^{n+1},$$

    donc $\mathcal{P}_{n+1}$ est vraie. Ainsi, $\mathcal{P}_n$ est vraie pour tout $n \geqslant 0$ d'après le principe de récurrence.

    Pour tout $n \in \mathbb{N}$, la variable aléatoire X_n admet donc une espérance donnée par $g_n'(1) = \mathbb{E}(X_1)^n$. La moyenne de la population croît donc de façon géométrique, ce qui rappelle un modèle de population déterministe.

13. Supposons que les $\xi_{i,j}$ n'admettent pas d'espérance (ce qui rend g non dérivable en 1). On peut alors raisonner de deux façons pour montrer que la probabilité d'extinction est non nulle.

    Solution analytique : On va tout d'abord utiliser la divergence de la série de terme général kp_k pour montrer que

$$\frac{g(1) - g(x)}{1 - x} \xrightarrow[x \to 1^-]{} +\infty. \tag{3}$$

    Pour établir cette relation[56], on fixe $A > 0$ et on remarque qu'il existe $N \in \mathbb{N}$ tel que $\sum_{k=1}^{N} kp_k > A$. Or $\sum_{k=1}^{N} kp_k s^{k-1} \xrightarrow[s \to 1^-]{} \sum_{k=1}^{N} kp_k$ par continuité des fonctions

56. Cet argument est à comparer à celui utilisé dans la question 5(c) de l'exercice 63 dans le cas d'une variable admettant une espérance.

polynomiales, donc il existe $\delta > 0$ tel que pour tout $s \in [1-\delta, 1[$ on ait $\sum_{k=1}^{N} kp_k s^{k-1} > A$, donc

$$g'(s) \geqslant \sum_{k=1}^{N} kp_k s^{k-1} > A.$$

Soit $x \in [1-\delta, 1[$; en appliquant le théorème des accroissements finis à g entre x et 1 (ce qui est possible puisque g est continue sur $[x, 1]$ et dérivable sur $]x, 1[$) on obtient l'existence d'un $s_x \in]x, 1[$ tel que

$$\frac{g(1)-g(x)}{1-x} = g'(s_x), \quad \text{d'où} \quad \frac{g(1)-g(x)}{1-x} > A.$$

Comme cette dernière inégalité est valide pour tout x dans un voisinage à gauche de 1 et comme A est choisi arbitrairement, on a bien montré la relation (3).

On peut alors conclure à l'existence d'un point fixe $c < 1$ de g comme dans la question 1. Tout le reste du raisonnement s'applique et montre que c est la probabilité d'extinction de la population.

Solution probabiliste (couplage) : on va formaliser l'argument selon lequel « si la population ne s'éteint pas toujours lorsque les $\xi_{i,j}$ sont d'espérance finie supérieure à 1, alors elle ne s'éteint pas non plus toujours lorsque les $\xi_{i,j}$ prennent des valeurs suffisamment grandes pour ne pas admettre d'espérance ». Pour ce faire, on va comparer notre population à une population imaginaire plus petite dont l'extinction n'est pas quasi-certaine.

Comme la série de terme général positif kp_k diverge, il existe $N \in \mathbb{N}$ tel que $\sum_{k=0}^{N} kp_k > 1$. En définissant $\xi'_{i,j} := \xi_{i,j} \mathbb{1}_{\xi_{i,j} \leqslant N}$ pour tout $(i,j) \in (\mathbb{N}^*)^2$ (c'est-à-dire que $\xi'_{i,j}$ prend la même valeur que $\xi_{i,j}$ si l'individu j de la génération $i-1$ a moins de N enfants, et la valeur 0 sinon), on voit que les $\xi'_{i,j}$ admettent une espérance

$$\mathbb{E}(\xi'_{i,j}) = \sum_{k=0}^{N} kp_k$$

finie mais strictement supérieure à 1.

Considérons le processus de branchement $(X'_n)_{n\in\mathbb{N}}$ construit à partir des lois de reproduction $\xi'_{i,j}$ (c'est-à-dire que ce processus décrit une population telle que $X'_0 = 1$ et telle que pour tout $n \in \mathbb{N}$ on ait

$$X'_{n+1} = \sum_{j=1}^{X'_n} \xi'_{n,j}.$$

On a donc trivialement $0 \leqslant X'_n \leqslant X_n$ pour tout $n \in \mathbb{N}$ (puisque les individus constituant la population décrite par les X'_n ont structurellement moins d'enfants que les individus correspondants dans la population d'origine), si bien que l'extinction du processus initial implique celle de celui décrit par les X'_n. Ainsi, on a la comparaison suivante entre les probabilités d'extinction :

$$\mathbb{P}(\exists n \in \mathbb{N} \,:\, X_n = 0) \leqslant \mathbb{P}(\exists n \in \mathbb{N} \,:\, X'_n = 0).$$

Or $(X'_n)_{n\in\mathbb{N}}$ est un processus de branchement associé à des lois de reproduction d'espérance finie strictement supérieure à 1, donc $\mathbb{P}(\exists n \in \mathbb{N} \,:\, X'_n = 0) < 1$. Ainsi, on a

$$\mathbb{P}(\exists n \in \mathbb{N} \,:\, X_n = 0) < 1,$$

ce qui signifie que la population étudiée s'éteint avec probabilité strictement inférieure à 1. Ceci clôt la preuve.