Espaces probabilisés.

1. Déterminer une distribution de probabilités $(p_\omega)_{\omega\in\Omega}$ sur l'ensemble $\Omega = [\![1, n]\!]$, telle que la probabilité de l'événement $[\![1, k]\!]$ soit proportionnelle à k^2.

- -

Une distribution de probabilités sur Ω est une suite finie $(p_i)_{1\le i\le n}$ de réels positifs telle que $\displaystyle\sum_{i=1}^{n} p_i = 1$. La probabilité associée sur Ω est alors définie par $P(I) = \displaystyle\sum_{i\in I} p_i$ pour toute partie I de Ω. Il doit, de plus, exister un réel α strictement positif tel que $P\big([\![1, k]\!]\big) = \alpha k^2$ pour tout k, soit $\displaystyle\sum_{i=1}^{k} p_i = \alpha k^2$. Pour $k = n$, on voit que nécessairement $\alpha = \dfrac{1}{n^2}$. Puis, pour tout $k \in [\![1, n]\!]$,

$$p_k = P\big(\{k\}\big) = P\big([\![1, k]\!]\big) - P\big([\![1, k-1]\!]\big) = \frac{1}{n^2}\left(k^2 - (k-1)^2\right) = \frac{2k-1}{n^2}\ .$$

Réciproquement, le lecteur est invité à vérifier que cette suite de nombres $(p_k)_{1\le k\le n}$ convient.

---

2. En Palombie, lorsqu'il fait soleil un jour, il fait soleil le lendemain une fois sur 10 et, lorsqu'il pleut un jour, il pleut le lendemain 6 fois sur 10.

a. Il a fait soleil un lundi. Quelle est la probabilité qu'il fasse soleil le jeudi qui suit ?

b. Sachant qu'il a fait soleil le jour J, on note p_n la probabilité qu'il fasse soleil le jour $J + n$. Calculer p_n et $\displaystyle\lim_{n\to+\infty} p_n$. On pourra considérer pour tout n l'événement E_n = "il fait soleil le jour $J + n$" et travailler sur le vecteur-colonne $X_n = \begin{pmatrix} P(E_n) \\ P(\overline{E_n}) \end{pmatrix}$.

- -

a. et b. Disons que le lundi est le jour 0. En notant E_n l'événement "'il fait soleil le jour n", on a les relations $P(E_{n+1}|E_n) = \dfrac{1}{10}$ et $P(\overline{E_{n+1}}|\overline{E_n}) = \dfrac{3}{5}$, donc $P(E_{n+1}|\overline{E_n}) = \dfrac{2}{5}$. Pour tout entier naturel n, $(E_n, \overline{E_n})$ est un système complet d'événements. La formule des probabilités totales donne donc

$$\begin{aligned} P(E_{n+1}) &= P(E_{n+1}|E_n)\, P(E_n) + P(E_{n+1}|\,\overline{E_n})\, P(\overline{E_n}) \\ &= \frac{1}{10}\, P(E_n) + \frac{2}{5}\, P(\overline{E_n}) \\ &= \frac{1}{10}\, P(E_n) + \frac{2}{5}\left(1 - P(E_n)\right) \\ &= \frac{2}{5} - \frac{3}{10}\, P(E_n)\ . \end{aligned}$$

Ainsi, s'il fait soleil le jour 0, la suite (p_n), où $p_n = P(E_n)$ est la probabilité d'ensoleillement, vérifie $p_0 = 1$ et la relation de récurrence $p_{n+1} = \dfrac{2}{5} - \dfrac{3}{10}\, p_n$. On reconnaît une suite arithmético-géométrique. L'équation $l = \dfrac{2}{5} - \dfrac{3}{10}\, l$ admet pour solution $l = \dfrac{4}{13}$ et, en posant $r_n = p_n - \dfrac{4}{13}$, on vérifie la relation $r_{n+1} = -\dfrac{3}{10} r_n$, d'où $r_n = \left(-\dfrac{3}{10}\right)^n r_0 = \dfrac{9}{13} \times \left(-\dfrac{3}{10}\right)^n$, puis $p_n = r_n + \dfrac{4}{13} = \dfrac{9}{13} \times \left(-\dfrac{3}{10}\right)^n + \dfrac{4}{13}$, ce qui répond à la question **b.** sans utiliser de calcul matriciel. On en déduit aussi $\displaystyle\lim_{n\to+\infty} p_n = \frac{4}{13} \simeq 0,308$.

Pour répondre à la question **a.**, on évalue pour $n = 3$: $p_3 = \dfrac{9}{13} \times \left(-\dfrac{3}{10}\right)^3 + \dfrac{4}{13} = \dfrac{289}{1000} =$

$0,289$. On peut retrouver ce résultat en faisant un arbre de probabilité, ils poussent bien en Palombie.

Palons peu, palombien! Passons au calcul matriciel! On a obtenu plus haut la relation $P(E_{n+1}) = \dfrac{1}{10}\,P(E_n) + \dfrac{2}{5}\,P(\overline{E_n})$, on obtient de même $P(\overline{E_{n+1}}) = \dfrac{9}{10}\,P(E_n) + \dfrac{3}{5}\,P(\overline{E_n})$. Les matrices-colonnes X_n vérifient alors la relation de récurrence $X_{n+1} = AX_n$, avec $A = \begin{pmatrix} \dfrac{1}{10} & \dfrac{2}{5} \\ \dfrac{9}{10} & \dfrac{3}{5} \end{pmatrix} = \begin{pmatrix} 0,1 & 0,4 \\ 0,9 & 0,6 \end{pmatrix}$. Donc $X_n = A^n X_0$ avec $X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Et on diagonalise A, pardi! Je n'écris pas les détails de calcul, on obtient $A = PDP^{-1}$ avec $D = \mathrm{diag}\Big(-\dfrac{3}{10}, 1\Big)$ et $P = \begin{pmatrix} -1 & \dfrac{4}{9} \\ 1 & 1 \end{pmatrix}$. Après un peu de calcul, on a

$$A^n = PD^nP^{-1} = \begin{pmatrix} \dfrac{9}{13} \times \Big(-\dfrac{3}{10}\Big)^n + \dfrac{4}{13} & -\dfrac{4}{13} \times \Big(-\dfrac{3}{10}\Big)^n + \dfrac{4}{13} \\ -\dfrac{9}{13} \times \Big(-\dfrac{3}{10}\Big)^n + \dfrac{9}{13} & \dfrac{4}{13} \times \Big(-\dfrac{3}{10}\Big)^n + \dfrac{9}{13} \end{pmatrix},$$

ce qui permet de retrouver les résultats déjà obtenus ci-dessus.

---

3. Une succession d'individus $A_1, \cdots, A_n$ se transmet une information binaire du type "oui" ou "non". Chaque individu A_k transmet l'information qu'il a reçue avec la probabilité p à l'individu A_{k+1} ou la transforme en son contraire avec la probabilité $q = 1 - p$. Chaque individu se comporte indépendamment des autres. Calculer la probabilité π_n pour que l'information reçue par A_n soit identique à celle émise par A_1. Quelle est la limite de π_n quand n tend vers l'infini ?

- -

Pour $k \in [\![1, n]\!]$, notons B_k l'événement: "l'individu A_k a reçu la bonne information (celle émise par A_1)". On a alors $P(B_1) = 1$ et, pour $k \geq 1$, par la formule des probabilités totales:

$$\begin{aligned} \pi_{k+1} = P(B_{k+1}) &= P(B_{k+1}|B_k)\,P(B_k) + P(B_{k+1}|\overline{B_k})\,P(\overline{B_k}) \\ &= p\,P(B_k) + (1-p)\,\big(1 - P(B_k)\big) \\ &= (2p-1)\,\pi_k + (1-p)\,. \end{aligned}$$

On reconnaît une suite arithmético-géométrique. L'équation $l = (2p-1)l + (1-p)$ a pour solution $l = \dfrac{1}{2}$, donc la suite (v_k) définie par $v_k = P(B_k) - \dfrac{1}{2}$ est géométrique, de raison $2p - 1$, ce que l'improbable lecteur se fera un plaisir de vérifier par le calcul. Ainsi, $v_k = (2p-1)^{k-1} v_1 = \dfrac{1}{2}(2p-1)^{k-1}$. Puis $\pi_n = P(B_n) = \dfrac{1}{2}\Big(1 + (2p-1)^{n-1}\Big)$.

Si on a $0 < p < 1$, alors $-1 < 2p - 1 < 1$ et $\lim\limits_{n\to+\infty} \pi_n = \dfrac{1}{2}$.

---

4. Soient $A_1, \cdots, A_n$ des événements mutuellement indépendants. Montrer que la probabilité pour qu'aucun d'eux ne soit réalisée est majorée par $M = \exp\Big(-\sum\limits_{k=1}^{n} P(A_k)\Big)$.

- -

On étudie $P\Big(\bigcap_{k=1}^{n}\overline{A_k}\Big)$. Par indépendance des $\overline{A_k}$, on a

$$P\Big(\bigcap_{k=1}^{n}\overline{A_k}\Big)=\prod_{k=1}^{n}P(\overline{A_k})=\prod_{k=1}^{n}\big(1-P(A_k)\big)\ .$$

Or, pour $x\in[0,1]$, on a $0\le 1-x\le e^{-x}$, d'où

$$P\Big(\bigcap_{k=1}^{n}\overline{A_k}\Big)\le\prod_{k=1}^{n}e^{-P(A_k)}=\exp\Big(-\sum_{k=1}^{n}P(A_k)\Big)=M\ .$$

---

5. Soient m, N, k des entiers naturels au moins égaux à 2, avec $k\le N$. Une urne contient mN boules dont N sont blanches. On tire k boules dans cette urne, et on s'intéresse à la probabilité de tirer j boules blanches ($0\le j\le k$). On pourra poser $p=\dfrac{1}{m}$, le nombre $p\in[0,1]$ est donc la proportion initiale de boules blanches.

a. Quelle est cette probabilité si le tirage est avec remise ? On constate qu'elle ne dépend pas de N, mais qu'elle ne dépend que de la proportion p de boules blanches dans l'urne.

b. Quelle est cette probabilité si le tirage est sans remise ? Elle dépend alors de N.

c. Si N devient très grand devant le nombre k de boules tirées, on a l'intuition que le fait que le tirage s'effectue avec ou sans remise ne va pas changer grand-chose. Le vérifier mathématiquement.

- -

a. À chaque tirage, la probabilité de "succès", c'est-à-dire de tirer une boule blanche, est p, et il y a k tirages indépendants, le nombre de boules blanches tirées suit donc une loi binomiale de paramètres k et p. La probabilité de tirer j boules blanches exactement est donc $\dbinom{k}{j}p^j(1-p)^{k-j}$.

b. Ici, les tirages possibles sont les parties à k éléments de l'ensemble $[\![1,mN]\!]$ de toutes les boules de l'urne, ces tirages sont au nombre de $\dbinom{mN}{k}$ avec équiprobabilité. Les tirages favorables sont ceux constitués d'une partie à j éléments de l'ensemble des N boules blanches et d'une partie à $k-j$ éléments de l'ensemble des $(m-1)N$ boules non blanches. Ils sont au nombre de $\dbinom{N}{j}\dbinom{(m-1)N}{k-j}$, d'où la probabilité

$$\pi_N=\frac{\dbinom{N}{j}\dbinom{(m-1)N}{k-j}}{\dbinom{mN}{k}}\ .$$

c. Il s'agit de vérifier que $\lim\limits_{N\to+\infty} \pi_N = \binom{k}{j} p^j (1-p)^{k-j}$. Pour cela, notons que

$$\binom{N}{j} = \frac{N!}{j!(N-j)!} = \frac{N(N-1)\cdots(N-j+1)}{j!} \underset{N\to+\infty}{\sim} \frac{N^j}{j!} \ :$$

en effet, j étant fixé,le coefficient binomial $\binom{N}{j}$ est une fonction polynomiale de la variable N, de terme dominant $\dfrac{N^j}{j!}$. On obtient de même $\binom{(m-1)N}{k-j} \underset{N\to+\infty}{\sim} \dfrac{(m-1)^{k-j}\,N^{k-j}}{(k-j)!}$ et $\binom{mN}{k} \underset{N\to+\infty}{\sim} \dfrac{m^k\,N^k}{k!}$. Par opérations (licites!) sur les équivalents, on déduit

$$\pi_N \underset{N\to+\infty}{\sim} \frac{N^j}{j!}\ \frac{(m-1)^{k-j}\,N^{k-j}}{(k-j)!}\ \frac{k!}{m^k\,N^k}$$

et, comme on constate que cette expression après simplifications ne dépend en fait pas de N, on peut écrire

$$\lim_{N\to+\infty} \pi_N = \frac{k!}{j!(k-j)!}\ \frac{(m-1)^{k-j}}{m^k} = \binom{k}{j} \frac{\left(\dfrac{1}{p}-1\right)^{k-j}}{\left(\dfrac{1}{p}\right)^k} = \binom{k}{j}\ p^j\ (1-p)^{k-j}\ ,$$

ce que l'on souhaitait obtenir.

6*. On s'intéresse à la survie d'une espèce pour laquelle un individu admet trois descendants avec la probabilité $\dfrac{1}{8}$, un ou deux descendants avec la probabilité $\dfrac{3}{8}$, et enfin aucun descendant avec la probabilité $\dfrac{1}{8}$. On suppose qu'à l'instant initial, la population est composée d'un seul individu. Par conséquent, l'espèce s'éteindra au bout de la première génération avec une probabilité de $x_1 = \dfrac{1}{8}$.

a. Quelle est la probabilité x_2 pour que l'espèce ait disparu à l'issue de la deuxième génération ?

b. Pour tout n entier naturel non nul, on note x_n la probabilité que l'espèce ait disparu à l'issue de la n-ème génération. Montrer que la suite (x_n) vérifie la relation

$$\forall n \in \mathbb{N}^* \qquad x_{n+1} = \frac{1}{8}\left(x_n^3 + 3x_n^2 + 3x_n + 1\right).$$

c. Déterminer $\lim\limits_{n\to+\infty} x_n$.

- -

a. Notons X_k le nombre d'individus de la génération k. Ainsi, $X_0 = 1$ et $x_1 = P(X_1 = 0) = \dfrac{1}{8}$. On a alors

$$\begin{aligned} x_2 &= P(X_2 = 0) \\ &= P(X_1=0) + P\big(\{X_1=1\} \cap \{X_2=0\}\big) + P\big(\{X_1=2\} \cap \{X_2=0\}\big) + P\big(\{X_1=3\} \cap \{X_2=0\}\big) \end{aligned}$$

Mais $P\big(\{X_1=1\} \cap \{X_2=0\}\big) = P(X_2=0|X_1=1)\,P(X_1=1) = \dfrac{1}{8} \times \dfrac{3}{8}$;

$$P\big(\{X_1=2\}\ \cap\ \{X_2=0\}\big)=P(X_2=0|X_1=2)\,P(X_1=2)=\left(\frac{1}{8}\right)^2\times\frac{3}{8}\ ;$$

$$P\big(\{X_1=3\}\ \cap\ \{X_2=0\}\big)=P(X_2=0|X_1=3)\,P(X_1=3)=\left(\frac{1}{8}\right)^3\times\frac{1}{8}.$$

Finalement, $x_2=\dfrac{1}{8}+\dfrac{3}{8}\times\dfrac{1}{8}+\dfrac{3}{8}\times\left(\dfrac{1}{8}\right)^2+\dfrac{1}{8}\times\left(\dfrac{1}{8}\right)^3=\dfrac{729}{4096}\simeq 0,178.$

b. On a $P(X_{n+1}=0)=P(X_1=0)+\displaystyle\sum_{k=1}^{3}P\big(\{X_1=k\}\ \cap\ \{X_{n+1}=0\}\big)$ et, pour tout $k\in\{1,2,3\}$,

$$P\big(\{X_1=k\}\cap\{X_{n+1}=0\}\big)=P(X_{n+1}=0|X_1=k)P(X_1=k)=P(X_n=0|X_0=k)P(X_1=k)\ .$$

Enfin, avec un peu de bon sens, on a $P(X_n=0|X_0=k)=\big(P(X_n=0|X_0=1)\big)^k=x_n^k$. On obtient donc

$$x_{n+1}=P(X_{n+1}=0)=\frac{1}{8}+\frac{3}{8}x_n+\frac{3}{8}x_n^2+\frac{1}{8}x_n^3\ ,$$

ce qui est la relation demandée.

c. Soit $f:x\mapsto\dfrac{1}{8}\big(x^3+3x^2+3x+1\big)=\dfrac{(x+1)^3}{8}$. La fonction f est croissante sur $[0,1]$ avec $f(0)=\dfrac{1}{8}=0,125$, $f(1)=1$, donc l'intervalle $[0,1]$ est stable par f. L'équation $f(x)=x$ admet deux solutions dans $[0,1]$, qui sont le nombre 1 et le nombre $\alpha=\sqrt{5}-2\simeq 0,236$. En fait, l'intervalle $J=[0,\alpha]$ est stable par f puisque son image est $\left[\dfrac{1}{8},\alpha\right]\subset J$. Comme $x_1=\dfrac{1}{8}\in J$, on a $x_n\in J$ pour tout $n\in\mathbb{N}^*$. Sur J, on a $f(x)\geq x$ (il est facile d'étudier le signe de $f(x)-x$ après avoir factorisé par $x-1$), donc la suite (x_n) est croissante. Elle est majorée par α, donc elle converge. Le seul point fixe de f adhérent à l'intervalle J est α, donc $\displaystyle\lim_{n\to+\infty}x_n=\alpha=\sqrt{5}-2$.

---

7. Un groupe de n chasseurs tire simultanément et indépendamment sur n canards. Chaque chasseur ne tire qu'une seule fois et atteint toujours sa cible.

a. Quelle est la probabilité p_n qu'au moins un canard survive ? Déterminer $\displaystyle\lim_{n\to+\infty}p_n$.

b. L'un des canards s'appelle Saturnin. Quelle est la probablité q_n qu'il survive ? Déterminer $\displaystyle\lim_{n\to+\infty}q_n$.

- -

a. On peut modéliser le résultat du tir par une application de l'ensemble $\mathcal{H}_n$ des n chasseurs vers l'ensemble $\mathcal{C}_n$ des n canards, puisque chaque chasseur atteint un et un seul canard. Donc $|\Omega|=n^n$. Au moins un canard survit si et seulement si cette application n'est pas surjective, i.e. si et seulement si elle n'est pas bijective. Or le nombre d'applications bijectives de $\mathcal{H}_n$ vers $\mathcal{C}_n$ est $n!$. La probabilité qu'au moins un canard survive est donc $p_n=1-\dfrac{n!}{n^n}$, et $\displaystyle\lim_{n\to+\infty}p_n=1$ puisque $n!=o(n^n)$ lorsque n tend vers $+\infty$.

b. Saturnin s'en tire si et seulement si l'application correspondant au tir est à valeurs dans l'ensemble $\mathcal{C}_n \setminus \{\text{Saturnin}\}$, de cardinal $n-1$. Comme il y a $(n-1)^n$ applications de l'ensemble $\mathcal{H}_n$ de cardinal n vers l'ensemble $\mathcal{C}_n \setminus \{\text{Saturnin}\}$, de cardinal $n-1$, la probabilité de survie de Saturnin est $q_n = \dfrac{(n-1)^n}{n^n} = \left(1 - \dfrac{1}{n}\right)^n$. Alors $\lim\limits_{n \to +\infty} q_n = e^{-1}$.

Variables aléatoires.

8. Deux variables aléatoires indépendantes X et Y suivent des lois binomiales de tailles n et m et de même paramètre p. Quelle est la loi suivie par la variable aléatoire $Z = X + Y$?

- -

On a clairement $X(\Omega) = [\![1, n]\!]$, $Y(\Omega) = [\![1, m]\!]$ et $Z(\Omega) = [\![1, n+m]\!]$. Pour $k \in [\![1, n+m]\!]$, on a

$$\{Z = k\} = \bigsqcup_{j=0}^{k} \big(\{X = j\} \cap \{Y = k - j\}\big) .$$

Donc $P(Z = k) = \displaystyle\sum_{j=0}^{k} P\big(\{X = j\} \cap \{Y = k - j\}\big) = \sum_{j=0}^{k} P(X = j)\, P(Y = k - j)$ puisque les variables X et Y sont supposées indépendantes. On obtient donc

$$\begin{aligned} P(Z = k) &= \sum_{j=0}^{k} \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{m-(k-j)} \\ &= p^k\, (1-p)^{n+m-k} \sum_{j=0}^{k} \binom{n}{j} \binom{m}{k-j} . \end{aligned}$$

On conclut grâce à l'identité de Vandermonde: $\displaystyle\sum_{j=0}^{k} \binom{n}{j} \binom{m}{k-j} = \binom{n+m}{k}$, que

$$\forall k \in [\![0, n+m]\!] \qquad P(Z = k) = \binom{n+m}{k} p^k\, (1-p)^{(n+m)-k} ,$$

donc Z suit la loi binomiale de paramètres $n + m$ et p.

Remarque. On peut démontrer l'identité de Vandermonde:

- de façon algébrique, en écrivant de deux façons le coefficient de X^k dans le polynôme $(1 + X)^{n+m} = (1 + X)^n (1 + X)^m$;

- de façon combinatoire, en écrivant le nombre de façons de choisir k éléments dans un ensemble E de cardinal $n+m$ que l'on considérerait comme réunion disjointe d'un ensemble F de cardinal n et d'un ensemble G de cardinal m (pour un j donné dans $[\![0, k]\!]$, on doit choisir j éléments dans F et $k - j$ éléments dans G).

---

9. Une urne contient n boules blanches et n boules rouges. On tire simultanément n boules dans celle-ci et on note X le nombre de boules rouges obtenues lors de ce tirage. Quelle est la loi de X, son espérance, sa variance ?

- -

Clairement, $X(\Omega) = [\![0, n]\!]$. On peut d'ailleurs choisir comme univers l'ensemble $\mathcal{P}_n\big([\![1, 2n]\!]\big)$ des parties à n éléments de l'intervalle entier $[\![1, 2n]\!]$ si l'on convient de numéroter les boules de 1 à $2n$, muni de la probabilité uniforme. Le nombre de tirages possibles est $|\Omega| = \dbinom{2n}{n}$. Pour $k \in [\![0, n]\!]$, l'événement $\{X = k\}$ est réalisé si l'on tire k boules parmi les n rouges, et $n - k$ boules parmi les n blanches, le nombre de tirages "favorables" est alors $\dbinom{n}{k}\dbinom{n}{n-k} = \dbinom{n}{k}^2$. Comme on est en situation d'équiprobabilité, on déduit

$$\forall k \in [\![0, n]\!] \qquad P(X = k) = \frac{\dbinom{n}{k}^2}{\dbinom{2n}{n}} \ .$$

Il est possible à partir de cela de calculer $\mathrm{E}(X)$, puis $\mathrm{E}(X^2)$, puis la variance mais bôf! Voici une meilleure idée: considérons que les boules rouges sont numérotées de 1 à n. Pour tout $i \in [\![1, n]\!]$, soit U_i la variable aléatoire qui vaut 1 si la boule i a été tirée, et 0 sinon. Il est clair que $P(U_i = 1) = \dfrac{1}{2}$ (comme on tire la moitié des boules, la boule i a une chance sur deux de figurer dans le tirage), donc U_i suit la loi de Bernoulli $\mathcal{B}\Big(\dfrac{1}{2}\Big)$ et $\mathrm{E}(U_i) = \dfrac{1}{2}$. Enfin, $X = \displaystyle\sum_{i=1}^{n} U_i$, donc par linéarité de l'espérance, $\mathrm{E}(X) = \displaystyle\sum_{i=1}^{n} \frac{1}{2} = \frac{n}{2}$, résultat assez évident intuitivement.

Ensuite, $X^2 = \Big(\displaystyle\sum_{i=1}^{n} U_i\Big)^2 = \displaystyle\sum_{i=1}^{n} U_i^2 + \sum_{i \neq j} U_i U_j$. Comme U_i suit une loi de Bernoulli, $U_i^2 = U_i$. Si $i \neq j$, la variable $U_i U_j$ suit aussi une loi de Bernoulli (les valeurs possibles sont 0 et 1) et, par un raisonnement simple de dénombrement,

$$\mathrm{E}(U_i U_j) = P(U_i U_j = 1) = \frac{\dbinom{2n-2}{2n}}{\dbinom{2n}{n}} = \frac{n-1}{2(2n-1)} \ .$$

Par linéarité de l'espérance,

$$\mathrm{E}(X^2) = \sum_{i=1}^{n} \mathrm{E}(U_i^2) + \sum_{i \neq j} \mathrm{E}(U_i U_j) = n\,\frac{1}{2} + n(n-1)\,\frac{n-1}{2(2n-1)} = \frac{n^3}{2(2n-1)}$$

après réduction sur feu moyen. Et enfin, par la formule de Koenig-Huygens,

$$\mathrm{V}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{n^2}{4(2n-1)} \ .$$

---

10. Une urne contient N boules numérotées de 1 à N. On tire n boules sans remise ($1 \leq n \leq N$). On note X et Y le plus petit et le plus grand numéros obtenus.

a. Pour $k \in [\![1, N]\!]$, déterminer $P(X \geq k)$. En déduire la loi de X.

b. Pour $l \in [\![1, N]\!]$, déterminer $P(Y \leq l)$. En déduire la loi de l.

c. Quelle est la loi conjointe du couple $U = (X, Y)$?

- -

On peut considérer que l'univers est l'ensemble des parties à n éléments de l'intervalle $[\![1, N]\!]$, noté $\mathcal{P}_n\big([\![1, N]\!]\big)$, de cardinal $\dbinom{N}{n}$, muni de l'équiprobabilité.

a. L'événement $\{X \geq k\}$ coïncide donc avec l'ensemble $\mathcal{P}_n\big([\![k, N]\!]\big)$, de cardinal $\dbinom{N-k+1}{n}$, puisque cela signifie que l'on tire n boules dont les numéros sont compris entre k et N. Notons que cet ensemble est vide si $n > N-k+1$, mais ceci est cohérent avec la convention $\dbinom{n}{p} = 0$ lorsque $p > n$. On a donc $P(X \geq k) = \dfrac{\dbinom{N-k+1}{n}}{\dbinom{N}{n}}$. On en déduit

$$P(X = k) = P(X \geq k) - P(X \geq k+1) = \frac{\dbinom{N-k+1}{n} - \dbinom{N-k}{n}}{\dbinom{N}{n}} = \frac{\dbinom{N-k}{n-1}}{\dbinom{N}{n}}$$

en utilisant la formule de Pascal.

Remarque. On retrouve ce résultat plus simplement en voyant que l'événement $\{X = k\}$ correspond aux parties de cardinal n de l'intervalle $[\![1, N]\!]$ qui contiennent l'élément k et qui contiennent $n-1$ éléments distincts parmi les $N-k$ éléments de l'intervalle $[\![k+1, N]\!]$, il y a donc $\dbinom{N-k}{n-1}$ telles parties.

b. De façon analogue, $\{Y \leq l\} = \mathcal{P}_n\big([\![1, l]\!]\big)$, donc $P(Y \leq l) = \dfrac{\dbinom{l}{n}}{\dbinom{N}{n}}$, puis

$$P(Y = l) = P(Y \leq l) - P(Y \leq l-1) = \frac{\dbinom{l}{n} - \dbinom{l-1}{n}}{\dbinom{N}{n}} = \frac{\dbinom{l-1}{n-1}}{\dbinom{N}{n}}$$

c. L'événement $\big\{U = (k, l)\big\} = \{X = k\} \cap \{Y = l\}$ correspond aux parties de cardinal n de l'intervalle $[\![1, N]\!]$ contenant l'élément k, l'élément l, et $n-2$ éléments parmi les $l-k-1$ de l'intervalle $[\![k+1, l-1]\!]$. Il y a $\dbinom{l-k-1}{n-2}$ telles parties, donc

$$\forall (k,l) \in [\![1,N]\!]^2 \qquad P\big((X,Y)=(k,l)\big) = \frac{\dbinom{l-k-1}{n-2}}{\dbinom{N}{n}} \ .$$

---

11. On considère une suite de tirages avec remise dans une urne contenant N boules numérotées de 1 à N. Pour tout $n \geq 1$, on note X_n le nombre de numéros non encore sortis à l'issue du n-ème tirage.

a. Que dire de la variable aléatoire X_1 ?

b. Pour $n \geq 1$, quel est l'ensemble $X_n(\Omega)$?

c. En utilisant la formule des probabilités totales, montrer la relation, pour $n \geq 2$ et $k \in [\![0, N-1]\!]$,

$$P(X_n = k) = \frac{N-k}{N}\, P(X_{n-1} = k) + \frac{k+1}{N}\, P(X_{n-1} = k+1)\ .$$

d. En déduire que la suite $\big(\mathrm{E}(X_n)\big)_{n\geq 1}$ est géométrique, et donner l'expression de $\mathrm{E}(X_n)$.

- -

a. La variable aléatoire X_1 est constante, de valeur $N-1$.

b. On a $X_n(\Omega) = [\![N-n, N-1]\!]$ si $1 \leq n \leq N$, et $X_n(\Omega) = [\![0, N-1]\!]$ si $n \geq N$.

c. Pour $n \geq 2$ et $k \in [\![0, N-1]\!]$, on a

$$\{X_n = k\} = \big(\{X_n = k\} \cap \{X_{n-1} = k\}\big) \sqcup \big(\{X_n = k\} \cap \{X_{n-1} = k+1\}\big)$$

(réunion disjointe). On a donc

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

Or, $P(X_n = k | X_{n-1} = k) = \dfrac{N-k}{N}$ puisque la n-ème boule doit alors être tirée parmi les $N-k$ boules déjà tirées parmi les N boules de l'urne. Et $P(X_n = k | X_{n-1} = k+1) = \dfrac{k+1}{N}$ puisque la n-ème boule doit alors être tirée parmi les $k+1$ boules non encore obtenues parmi les N boules de l'urne. On obtient bien la relation

$$P(X_n = k) = \frac{N-k}{N}\, P(X_{n-1} = k) + \frac{k+1}{N}\, P(X_{n-1} = k+1)\ .$$

d. Pour $n \geq 2$, on a

$$\begin{aligned} \mathrm{E}(X_n) &= \sum_{k=0}^{N-1} k\, P(X_n = k) \ = \sum_{k=1}^{N-1} k\, P(X_n = k) \\ &= \sum_{k=1}^{N-1} k \frac{N-k}{N}\, P(X_{n-1} = k) + \sum_{k=0}^{N-2} k\, \frac{k+1}{N}\, P(X_{n-1} = k+1) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^{N-1} k\,P(X_{n-1}=k) - \frac{1}{N}\sum_{k=1}^{N-1} k^2\,P(X_{n-1}=k) + \sum_{k=1}^{N-1} \frac{(k-1)k}{N}\,P(X_{n-1}=k)\\
&= \sum_{k=1}^{N-1} k\,P(X_{n-1}=k) - \frac{1}{N}\sum_{k=1}^{N-1} k^2\,P(X_{n-1}=k) + \frac{1}{N}\sum_{k=1}^{N-1} k^2\,P(X_{n-1}=k) - \frac{1}{N}\sum_{k=1}^{N-1} k\,P(X_{n-1}=k)\\
&= \left(1-\frac{1}{N}\right)\sum_{k=1}^{N-1} k\,P(X_{n-1}=k)\\
&= \left(1-\frac{1}{N}\right)\mathrm{E}(X_{n-1})\,.
\end{aligned}$$

La suite $\big(\mathrm{E}(X_n)\big)_{n\geq 1}$ est donc géométrique, de raison $1-\dfrac{1}{N}$. Comme $\mathrm{E}(X_1)=N-1$, on déduit, pour $n\geq 1$,

$$\mathrm{E}(X_n) = (N-1)\left(1-\frac{1}{N}\right)^{n-1} = \frac{(N-1)^n}{N^{n-1}}\,.$$

---

12. Soient X et Y deux variables aléatoires à valeurs dans $[\![1,n+1]\!]$. On suppose que

$$\forall (i,j)\in[\![1,n+1]\!]^2 \qquad P\big(\{X=i\}\,\cap\,\{Y=j\}\big) = \frac{1}{2^{2n}}\binom{n}{i-1}\binom{n}{j-1}\,.$$

a. Déterminer la loi de X. Reconnaître la loi de $X-1$, en déduire $\mathrm{E}(X)$.

b. Les variables X et Y sont-elles indépendantes ?

c. Soit $M=(m_{i,j})\in\mathcal{M}_{n+1}(\mathbb{R})$, avec $m_{i,j}=P\big(\{X=i\}\,\cap\,\{Y=j\}\big)$. Calculer M^2, en déduire le spectre de M.

Indication. On pourra utiliser (et éventuellement prouver) la relation $\displaystyle\sum_{k=0}^{n}\binom{n}{k}^2=\binom{2n}{n}$.

- -

a. Si l'on fixe $i\in[\![1,n+1]\!]$, alors

$$P(X=i)=\sum_{j=1}^{n+1} a_{i,j} = \frac{1}{4^n}\binom{n}{i-1}\sum_{j=1}^{n+1}\binom{n}{j-1} = \frac{1}{2^n}\binom{n}{i-1}\,.$$

Soit $Z=X-1$, alors $Z(\Omega)=[\![1,n]\!]$ et

$$\forall k\in[\![1,n]\!] \qquad P(Z=k)=P(X=k+1)=\frac{1}{2^n}\binom{n}{k}\,.$$

On reconnaît une loi binomiale de paramètres n et $\dfrac{1}{2}$. Donc $\mathrm{E}(X-1)=\dfrac{n}{2}$ d'après le cours, puis $\mathrm{E}(X)=\dfrac{n}{2}+1$.

b. De façon symétrique, pour $j\in[\![1,n+1]\!]$, $P(Y=j)=\dfrac{1}{2^n}\dbinom{n}{j-1}$. Les variables X et Y ont la même loi. On voit immédiatement alors que

$$\forall (i,j) \in [\![1, n+1]\!]^2 \qquad m_{i,j} = P\big(\{X=i\} \cap \{Y=j\}\big) = P(X=i)\,P(Y=j)\ ,$$

donc les variables X et Y sont indépendantes.

c. Posons $M^2 = A = (a_{i,j})_{1\leq i,j\leq n+1}$. Alors $a_{i,k} = \displaystyle\sum_{j=1}^{n+1} m_{i,j}m_{j,k}$, soit

$$a_{i,k} = \frac{1}{16^n}\sum_{j=1}^{n+1}\binom{n}{i-1}\binom{n}{j-1}^2\binom{n}{k-1} = \frac{1}{16^n}\binom{n}{i-1}\binom{n}{k-1}\sum_{l=0}^{n}\binom{n}{l}^2 = \frac{\binom{2n}{n}}{16^n}\binom{n}{i-1}\binom{n}{k-1}\ .$$

En changeant de notations et en posant $K = \dfrac{\binom{2n}{n}}{4^n}$, on voit que $a_{i,j} = K\,m_{i,j}$, autrement dit $M^2 = K\,M$. Le polynôme $P = X^2 - KX = X(X-K)$ est donc annulateur de M, on en déduit que M est diagonalisable (*mais on pouvait aussi noter que M est symétrique réelle*) et $\mathrm{Sp}(M) \subset \{0, K\}$. Comme une matrice diagonalisable avec une seule valeur propre est scalaire (i.e. de la forme λI_{n+1}), ce qui n'est pas le cas ici, on déduit que les nombres 0 et $K = \dfrac{\binom{2n}{n}}{4^n}$ sont tous deux valeurs propres de M et en constituent le spectre.

Quelques remarques. La relation $\displaystyle\sum_{k=0}^{n}\binom{n}{k}^2 = \binom{2n}{n}$ s'obtient facilement en observant le coefficient de X^n dans le polynôme $(X+1)^{2n}$, que l'on peut écrire aussi $(X+1)^n(X+1)^n$.

L'écriture de $m_{i,j}$ sous la forme $m_{i,j} = c_ic_j$ avec $c_i = \dfrac{\binom{n}{i-1}}{2^n}$ montre que la matrice M est de rang 1 ; en effet, ses lignes sont toutes proportionnelles. La valeur propre 0 est donc de multiplicité $n-1$, et la valeur propre $K = \dfrac{\binom{2n}{n}}{4^n}$ (dont on peut remarquer que c'est la trace de M) est simple.