CORRIGÉ – CHAPITRE 21

DÉNOMBREMENT, PROBABILITÉS ET VARIABLES ALÉATOIRES

⋆ *application directe* ⋆⋆ *mise en œuvre* ⋆⋆⋆ *problème* ⋆⋆⋆⋆ *problème long, en autonomie*

Les renvois « proposition n », « théorème n » et les numéros de section 21.x désignent les énoncés du chapitre 21 du cours ; ceux des chapitres antérieurs sont précisés, de même que les exercices des feuilles précédentes. Pour dénombrer, on suit la méthode de la section 21.2.6 : on décrit précisément les objets à compter, on se demande si l'ordre compte et si les répétitions sont permises, on découpe en morceaux disjoints (proposition 2 (i)) ou l'on construit les objets par choix successifs, en vérifiant que le nombre de possibilités à chaque étape ne dépend pas des choix précédents (lemme des bergers, proposition 2 (v)). Le plus sûr est presque toujours d'exhiber une *bijection* entre l'ensemble à compter et un ensemble dont on connaît le cardinal (proposition 1 (i)). Lorsqu'un énoncé parle de hasard sans préciser le modèle, le corrigé commence par fixer l'univers et la probabilité. On note, comme dans la feuille, $\mathrm{P}_B(A)$ la probabilité de A sachant un événement B, et $\mathrm{P}(A \mid B_1 \cap B_2)$ lorsque l'événement conditionnant est une intersection.

Dénombrement

EXERCICE 1 ⋆

Un élément de E est déterminé par la liste de ses sept chiffres, lus de gauche à droite, et toute 7-liste d'éléments de $[\![1\,,9]\!]$ est la liste des chiffres d'un unique élément de E : comme le chiffre 0 est exclu, aucun nombre ne commence par un zéro « parasite ». L'application qui associe à un nombre la liste de ses chiffres est donc une bijection de E sur $[\![1\,,9]\!]^7$, et la proposition 2 (iv) donne

$$\boxed{|E| = 9^7 = 4\,782\,969}$$

(a) Par la même bijection, E_1 correspond aux 7-listes d'éléments distincts de $[\![1\,,9]\!]$, dont le théorème 1 donne le nombre :

$$\boxed{|E_1| = A_9^7 = 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 = \frac{9!}{2!} = 181\,440}$$

(b) Un nombre est pair si et seulement si son chiffre des unités l'est, c'est-à-dire appartient à $\{2, 4, 6, 8\}$. Les éléments de E_2 correspondent aux listes de $[\![1\,,9]\!]^6 \times \{2, 4, 6, 8\}$:

$$\boxed{|E_2| = 9^6 \times 4 = 2\,125\,764}$$

(c) Une liste strictement croissante est formée de sept chiffres distincts, et elle est déterminée par l'*ensemble* de ses chiffres : il n'y a qu'une façon de ranger sept chiffres distincts dans l'ordre croissant. L'application qui associe à un élément de E_3 l'ensemble de ses chiffres est donc une bijection de E_3 sur l'ensemble des parties à sept éléments de $[\![1\,,9]\!]$, et le théorème 2 donne

$$\boxed{|E_3| = \binom{9}{7} = \binom{9}{2} = 36}$$

À retenir. Les questions (a) et (c) comptent les mêmes chiffres, distincts, mais l'ordre est libre dans la première et imposé dans la seconde : on passe de l'une à l'autre en divisant par $7! = 5\,040$, puisque $\dfrac{181\,440}{5\,040} = 36$. C'est exactement la démonstration du théorème 2. Une énumération exhaustive des 9^7 listes confirme les quatre résultats.

EXERCICE 2 ⋆

1. Numérotons les boulangeries de 1 à 4 et les jours de la semaine de 1 à 7. Un choix des jours de fermeture est une application $f : [\![1\,,4]\!] \to [\![1\,,7]\!]$, où $f(i)$ est le jour de fermeture de la boulangerie i. La proposition 3 donne 7^4 choix. Lorsque deux boulangeries ne ferment jamais le même jour, f est injective, et le théorème 1 donne A_7^4 choix.

$$\boxed{7^4 = 2\,401 \text{ choix sans contrainte,} \qquad A_7^4 = 7 \times 6 \times 5 \times 4 = 840 \text{ choix de jours distincts}}$$

Pour la dernière condition, on passe au complémentaire, comme le recommande la méthode de la section 21.2.6. Un jour donné, les quatre boulangeries sont fermées si et seulement si elles ferment toutes ce jour-là : la condition échoue exactement pour les 7 applications constantes.

$$\boxed{7^4 - 7 = 2\,394 \text{ choix laissant chaque jour une boulangerie ouverte}}$$

Piège. L'erreur la plus répandue est de lire « chaque jour, au moins une boulangerie ouverte » comme « chaque jour, au moins une boulangerie fermée », ce qui demanderait que f soit surjective, impossible avec quatre boulangeries et sept jours. La condition porte sur les boulangeries *ouvertes*, et elle n'exclut que les fermetures simultanées des quatre.

2. Un choix de boissons pour n matins est une n-liste d'éléments de $\{T, C\}$, où T désigne le thé et C le café.

(a) Sans contrainte, la proposition 3 donne le nombre de listes :

$$\boxed{2^n \text{ choix}}$$

(b) On retire les deux listes constantes, celle qui ne contient que T et celle qui ne contient que C :

$$\boxed{2^n - 2 \text{ choix}}$$

(c) Le nombre de matins avec thé est égal à celui des matins avec café, donc à $\dfrac{n}{2}$: c'est impossible si n est impair. Si n est pair, une liste convenable est déterminée par l'ensemble des matins où l'on boit du thé, qui est une partie à $\dfrac{n}{2}$ éléments de $[\![1\,,n]\!]$, et toute partie de ce cardinal convient.

$$\boxed{\binom{n}{n/2} \text{ choix si } n \text{ est pair,} \qquad 0 \text{ si } n \text{ est impair}}$$

(d) La boisson du premier matin se choisit de deux façons, et chaque boisson suivante est imposée, puisqu'elle doit différer de la précédente : les seules listes possibles sont $TCTC\cdots$ et $CTCT\cdots$.

$$\boxed{2 \text{ choix}}$$

Les résultats ont été contrôlés par énumération pour $n \leqslant 8$. Pour $n = 4$, par exemple, on trouve 16, 14, $\binom{4}{2} = 6$ et 2 choix.

EXERCICE 3 ⋆

1. Les cinq lettres de MATHS sont distinctes : une anagramme est une façon de les ordonner, c'est-à-dire une permutation de ces cinq lettres, et il y en a $5! = 120$ (théorème 1). Pour RIRE, on choisit les deux positions du R parmi les quatre, puis on place I et E dans les deux positions restantes, de 2 façons ; pour ANANAS, on choisit les trois positions des A, puis les deux positions des N parmi les trois restantes, le S occupant la dernière.

$$\boxed{\text{MATHS} : 5! = 120, \qquad \text{RIRE} : \binom{4}{2} \times 2 = 12, \qquad \text{ANANAS} : \binom{6}{3}\binom{3}{2} = 60}$$

2. Notons L l'ensemble des anagrammes du mot. On rend les n lettres discernables en indiçant les exemplaires de chaque lettre : la i-ième lettre ℓ_i devient $\ell_{i,1}$, …, ℓ_{i,k_i}. On obtient n symboles distincts, et l'ensemble Λ des façons de les ranger en un mot de longueur n a $n!$ éléments (théorème 1). L'application $\pi : \Lambda \to L$ qui efface les indices est surjective. Fixons une anagramme $w \in L$: la lettre ℓ_i y occupe k_i positions, et un antécédent de w par π est déterminé par la façon de distribuer les indices $1, \ldots, k_i$ dans ces positions, pour chaque i, soit une permutation de $[\![1\,, k_i]\!]$ pour chaque i. La fibre $\pi^{-1}(\{w\})$ a donc $k_1! \cdots k_r!$ éléments (proposition 2 (iv) et théorème 1), quel que soit w, et le lemme des bergers donne $n! = |L| \times k_1! \cdots k_r!$.

$$\boxed{|L| = \frac{n!}{k_1!\, k_2! \cdots k_r!}}$$

On peut aussi choisir successivement les positions des lettres : $\binom{n}{k_1}$ choix pour la première, puis $\binom{n-k_1}{k_2}$ pour la deuxième, et ainsi de suite ; le produit se simplifie en cascade et donne le même nombre. Avec MATHS, RIRE et ANANAS, on retrouve $\dfrac{5!}{1!^5} = 120$, $\dfrac{4!}{2!\,1!\,1!} = 12$ et $\dfrac{6!}{3!\,2!\,1!} = 60$ ✓.

L'exercice 37 de la feuille du chapitre 3 a montré que le coefficient de $x^p y^q z^r$ dans $(x+y+z)^n$ est $\dfrac{n!}{p!\,q!\,r!}$: c'est le nombre d'anagrammes d'un mot formé de p lettres x, q lettres y et r lettres z. Ce n'est pas une coïncidence : en développant le produit des n facteurs $(x+y+z)$, on choisit une lettre dans chaque facteur, c'est-à-dire un mot de longueur n, et le terme $x^p y^q z^r$ apparaît une fois pour chaque mot qui contient p fois x, q fois y et r fois z. C'est la démonstration combinatoire de la formule du binôme (théorème 3), étendue à trois termes.

EXERCICE 4 ⋆

Chaque étudiant a exactement une langue et exactement une option : les six ensembles d'étudiants définis par un couple (langue, option) sont deux à deux disjoints et recouvrent le groupe. On note a_I, a_C, a_A les effectifs des étudiants d'anglais qui suivent l'informatique, la chimie et l'astronomie, et g_I, g_C, g_A ceux des étudiants d'allemand. Par la proposition 2 (i), l'effectif d'une option est la somme des effectifs de ses deux cases, et celui d'une langue la somme de ses trois cases. Les données s'écrivent

$$a_\mathrm{A} = 8, \quad a_\mathrm{I} = 3, \quad g_\mathrm{C} = 6, \quad a_\mathrm{A} + g_\mathrm{A} = 12, \quad a_\mathrm{C} + g_\mathrm{C} = 15, \quad g_\mathrm{I} + g_\mathrm{C} + g_\mathrm{A} = 16,$$

d'où successivement $g_\mathrm{A} = 12 - 8 = 4$, $a_\mathrm{C} = 15 - 6 = 9$ et $g_\mathrm{I} = 16 - 6 - 4 = 6$.

	Informatique	Chimie	Astronomie	Total
Anglais	3	9	8	20
Allemand	6	6	4	16
Total	9	15	12	**36**

On lit dans la dernière case l'effectif du groupe, obtenu en sommant indifféremment la dernière ligne ou la dernière colonne.

$$\boxed{\text{Le groupe compte 36 étudiants, dont 20 étudient l'anglais}}$$

Piège. Sans l'hypothèse « exactement une langue et exactement une option », le problème n'a pas de solution unique : un étudiant qui suivrait deux options serait compté deux fois dans les effectifs des options. C'est l'hypothèse de partition qui autorise l'addition des cases.

EXERCICE 5 ⋆

On note N_k le nombre de parties de $[\![1\,,n]\!]$ de diamètre k. Le diamètre d'une partie non vide est compris entre 0 et $n-1$, donc $N_k = 0$ pour $k \geqslant n$. Une partie de diamètre 0 a son minimum égal à son maximum : c'est un singleton, et $N_0 = n$.

Soit $k \in [\![1\,,n-1]\!]$. Une partie A de diamètre k est déterminée par son minimum a et par l'ensemble $A \cap [\![a+1\,,a+k-1]\!]$ de ses éléments compris strictement entre a et $a+k$. Réciproquement, pour $a \in [\![1\,,n-k]\!]$ et B une partie quelconque de l'intervalle $[\![a+1\,,a+k-1]\!]$, qui a $k-1$ éléments, la partie $\{a\} \cup B \cup \{a+k\}$ est contenue dans $[\![1\,,n]\!]$ et a pour diamètre k. On a ainsi une bijection entre les parties de diamètre k et les couples (a, B) ; il y a $n-k$ choix de a et, pour chacun, 2^{k-1} choix de B (proposition 3).

$$\boxed{N_0 = n, \qquad N_k = (n-k)\,2^{k-1} \text{ pour } 1 \leqslant k \leqslant n-1, \qquad N_k = 0 \text{ pour } k \geqslant n}$$

Pour la vérification, on écrit $n-k = \displaystyle\sum_{j=k}^{n-1} 1$ et l'on intervertit les sommes (chapitre 3, proposition 8) :

$$\sum_{k=1}^{n-1}(n-k)2^{k-1} = \sum_{j=1}^{n-1}\sum_{k=1}^{j} 2^{k-1} = \sum_{j=1}^{n-1}(2^j-1) = (2^n-2)-(n-1) = 2^n-n-1.$$

$$\boxed{\sum_{k\geqslant 0} N_k = n + 2^n - n - 1 = 2^n - 1}$$

C'est bien le nombre de parties non vides de $[\![1\,,n]\!]$ (proposition 3) ✓ : les parties non vides se répartissent selon leur diamètre en morceaux disjoints. Lu dans l'autre sens, ce dénombrement est une démonstration combinatoire de l'identité $\displaystyle\sum_{k=1}^{n-1}(n-k)2^{k-1} = 2^n-n-1$.

EXERCICE 6 ⋆

Une main est une partie à cinq éléments du jeu de 32 cartes ; il y en a $\binom{32}{5} = 201\,376$ (théorème 2). Pour chaque question, on découpe l'ensemble des mains à compter en morceaux disjoints, et l'on compte chaque morceau en choisissant séparément les cartes de chaque catégorie, ce que justifie la proposition 2 (iv). Dans le jeu, une couleur compte 8 cartes, une hauteur 4 cartes.

(a) Les mains de cinq cœurs et celles de cinq trèfles forment deux ensembles disjoints, de $\binom{8}{5}$ éléments chacun.

$$\boxed{2\binom{8}{5} = 112}$$

(b) On choisit deux cœurs parmi huit, puis trois trèfles parmi huit.

$$\boxed{\binom{8}{2}\binom{8}{3} = 28 \times 56 = 1\,568}$$

(c) On sépare les mains sans as, qui sont des parties à cinq éléments des 28 cartes qui ne sont pas des as, et celles qui ont exactement un as, choisi de 4 façons, avec quatre autres cartes prises parmi les 28.

$$\boxed{\binom{28}{5} + 4\binom{28}{4} = 98\,280 + 81\,900 = 180\,180}$$

(d) L'as de cœur est à la fois un as et un cœur : on distingue selon qu'il figure dans la main. On note qu'il y a 3 as qui ne sont pas des cœurs, 7 cœurs qui ne sont pas des as, et $32 - 4 - 7 = 21$ cartes qui ne sont ni l'un ni l'autre.

— Si l'as de cœur figure dans la main, il faut un autre as, parmi 3, deux autres cœurs, parmi 7, et une dernière carte, ni as ni cœur, parmi 21 : $3 \times \binom{7}{2} \times 21 = 1\,323$ mains.

— Sinon, il faut deux as parmi 3 et trois cœurs parmi 7, ce qui fait déjà cinq cartes : $\binom{3}{2}\binom{7}{3} = 105$ mains.

$$\boxed{1\,323 + 105 = 1\,428}$$

(e) On procède de même avec l'as de carreau ; il y a 7 carreaux qui ne sont pas des as et 21 cartes qui ne sont ni des as ni des carreaux.

— Si l'as de carreau figure dans la main, c'est le seul as, et il faut au moins deux autres carreaux parmi les quatre autres cartes : $\sum_{j=2}^{4} \binom{7}{j}\binom{21}{4-j} = 4\,410 + 735 + 35 = 5\,180$ mains.

— Sinon, on choisit l'as parmi les 3 autres, puis au moins trois carreaux parmi les quatre autres cartes : $3\Big(\binom{7}{3} \times 21 + \binom{7}{4}\Big) = 3 \times 770 = 2\,310$ mains.

$$\boxed{5\,180 + 2\,310 = 7\,490}$$

(f) Le roi de cœur est à la fois un roi et un cœur. Il y a 3 autres rois, 7 cœurs qui ne sont pas des rois, et 21 cartes qui ne sont ni des rois ni des cœurs.

— Si le roi de cœur figure dans la main, aucun autre cœur n'est permis, et il faut au moins un autre roi : $\sum_{j=1}^{3} \binom{3}{j}\binom{21}{4-j} = 3\,990 + 630 + 21 = 4\,641$ mains.

— Sinon, il faut $j \in \{2, 3\}$ rois parmi 3 et $h \in \{0, 1\}$ cœur parmi 7, le reste parmi 21, soit

$$\sum_{j=2}^{3} \sum_{h=0}^{1} \binom{3}{j}\binom{7}{h}\binom{21}{5-j-h} = 3 \times 2\,800 + 357 = 8\,757 \text{ mains.}$$

$$\boxed{4\,641 + 8\,757 = 13\,398}$$

Une énumération exhaustive des 201 376 mains confirme les six résultats.

Piège. L'erreur la plus répandue, en (d), est de multiplier $\binom{4}{2}$ par $\binom{8}{3}$, en choisissant les as parmi les quatre et les cœurs parmi les huit : l'as de cœur peut alors être choisi deux fois, une fois comme as et une fois comme cœur, et l'on obtient des « mains » de quatre cartes, tandis que les vraies mains contenant l'as de cœur, qui ont une cinquième carte ni as ni cœur, sont oubliées. Le nombre $6 \times 56 = 336$ ainsi obtenu ne compte rien. Le découpage selon la carte « double » rend les catégories disjointes.

EXERCICE 7 ⋆⋆

1. On compte le complémentaire : une main dont les cinq cartes ont des hauteurs deux à deux distinctes est déterminée par l'ensemble de ses cinq hauteurs, choisi parmi 8, et par la couleur de chacune des cinq cartes, soit 4^5 choix (proposition 3).

$$\boxed{\binom{32}{5} - \binom{8}{5}4^5 = 201\,376 - 57\,344 = 144\,032}$$

Pour la seconde question, on regroupe les cartes d'une main par hauteur : les effectifs des hauteurs présentes forment l'une des répartitions $1+1+1+1+1$, $2+1+1+1$, $2+2+1$, $3+1+1$, $3+2$ ou $4+1$. Une main contient deux paires de hauteurs différentes exactement pour les répartitions $2+2+1$ et $3+2$.

— Répartition $2+2+1$: on choisit l'*ensemble* des deux hauteurs des paires, de $\binom{8}{2}$ façons, deux cartes de chacune, de $\binom{4}{2}^2$ façons, puis la cinquième carte parmi les $32-8=24$ cartes des autres hauteurs, soit $28\times 36\times 24 = 24\,192$ mains.

— Répartition $3+2$: ce sont les fulls, au nombre de $8\times\binom{4}{3}\times 7\times\binom{4}{2} = 1\,344$ (exemple de la section 21.2.6).

$$\boxed{24\,192 + 1\,344 = 25\,536}$$

Piège. Pour les deux paires, l'erreur la plus répandue est de choisir « la hauteur de la première paire, puis celle de la seconde », de 8×7 façons : les deux paires jouent le même rôle, et chaque main est alors comptée deux fois. Pour un full, au contraire, le brelan et la paire jouent des rôles différents, et le produit 8×7 est juste.

2. Il y a $\binom{52}{5} = 2\,598\,960$ mains. On numérote les hauteurs de 1 (as) à 13 (roi). Les quintes ont pour ensembles de hauteurs les dix ensembles $\{h, h+1, \dots, h+4\}$ pour $h \in [\![1\,,9]\!]$, et $\{10, 11, 12, 13, 1\}$.

(a) Une quinte flush est déterminée par sa suite de hauteurs, parmi 10, et sa couleur, parmi 4 :

$$\boxed{10\times 4 = 40 \text{ quintes flush}}$$

(b) Un carré est déterminé par la hauteur des quatre cartes, parmi 13, et la cinquième carte, parmi 48 :

$$\boxed{13\times 48 = 624 \text{ carrés}}$$

(c) On choisit la hauteur du brelan et ses trois cartes, puis la hauteur de la paire, différente, et ses deux cartes :

$$\boxed{13\binom{4}{3}\times 12\binom{4}{2} = 13\times 4\times 12\times 6 = 3\,744 \text{ fulls}}$$

(d) Une quinte est déterminée par sa suite de hauteurs et par la couleur de chacune des cinq cartes, soit 10×4^5 mains, dont 40 quintes flush :

$$\boxed{10(4^5-4) = 10\,200 \text{ quintes qui ne sont pas des quintes flush}}$$

(e) On choisit la hauteur du brelan, ses trois cartes, puis l'*ensemble* des deux autres hauteurs, de $\binom{12}{2}$ façons, et une carte de chacune :

$$\boxed{13\times 4\times\binom{12}{2}\times 4^2 = 54\,912 \text{ brelans}}$$

Les cinq nombres ont été confirmés par une énumération exhaustive des 2 598 960 mains. Rapportés au nombre total de mains, ils donnent la hiérarchie du poker : une quinte flush apparaît environ une fois sur 65 000 donnes, un carré une fois sur 4 165, un brelan une fois sur 47 ; plus une combinaison est rare, plus elle est forte.

EXERCICE 8 ⋆⋆

1. Une donne est une suite (N, S, E, O) de quatre parties à 13 éléments du jeu, deux à deux disjointes. On choisit la main de Nord parmi les 52 cartes, puis celle de Sud parmi les 39 restantes, puis celle d'Est parmi les 26 restantes, et Ouest reçoit les 13 dernières ; le nombre de choix à chaque étape ne dépend pas des choix précédents.

$$\boxed{\binom{52}{13}\binom{39}{13}\binom{26}{13} = \frac{52!}{(13!)^4} \approx 5{,}36 \cdot 10^{28} \text{ donnes}}$$

C'est le nombre d'anagrammes d'un mot formé de treize lettres N, treize S, treize E et treize O (exercice 3) : on écrit, sous chacune des 52 cartes rangées dans un ordre fixé, l'initiale du joueur qui la reçoit.

2. Si chaque joueur n'a que des cartes d'une couleur, il a les treize cartes de cette couleur, et les quatre joueurs ont des couleurs deux à deux distinctes. Une telle donne est donc déterminée par une bijection de l'ensemble des joueurs sur l'ensemble des couleurs, et réciproquement toute bijection en fournit une.

$$\boxed{4! = 24 \text{ donnes}}$$

3. Disons qu'un joueur est *unicolore* s'il n'a que des cartes d'une couleur ; il possède alors toute cette couleur. Pour une donne où exactement deux joueurs sont unicolores, on choisit ces deux joueurs, de $\binom{4}{2} = 6$ façons, puis leurs couleurs, distinctes, de $4 \times 3 = 12$ façons. Il reste 26 cartes, de deux couleurs, à partager en deux mains de 13 entre les deux autres joueurs, de $\binom{26}{13}$ façons, dont exactement deux rendent ces joueurs unicolores : celles où l'un reçoit la première couleur entière et l'autre la seconde.

$$\boxed{6 \times 12 \times \left(\binom{26}{13} - 2\right) = 748\,843\,056 \text{ donnes}}$$

4. Notons $\mathcal{N}$ l'ensemble des donnes où Nord et Sud sont tous deux unicolores, et $\mathcal{E}$ celui où Est et Ouest le sont. Comme à la question 3, $|\mathcal{N}| = 12\binom{26}{13}$: on choisit les couleurs de Nord et de Sud, puis la main d'Est parmi les 26 cartes restantes, sans contrainte. De même $|\mathcal{E}| = 12\binom{26}{13}$, et $\mathcal{N} \cap \mathcal{E}$ est l'ensemble des 24 donnes de la question 2. La proposition 2 (iii) donne

$$\boxed{|\mathcal{N} \cup \mathcal{E}| = 2 \times 12\binom{26}{13} - 24 = 249\,614\,376 \text{ donnes}}$$

Ces formules ont été contrôlées par une énumération exhaustive sur un jeu réduit de quatre couleurs de 2 ou de 3 cartes, distribuées à quatre joueurs. Pour des couleurs de 2 cartes, par exemple, on trouve 2 520 donnes, dont 24, 288 et 120 satisfont les conditions des questions 2, 3 et 4, comme le prévoient les formules avec $\binom{4}{2} = 6$ au lieu de $\binom{26}{13}$.

Piège. L'erreur la plus répandue, à la question 4, est d'additionner $|\mathcal{N}|$ et $|\mathcal{E}|$: les 24 donnes entièrement unicolores sont comptées deux fois. À la question 3, c'est d'oublier de retirer les deux partages unicolores des 26 dernières cartes, qui donneraient quatre joueurs unicolores et non deux.

EXERCICE 9 ⋆⋆

1. Les mots de longueur 1 sont a et b, et ceux de longueur 2 sans facteur aa sont ab, ba et bb : $u_1 = 2$ et $u_2 = 3$. Soit $n \geqslant 1$, et w un mot de longueur $n+2$ sans facteur aa. On distingue selon sa dernière lettre.

— Si w se termine par b, alors $w = w'b$, où w' est un mot de longueur $n+1$ sans facteur aa ; réciproquement, pour tout tel mot w', le mot $w'b$ n'a pas de facteur aa, puisqu'un tel facteur se trouverait dans w'. Il y a u_{n+1} tels mots.

— Si w se termine par a, son avant-dernière lettre est b, et $w = w''ba$, où w'' est un mot de longueur n sans facteur aa ; réciproquement, tout mot $w''ba$ de ce type convient. Il y a u_n tels mots.

Les deux cas sont disjoints, et la proposition 2 (i) donne

$$\boxed{u_{n+2} = u_{n+1} + u_n \text{ pour tout } n \geqslant 1}$$

Comme $u_1 = 2 = F_3$ et $u_2 = 3 = F_4$, une récurrence double montre que $u_n = F_{n+2}$ pour tout $n \geqslant 1$: si $u_n = F_{n+2}$ et $u_{n+1} = F_{n+3}$, alors $u_{n+2} = F_{n+3} + F_{n+2} = F_{n+4}$.

$$\boxed{u_n = F_{n+2} \text{ pour tout } n \geqslant 1}$$

La formule vaut aussi pour $n = 0$, si l'on convient que le mot vide compte : $u_0 = 1 = F_2$.

2. Un mot $w = w_1 \cdots w_n$ n'a pas de facteur ab si et seulement si aucune lettre a n'est suivie, plus loin, d'une lettre b. En effet, s'il existe $i < j$ tels que $w_i = a$ et $w_j = b$, notons j' le plus petit indice $j' > i$ tel que $w_{j'} = b$: les lettres w_i, …, $w_{j'-1}$ sont des a, et $w_{j'-1}w_{j'} = ab$ est un facteur. Les mots sans facteur ab sont donc les mots $b^k a^{n-k}$, formés de k lettres b suivies de $n-k$ lettres a, avec $k \in [\![0\,,n]\!]$, et ils sont deux à deux distincts.

$$\boxed{n+1 \text{ mots sans facteur } ab}$$

Pour aab, on sépare les mots qui n'ont pas de facteur aa, au nombre de $u_n = F_{n+2}$, et ceux qui en ont un. Soit w un mot sans facteur aab qui contient le facteur aa, et i la position de la première occurrence de ce facteur. Toutes les lettres qui suivent sont des a : sinon, la première lettre b située après la position i serait précédée de deux a au moins. Le mot w s'écrit donc a^n, ou bien vba^m avec $m \in [\![2\,,n-1]\!]$ et v un mot de longueur $n-1-m$ sans facteur aa ; réciproquement, ces mots n'ont pas de facteur aab et contiennent aa. Pour $n \geqslant 2$, il y en a donc, en posant $j = n-1-m$ et avec $u_0 = F_2 = 1$,

$$1 + \sum_{m=2}^{n-1} u_{n-1-m} = 1 + \sum_{j=0}^{n-3} F_{j+2} = 1 + (F_{n+1} - 2) = F_{n+1} - 1,$$

puisque $\displaystyle\sum_{i=0}^{N} F_i = F_{N+2} - 1$: les égalités $F_i = F_{i+2} - F_{i+1}$ donnent une somme télescopique (chapitre 3, proposition 3), égale à $F_{N+2} - F_1$. Au total, $F_{n+2} + F_{n+1} - 1 = F_{n+3} - 1$, valeur qui reste juste pour $n = 1$, où les deux mots conviennent.

$$\boxed{F_{n+3} - 1 \text{ mots sans facteur } aab}$$

Les trois formules ont été vérifiées par énumération pour $n \leqslant 14$. Pour $n = 4$, on trouve 8, 5 et 12 mots.

3. À une partie A de $[\![1\,,n]\!]$, on associe le mot $w_1 \cdots w_n$ défini par $w_i = a$ si $i \in A$ et $w_i = b$ sinon : c'est l'indicatrice de A, écrite avec les lettres a et b, et cette application est une bijection de $\mathcal{P}([\![1\,,n]\!])$ sur $\{a,b\}^n$ (démonstration de la proposition 3). La partie A contient deux entiers consécutifs i et $i+1$ si et seulement si $w_i w_{i+1} = aa$. La bijection envoie donc les parties cherchées sur les mots sans facteur aa.

$$\boxed{F_{n+2} \text{ parties de } [\![1\,,n]\!] \text{ sans deux entiers consécutifs}}$$

4. La fonction suivante calcule u_n par la récurrence de la question 1, à partir de $u_0 = 1$ et $u_1 = 2$.

```python
def u(n):
    a, b = 1, 2            # a = u_k, b = u_(k+1), pour k = 0
    for _ in range(n):     # on passe de k a k + 1
        a, b = b, a + b
    return a

print(u(20))               # affiche 17711
```

Les résultats d'un candidat forment un mot de longueur 20 sur l'alphabet $\{0,1\}$, et la condition « jamais deux bonnes réponses consécutives » signifie que 11 n'en est pas un facteur : il y a $u_{20} = F_{22} = 17\,711$ listes de résultats possibles. Si les 17 712 candidats avaient des résultats deux à deux distincts, l'application qui associe à chaque candidat sa liste serait une injection d'un ensemble à 17 712 éléments dans un ensemble à 17 711 éléments, ce qu'interdit le principe des tiroirs (chapitre 2, théorème 2).

$$\boxed{\text{Oui : deux candidats au moins ont exactement les mêmes résultats}}$$

À retenir. Les mots sans facteur *aa* reviennent sous de nombreux déguisements : parties sans entiers consécutifs, pavages d'une bande par des dominos et des carrés, montées d'un escalier une ou deux marches à la fois. Chaque fois, la décomposition selon la fin de l'objet donne la récurrence de Fibonacci. C'est cette décomposition, et non la formule, qu'il faut retenir.

EXERCICE 10 ⋆⋆

La méthode est celle de la démonstration de la proposition 3 : à un couple (A,B) de parties de E, on associe l'application $x \mapsto (\mathbf{1}_A(x), \mathbf{1}_B(x))$ de E dans $\{0,1\}^2$. C'est une bijection de $\mathcal{P}(E)^2$ sur l'ensemble des applications de E dans $\{0,1\}^2$, puisqu'une partie est déterminée par son indicatrice. Une condition sur le couple (A,B) qui se lit élément par élément devient une contrainte sur la valeur prise en chaque point x, et l'on compte les applications qui la respectent.

1. On a $A \cap B = C$ si et seulement si, pour tout $x \in C$, x appartient à A et à B, et, pour tout $x \notin C$, x n'appartient pas aux deux. La valeur en x est donc imposée, égale à $(1,1)$, pour $x \in C$, et choisie librement parmi les trois couples $(0,0)$, $(1,0)$, $(0,1)$ pour $x \in E \setminus C$. Par la proposition 3, appliquée à l'ensemble $E \setminus C$, qui a $n - |C|$ éléments,

$$\boxed{|\{(A,B) \in \mathcal{P}(E)^2 \mid A \cap B = C\}| = 3^{n-|C|}}$$

Avec $C = \varnothing$, on retrouve les 3^n couples de parties disjointes de l'exercice 39 de la feuille du chapitre 1 ✓.

2. On associe de même à un triplet (X,Y,Z) de parties de E l'application

$$x \mapsto (\mathbf{1}_X(x), \mathbf{1}_Y(x), \mathbf{1}_Z(x)),$$

qui définit une bijection de $\mathcal{P}(E)^3$ sur l'ensemble des applications de E dans $\{0,1\}^3$. On a $X \subset Y \subset Z$ si et seulement si la valeur en chaque point est l'un des quatre triplets $(0,0,0)$, $(0,0,1)$, $(0,1,1)$ et $(1,1,1)$: chaque élément de E est hors de Z, dans $Z \setminus Y$, dans $Y \setminus X$ ou dans X.

$$\boxed{4^n \text{ triplets } (X,Y,Z) \text{ tels que } X \subset Y \subset Z}$$

Avec deux parties, le même raisonnement donne les 3^n couples tels que $X \subset Y$ de l'exercice 39 du chapitre 1 ; avec k parties emboîtées, on trouverait $(k+1)^n$ suites.

3. Les couples (X,Y) tels que $X \cap Y$ soit un singleton se répartissent, selon la valeur $\{c\}$ de ce singleton, en n ensembles deux à deux disjoints, dont chacun a 3^{n-1} éléments par la question 1.

$$\boxed{n\,3^{n-1} \text{ couples } (X,Y) \text{ tels que } X \cap Y \text{ soit un singleton}}$$

Les trois résultats ont été vérifiés par énumération pour $n \leqslant 5$.

EXERCICE 11 ⋆⋆

Les quatre sommes se calculent par double comptage : on écrit le cardinal comme une somme d'indicatrices, $|A| = \sum_{x\in E} \mathbf{1}_A(x)$, et l'on intervertit les sommes (chapitre 3, proposition 6). Pour un élément x fixé, il reste à compter les parties, ou les couples de parties, qui le contiennent.

Première somme. Pour $x \in E$, les parties qui contiennent x s'écrivent $\{x\} \cup A'$ avec $A' \subset E \setminus \{x\}$: il y en a 2^{n-1}. Ainsi

$$\sum_{A\in\mathcal{P}(E)} |A| = \sum_{A} \sum_{x\in E} \mathbf{1}_A(x) = \sum_{x\in E} |\{A \mid x \in A\}| = \sum_{x\in E} 2^{n-1}, \qquad \boxed{\sum_{A} |A| = n\,2^{n-1}}$$

On a compté de deux façons l'ensemble des couples (x, A) tels que $x \in A$. On peut aussi regrouper les parties selon leur cardinal, et l'on retrouve l'identité $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$ (chapitre 3, proposition 10).

Deuxième somme. Pour $x \in E$, les couples (A, B) tels que $x \in A \cap B$ sont formés de deux parties qui contiennent x : il y en a $2^{n-1} \times 2^{n-1} = 4^{n-1}$.

$$\boxed{\sum_{(A,B)} |A \cap B| = n\,4^{n-1}}$$

Troisième somme. Pour $x \in E$, les couples tels que $x \notin A \cup B$ sont formés de deux parties de $E \setminus \{x\}$: il y en a 4^{n-1}, et il y a donc $4^n - 4^{n-1} = 3 \times 4^{n-1}$ couples tels que $x \in A \cup B$.

$$\boxed{\sum_{(A,B)} |A \cup B| = 3n\,4^{n-1}}$$

On contrôle par la proposition 2 (iii), puisque $\sum_{(A,B)} |A| = 2^n \times n2^{n-1} = 2n4^{n-1}$:

$$\sum_{(A,B)} |A \cup B| = \sum_{(A,B)} |A| + \sum_{(A,B)} |B| - \sum_{(A,B)} |A \cap B| = 2 \times 2n4^{n-1} - n4^{n-1} = 3n4^{n-1} \ \checkmark$$

Quatrième somme. Pour X et Y disjoints, $|X| + |Y| = |X \cup Y|$ (proposition 2 (i)). Pour $x \in E$, les couples de parties disjointes tels que $x \in X \cup Y$ s'obtiennent en plaçant x dans X ou dans Y, de deux façons, puis en choisissant un couple de parties disjointes de $E \setminus \{x\}$, de 3^{n-1} façons (exercice 10).

$$\boxed{\sum_{X\cap Y=\varnothing} (|X| + |Y|) = 2n\,3^{n-1}}$$

Les quatre formules ont été vérifiées par énumération pour $n \leqslant 5$. Pour $n = 2$, par exemple, la première somme vaut $0 + 1 + 1 + 2 = 4 = 2 \times 2^1$ ✓.

À retenir. Le double comptage transforme une somme de cardinaux, difficile à organiser, en une somme sur les éléments, où tous les termes sont égaux par symétrie. C'est l'ancêtre de la méthode des indicatrices de la section 21.6.3 : si A est une partie de E choisie au hasard sous la probabilité uniforme, la première somme montre que $|A|$ vaut $\dfrac{n}{2}$ en moyenne.

EXERCICE 12 ⋆⋆

1. Une application strictement croissante $f : [\![1\,,p]\!] \to [\![1\,,n]\!]$ est injective, et son image $f([\![1\,,p]\!])$ est une partie à p éléments de $[\![1\,,n]\!]$. Réciproquement, une partie à p éléments $\{y_1 < \cdots < y_p\}$ est l'image d'une unique application strictement croissante, $k \mapsto y_k$: les valeurs d'une telle application sont nécessairement les éléments de son image rangés dans l'ordre croissant. L'application $f \mapsto f([\![1\,,p]\!])$ est donc une bijection, et le théorème 2 donne

$$\boxed{\binom{n}{p} \text{ applications strictement croissantes de } [\![1\,,p]\!] \text{ dans } [\![1\,,n]\!]}$$

Ce nombre est nul si $p > n$, ce qui est cohérent : il n'y a alors aucune injection (théorème 1).

2. Notons $\mathcal{C}$ l'ensemble des applications croissantes de $[\![1\,,p]\!]$ dans $[\![1\,,n]\!]$, et $\mathcal{S}$ celui des applications strictement croissantes de $[\![1\,,p]\!]$ dans $[\![1\,,n+p-1]\!]$.

▷ *φ est à valeurs dans $\mathcal{S}$.* Soit $f \in \mathcal{C}$ et $k \in [\![1\,,p]\!]$. On a $1 \leqslant f(k) \leqslant f(k)+k-1 \leqslant n+p-1$, et, pour $k \leqslant p-1$,

$$\varphi(f)(k+1) = f(k+1)+k \geqslant f(k)+k > f(k)+k-1 = \varphi(f)(k).$$

▷ *Une réciproque.* Pour $g \in \mathcal{S}$, on pose $\psi(g)(k) = g(k)-k+1$. Comme g est strictement croissante et à valeurs entières, $g(k+1) \geqslant g(k)+1$, donc $\psi(g)(k+1) = g(k+1)-k \geqslant g(k)-k+1 = \psi(g)(k)$: $\psi(g)$ est croissante. Par récurrence, $g(k) \geqslant g(1)+(k-1) \geqslant k$, et de même $g(k) \leqslant g(p)-(p-k) \leqslant n+k-1$; ainsi $\psi(g)(k) \in [\![1\,,n]\!]$, et $\psi(g) \in \mathcal{C}$. Enfin $\psi \circ \varphi$ et $\varphi \circ \psi$ sont les identités, puisque l'on ajoute puis retranche $k-1$. Par le chapitre 2, théorème 1, φ est une bijection de $\mathcal{C}$ sur $\mathcal{S}$.

La question 1, appliquée avec $n+p-1$ au lieu de n, donne

$$\boxed{|\mathcal{C}| = |\mathcal{S}| = \binom{n+p-1}{p}}$$

Pour $n=2$ et $p=2$, on trouve $\binom{3}{2} = 3$: ce sont les applications de valeurs $(1,1)$, $(1,2)$ et $(2,2)$ ✓.

L'exercice 46 de la feuille du chapitre 3 a dénombré les p-uplets d'entiers naturels de somme n : il y en a $\binom{n+p-1}{p-1}$. Une application croissante f est déterminée par les nombres $x_j = |f^{-1}(\{j\})|$ de fois qu'elle prend chaque valeur $j \in [\![1\,,n]\!]$: c'est un n-uplet d'entiers naturels de somme p, et l'on retrouve, en échangeant les rôles de n et p, le nombre $\binom{p+n-1}{n-1} = \binom{n+p-1}{p}$ ✓. C'est le nombre « hors programme » du tableau de la section 21.2.6 : choisir p éléments parmi n, sans ordre et avec répétitions, c'est choisir les valeurs d'une application croissante.

Piège. L'erreur la plus répandue est de croire qu'une application croissante est déterminée par son image, comme à la question 1 : les valeurs répétées sont perdues. Le décalage $k-1$ de φ sert précisément à séparer les valeurs égales, pour se ramener au cas strict.

EXERCICE 13 ⋆⋆

On considère l'application

$$\chi : S \to \{0,1\}^m, \qquad x \mapsto (\mathbf{1}_{A_1}(x), \ldots, \mathbf{1}_{A_m}(x)).$$

Elle est injective : si $x \neq y$, il existe i tel que A_i contienne l'un des deux éléments et pas l'autre, c'est-à-dire tel que $\mathbf{1}_{A_i}(x) \neq \mathbf{1}_{A_i}(y)$, et les images de x et de y diffèrent par leur i-ième coordonnée. Par le chapitre 2, théorème 2, et la proposition 2 (iv), on a $n \leqslant 2^m$, et la croissance du logarithme donne $\ln n \leqslant m \ln 2$.

$$\boxed{m \geqslant \frac{\ln n}{\ln 2}}$$

Pour l'égalité, on prend $S = \{0,1\}^m$, qui a $n = 2^m$ éléments, et $A_i = \{x \in S \mid x_i = 1\}$ pour $i \in [\![1\,,m]\!]$. Deux éléments distincts de S diffèrent par une coordonnée i, et A_i contient l'un et pas l'autre : les A_i séparent les points, et $m = \dfrac{\ln n}{\ln 2}$. Pour $m = 3$, on sépare ainsi les huit éléments de $\{0,1\}^3$ avec trois parties.

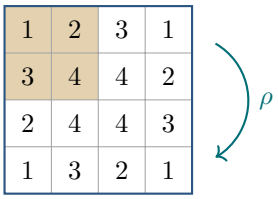
On peut le lire comme un jeu de questions : pour identifier un élément inconnu de S, on pose les questions « appartient-il à A_i ? », auxquelles on répond par oui ou non. Les parties séparent les points si et seulement si les réponses déterminent l'élément, et il faut au moins $\dfrac{\ln n}{\ln 2}$ questions, nombre de chiffres de l'écriture de n en base 2, à un près. C'est le principe de la recherche dichotomique, qui atteint ce minimum.

EXERCICE 14 ⋆⋆

On repère les cases par leurs coordonnées $(i,j) \in [\![0\,,n-1]\!]^2$, ligne et colonne, et l'on note ρ l'application qui envoie une case sur la case qu'elle occupe après un quart de tour, soit $\rho(i,j) = (j, n-1-i)$. C'est une permutation de l'ensemble des n^2 cases, et $\rho^4 = \mathrm{id}$. Si H est l'ensemble des trous, la grille tournée de r quarts de tour laisse voir les cases de $\rho^r(H)$, pour $r \in [\![0\,,3]\!]$.

▷ *Orbites.* On a $\rho^2(i,j) = (n-1-i, n-1-j)$: c'est la symétrie par rapport au centre. Une case est fixée par ρ^2 si et seulement si $2i = 2j = n-1$, ce qui n'arrive que pour n impair, et pour la seule case centrale. Si une case c vérifiait $\rho(c) = c$ ou $\rho^3(c) = c$, elle vérifierait $\rho^2(c) = c$, en composant par ρ ; ainsi, lorsque n est pair, les cases c, $\rho(c)$, $\rho^2(c)$, $\rho^3(c)$ sont deux à deux distinctes, et les orbites de ρ, au sens du chapitre 17, ont toutes quatre éléments. Il y en a $\dfrac{n^2}{4}$.

▷ *Caractérisation.* Une case c apparaît à travers un trou dans la position r si et seulement si $c \in \rho^r(H)$, c'est-à-dire $\rho^{-r}(c) \in H$. Le nombre de positions où elle apparaît est donc le nombre d'éléments de H parmi c, $\rho^{-1}(c)$, $\rho^{-2}(c)$, $\rho^{-3}(c)$, qui forment l'orbite de c. Lorsque n est pair, la grille est une grille de Fleissner si et seulement si H contient *exactement une* case de chaque orbite.



La figure numérote les cases d'une grille de côté 4 selon leur orbite : les quatre coins forment l'orbite 1, le carré central l'orbite 4. Une grille de Fleissner choisit un trou portant chacun des numéros 1 à 4, par exemple le quart supérieur gauche, grisé en ocre ; le quart de tour ρ envoie chaque case sur une case de même numéro.

1. Pour $n = 6$, il y a $\dfrac{36}{4} = 9$ orbites, donc 9 trous, et l'on choisit une case dans chaque orbite, de 4 façons indépendamment (proposition 2 (iv)).

$$\boxed{\text{9 trous, et } 4^9 = 262\,144 \text{ grilles de Fleissner de côté 6}}$$

2. Si n est pair, le même raisonnement donne $\dfrac{n^2}{4}$ trous et $4^{n^2/4} = 2^{n^2/2}$ grilles. Si n est impair, la case centrale c_0 est fixée par ρ : elle apparaît dans les quatre positions si $c_0 \in H$, et dans aucune sinon, jamais dans exactement une.

$$\boxed{\text{Il existe des grilles de Fleissner de côté } n \text{ si et seulement si } n \text{ est pair ; il y en a alors } 4^{n^2/4}}$$

Une énumération exhaustive confirme les valeurs 4 pour $n = 2$, 0 pour $n = 3$ et 256 pour $n = 4$. Le nombre de grilles croît très vite : une grille de côté 10 se choisit parmi $4^{25} \approx 1{,}1 \cdot 10^{15}$, ce qui rend vaine la recherche de la grille par essais successifs.

EXERCICE 15 ⋆⋆

On note $\mathrm{Surj}(E,F)$ l'ensemble des surjections de E sur F. Si $\alpha : E' \to E$ et $\beta : F \to F'$ sont des bijections, l'application $f \mapsto \beta \circ f \circ \alpha$ est une bijection de $\mathrm{Surj}(E,F)$ sur $\mathrm{Surj}(E',F')$ (chapitre 2, proposition 6) : c'est pourquoi $S(n,p)$ ne dépend que des cardinaux, et l'on peut raisonner avec $E = [\![1\,,n]\!]$ et $F = [\![1\,,p]\!]$.

1. S'il existe une surjection d'un ensemble à n éléments sur un ensemble à p éléments, alors $n \geqslant p$ (chapitre 2, théorème 2). Pour $p = n$, les surjections sont les bijections (proposition 1 (iii)), au nombre de $n!$ (théorème 1). Pour $p = 1$, la seule application est constante, et elle est surjective. Pour $p = 2$, une application de $[\![1\,,n]\!]$ dans $\{1,2\}$ est surjective si et seulement si elle n'est pas constante.

$$\boxed{S(n,p) = 0 \text{ si } p > n, \qquad S(n,n) = n!, \qquad S(n,1) = 1, \qquad S(n,2) = 2^n - 2}$$

2. Soit f une surjection de $[\![1\,,n+1]\!]$ sur $[\![1\,,n]\!]$. Ses n fibres $f^{-1}(\{y\})$ sont non vides, deux à deux disjointes, et recouvrent $[\![1\,,n+1]\!]$: leurs cardinaux, tous au moins égaux à 1, ont pour somme $n+1$, et exactement une fibre a deux éléments, les autres étant des singletons. On note $P(f)$ cette fibre à deux éléments. Fixons une paire $\{i,j\}$: les surjections telles que $P(f) = \{i,j\}$ sont celles qui prennent la même valeur en i et en j et des valeurs deux à deux distinctes, différentes de celle-là, aux $n-1$ autres points. Elles correspondent aux bijections de l'ensemble à n éléments $\{\{i,j\}\} \cup \{\{k\} \mid k \neq i,j\}$ sur $[\![1\,,n]\!]$: il y en a $n!$. Le lemme des bergers, appliqué à $f \mapsto P(f)$, donne

$$\boxed{S(n+1,n) = \binom{n+1}{2} n! = \frac{n\,(n+1)!}{2}}$$

Pour $n = 2$, on trouve $6 = 2^3 - 2$, en accord avec la question 1 ✓.

3. Soit $n \geqslant 2$, $p \geqslant 2$, et f une surjection de $[\![1\,,n]\!]$ sur $[\![1\,,p]\!]$; on note g sa restriction à $[\![1\,,n-1]\!]$.

— Si g est surjective sur $[\![1\,,p]\!]$, la valeur $f(n)$ est quelconque. Réciproquement, toute surjection g de $[\![1\,,n-1]\!]$ sur $[\![1\,,p]\!]$, prolongée par une valeur quelconque, donne une surjection f. Il y a $p\,S(n-1,p)$ telles surjections f (proposition 2 (iv)).

— Si g n'est pas surjective, elle manque une valeur y ; comme f est surjective, y est atteinte en n, donc $y = f(n)$, et y est la seule valeur manquée. Ainsi g est une surjection de $[\![1\,,n-1]\!]$ sur $[\![1\,,p]\!] \setminus \{f(n)\}$. Réciproquement, pour $y \in [\![1\,,p]\!]$ et g une surjection de $[\![1\,,n-1]\!]$ sur cet ensemble à $p-1$ éléments, le prolongement de g par $f(n) = y$ est une surjection sur $[\![1\,,p]\!]$ dont la restriction n'est pas surjective. Il y a $p\,S(n-1,p-1)$ telles surjections f.

Les deux cas sont disjoints, et la proposition 2 (i) donne

$$\boxed{S(n,p) = p(S(n-1,p) + S(n-1,p-1)) \quad \text{pour } n \geqslant 2 \text{ et } p \geqslant 2}$$

4. Le programme suit la question 3, avec pour cas de base ceux de la question 1.

```
def S(n, p):                # n >= 1 et p >= 1
    if p > n:
        return 0
    if p == 1:
        return 1
    return p * (S(n - 1, p) + S(n - 1, p - 1))

print(S(5, 3))              # affiche 150
```

Si un appel $S(n,p)$, avec $n \geqslant 1$ et $p \geqslant 1$, ne s'arrête pas dans l'un des deux tests, c'est que $2 \leqslant p \leqslant n$: il provoque deux appels avec le premier argument $n-1 \geqslant 1$ et des seconds arguments $p \geqslant 1$ et $p-1 \geqslant 1$, de sorte que l'hypothèse $n,p \geqslant 1$ est préservée. Le premier argument diminue de 1 à chaque appel et reste un entier au moins égal à 1 : une chaîne d'appels emboîtés a donc au plus n étages, et

le programme se termine. Lorsque $n = 1$, on a soit $p > 1$, soit $p = 1$, et l'un des deux tests s'applique. La valeur $S(5,3) = 150$ est celle que la remarque qui suit la proposition 6 a obtenue par la formule ✓.

5. Notons $R(n,p) = \sum_{k=0}^{p} (-1)^{p-k} \binom{p}{k} k^n$, et montrons par récurrence sur $n \geqslant 1$ la propriété : $S(n,p) = R(n,p)$ pour tout $p \geqslant 1$.

Initialisation : Pour $n = 1$, on utilise $k\binom{p}{k} = p\binom{p-1}{k-1}$ pour $k \geqslant 1$, puis la formule du binôme :

$$R(1,p) = \sum_{k=1}^{p} (-1)^{p-k}\, p \binom{p-1}{k-1} = p \sum_{j=0}^{p-1} (-1)^{p-1-j} \binom{p-1}{j} = p\,(1-1)^{p-1},$$

qui vaut 1 si $p = 1$ et 0 si $p \geqslant 2$: c'est $S(1,p)$ par la question 1.

Hérédité : Supposons la propriété vraie au rang $n-1$, pour un entier $n \geqslant 2$. Si $p = 1$, alors $R(n,1) = 1^n - 0^n = 1 = S(n,1)$. Pour $p \geqslant 2$, la question 3 et l'hypothèse de récurrence donnent

$$S(n,p) = p\,R(n-1,p) + p\,R(n-1,p-1), \qquad \text{soit}$$

$$S(n,p) = \sum_{k=0}^{p} (-1)^{p-k}\, p \binom{p}{k} k^{n-1} - \sum_{k=0}^{p-1} (-1)^{p-k}\, p \binom{p-1}{k} k^{n-1} = \sum_{k=0}^{p} (-1)^{p-k}\, p \left(\binom{p}{k} - \binom{p-1}{k} \right) k^{n-1},$$

avec la convention $\binom{p-1}{p} = 0$. Par la formule de Pascal, $\binom{p}{k} - \binom{p-1}{k} = \binom{p-1}{k-1}$ pour $k \geqslant 1$, et $p\binom{p-1}{k-1} = k\binom{p}{k}$; pour $k = 0$, le coefficient est nul. Ainsi

$$S(n,p) = \sum_{k=0}^{p} (-1)^{p-k}\, k \binom{p}{k} k^{n-1} = R(n,p).$$

Conclusion : La formule est vraie pour tous $n, p \geqslant 1$.

$$\boxed{S(n,p) = \sum_{k=0}^{p} (-1)^{p-k} \binom{p}{k} k^n}$$

La proposition 6 obtient la même formule en développant un produit, d'un seul coup ; la récurrence la retrouve sans la formule du crible, mais il faut connaître le résultat pour le démontrer. Les deux expressions ont été comparées à une énumération des surjections pour $n, p \leqslant 7$.

EXERCICE 16 ⋆⋆

1. Les k parties d'une partition de $[\![1\,,n]\!]$ en parties non vides sont deux à deux disjointes et ont chacune au moins un élément, donc $k \leqslant n$ par la proposition 2 (i). La seule partition en une partie est $\{[\![1\,,n]\!]\}$, et la seule partition en n parties est celle en singletons.

$$\boxed{T(n,k) = 0 \text{ pour } k > n, \qquad T(n,1) = T(n,n) = 1 \text{ pour } n \geqslant 1}$$

2. Soit $n, k \geqslant 1$. On sépare les partitions de $[\![1\,,n]\!]$ en k parties selon que $\{n\}$ en est une partie ou non.

— Si $\{n\}$ est une partie, les $k-1$ autres forment une partition de $[\![1\,,n-1]\!]$, et réciproquement toute partition de $[\![1\,,n-1]\!]$ en $k-1$ parties, complétée par $\{n\}$, convient. Il y a $T(n-1,k-1)$ telles partitions ; pour $n = 1$, la partition $\{\{1\}\}$ correspond à la partition vide de $\varnothing$, ce que traduit la convention $T(0,0) = 1$.

— Sinon, n appartient à une partie P qui a d'autres éléments. En retirant n de P, on obtient une partition de $[\![1\,,n-1]\!]$ en k parties non vides. Réciproquement, une partition de $[\![1\,,n-1]\!]$ en k parties est obtenue de cette façon à partir de exactement k partitions, selon la partie à laquelle on ajoute n. Le lemme des bergers donne $k\,T(n-1,k)$ partitions.

$$\boxed{T(n,k) = T(n-1,k-1) + k\,T(n-1,k) \quad \text{pour } n,k \geqslant 1}$$

Les conventions rendent la formule valable au bord. Pour $n = k = 1$, par exemple, le second membre vaut $T(0,0) + T(0,1) = 1$, qui est bien $T(1,1)$ ✓.

```python
def T(n, k):
    if n == 0 and k == 0:
        return 1
    if n == 0 or k == 0:
        return 0
    return T(n - 1, k - 1) + k * T(n - 1, k)

print([T(4, k) for k in range(5)])   # affiche [0, 1, 7, 6, 1]
```

Chaque appel non terminal diminue n de 1, et les appels avec $n = 0$ s'arrêtent : la fonction se termine, comme à l'exercice 15.

3. À une surjection f de $[\![1\,,n]\!]$ sur $[\![1\,,k]\!]$, on associe la partition $\pi(f)$ de $[\![1\,,n]\!]$ formée de ses k fibres $f^{-1}(\{y\})$, non vides puisque f est surjective. Fixons une partition $\{P_1,\dots,P_k\}$ en k parties : les surjections f telles que $\pi(f)$ soit cette partition sont constantes sur chaque P_i, et la valeur prise sur les parties définit une bijection de $\{P_1,\dots,P_k\}$ sur $[\![1\,,k]\!]$; réciproquement, chaque bijection fournit une telle surjection. Chaque fibre de π a donc $k!$ éléments, et le lemme des bergers donne

$$\boxed{S(n,k) = k!\,T(n,k) \quad \text{pour } n,k \geqslant 1}$$

Choisir une surjection, c'est choisir un partage en k paquets, puis étiqueter les paquets. On en déduit $T(n,2) = \dfrac{2^n-2}{2} = 2^{n-1}-1$, et $T(n+1,n) = \dbinom{n+1}{2}$ par l'exercice 15.

4. On a $B_1 = 1$, puis $B_2 = 2$, avec les partitions $\{\{1,2\}\}$ et $\{\{1\},\{2\}\}$, et $B_3 = 5$, avec $\{\{1,2,3\}\}$, les trois partitions $\{\{i,j\},\{k\}\}$ et la partition en singletons.

Soit $n \in \mathbb{N}$, et $\mathcal{P}$ une partition de $[\![1\,,n+1]\!]$. Notons P la partie qui contient $n+1$, et $Q = P\setminus\{n+1\}$, partie de $[\![1\,,n]\!]$. Les autres parties de $\mathcal{P}$ forment une partition de $[\![1\,,n]\!]\setminus Q$. Réciproquement, une partie Q de $[\![1\,,n]\!]$ et une partition de $[\![1\,,n]\!]\setminus Q$ déterminent une unique partition de $[\![1\,,n+1]\!]$, dont $Q\cup\{n+1\}$ est la partie contenant $n+1$. Si $[\![1\,,n]\!]\setminus Q$ a k éléments, il y a B_k partitions de cet ensemble, avec $B_0 = 1$ pour $Q = [\![1\,,n]\!]$; et il y a $\dbinom{n}{k}$ parties Q dont le complémentaire a k éléments (proposition 4). En regroupant selon k :

$$\boxed{B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k}$$

$$B_4 = 1 + 3\times 1 + 3\times 2 + 1\times 5 = 15, \qquad B_5 = 1 + 4\times 1 + 6\times 2 + 4\times 5 + 1\times 15 = 52.$$

$$\boxed{B_1 = 1, \quad B_2 = 2, \quad B_3 = 5, \quad B_4 = 15, \quad B_5 = 52}$$

L'exercice 28 de la feuille du chapitre 2 a dénombré les partitions de $\{1,2,3,4\}$, qui sont 15, et en a déduit le nombre de relations d'équivalence sur un ensemble à quatre éléments : le chapitre 2 (proposition 8 et théorème 4) a établi une bijection entre relations d'équivalence et partitions. Le nombre B_n est donc aussi le nombre de relations d'équivalence sur un ensemble à n éléments. On a encore $B_n = \sum_{k=0}^{n} T(n,k)$, en regroupant les partitions selon leur nombre de parties : par exemple $B_4 = 1+7+6+1$.

À retenir. Le corrigé de l'exercice 3 de la feuille du chapitre 20 avait trouvé les coefficients 1, 1, 2, 5, 15 de e dans $\sum_{n=0}^{+\infty} \dfrac{n^k}{n!}$, pour $k = 0, \dots, 4$, et annonçait qu'ils comptent les partitions d'un ensemble à k éléments : on reconnaît $B_0, \dots, B_4$. Le calcul de sympy confirme $\sum_{n=0}^{+\infty} \dfrac{n^k}{n!} = B_k$ e jusqu'à $k = 5$.

EXERCICE 17 ⋆⋆

Idée. Une partie stable par σ doit contenir, avec chacun de ses éléments, toute son orbite : les parties stables sont les réunions d'orbites, et l'on choisit librement quelles orbites on réunit.

Rappelons (chapitre 17, démonstration du théorème 1) que les orbites de σ sont les classes de la relation « il existe $k \in \mathbb{Z}$ tel que $j = \sigma^k(i)$ », et qu'elles forment une partition de $[\![1\,,n]\!]$; on note $\mathcal{O}_1$, …, $\mathcal{O}_\omega$ ces orbites, avec $\omega = \omega(\sigma)$.

▷ *Une réunion d'orbites est stable.* Pour une orbite $\mathcal{O}_j$ et $x \in \mathcal{O}_j$, l'élément $\sigma(x)$ est dans la même orbite : $\sigma(\mathcal{O}_j) \subset \mathcal{O}_j$. Comme σ est injective, $\sigma(\mathcal{O}_j)$ a autant d'éléments que $\mathcal{O}_j$, et l'inclusion est une égalité (proposition 1 (ii)). Pour une réunion $A = \bigcup_{j\in J} \mathcal{O}_j$, on obtient $\sigma(A) = \bigcup_{j\in J} \sigma(\mathcal{O}_j) = A$ (chapitre 2, proposition 2 (i)).

▷ *Une partie stable est une réunion d'orbites.* Soit A telle que $\sigma(A) = A$. Alors $\sigma^{-1}(A) = A$, et, par récurrence, $\sigma^k(A) = A$ pour tout $k \in \mathbb{Z}$. Pour $x \in A$, l'orbite de x, formée des $\sigma^k(x)$, est donc contenue dans A. Ainsi A est la réunion des orbites de ses éléments.

▷ *Dénombrement.* L'application $J \mapsto \bigcup_{j\in J} \mathcal{O}_j$, de $\mathcal{P}([\![1\,,\omega]\!])$ dans l'ensemble des parties stables, est surjective par l'étape précédente, et injective parce que les orbites sont non vides et deux à deux disjointes : on retrouve J comme l'ensemble des j tels que $\mathcal{O}_j \subset A$. Par la proposition 3,

$$\boxed{\Big|\{A \subset [\![1\,,n]\!] \mid \sigma(A) = A\}\Big| = 2^{\omega(\sigma)}}$$

Pour l'identité, $\omega = n$ et toutes les 2^n parties sont stables ; pour un n-cycle, $\omega = 1$ et seules $\varnothing$ et $[\![1\,,n]\!]$ le sont. Pour la permutation $\sigma = (1\,3\,5)(2\,6)$ de $\mathfrak{S}_6$, exemple du chapitre 17, les orbites sont $\{1,3,5\}$, $\{2,6\}$ et $\{4\}$, et il y a $2^3 = 8$ parties stables. Une énumération confirme le résultat pour toutes les permutations de $\mathfrak{S}_n$, $n \leqslant 5$.

Piège. L'erreur la plus répandue est de confondre « $\sigma(A) = A$ » et « $\sigma(x) = x$ pour tout $x \in A$ » : la seconde condition fait de A une partie de l'ensemble des points fixes, et il y en aurait $2^{n-|\mathrm{Supp}(\sigma)|}$. Une partie stable peut être déplacée point par point tout en restant globalement invariante.

EXERCICE 18 ⋆⋆

1. Pour $n = 1$, la seule permutation est l'identité, qui a un point fixe. Pour $n = 2$, la transposition est le seul dérangement. Pour $n = 3$, une permutation sans point fixe a pour seule orbite $[\![1\,,3]\!]$: c'est l'un des deux 3-cycles $(1\,2\,3)$ et $(1\,3\,2)$.

$$\boxed{D_1 = 0, \qquad D_2 = 1, \qquad D_3 = 2}$$

2. Soit F une partie à k éléments de $[\![1\,,n]\!]$. Si σ est une permutation dont l'ensemble des points fixes est exactement F, alors $\sigma(F) = F$, donc σ envoie aussi le complémentaire $\overline{F}$ sur lui-même, puisqu'elle est bijective ; sa restriction à $\overline{F}$ est une permutation de $\overline{F}$ sans point fixe. Réciproquement, un dérangement de $\overline{F}$, prolongé par l'identité sur F, est une permutation dont l'ensemble des points fixes est F. Il y a donc D_{n-k} telles permutations, puisque $\overline{F}$ a $n-k$ éléments, et $\binom{n}{k}$ choix de F.

$$\boxed{\binom{n}{k} D_{n-k} \text{ permutations de } [\![1\,,n]\!] \text{ ont exactement } k \text{ points fixes}}$$

3. Les permutations se répartissent selon leur nombre $k \in [\![0\,,n]\!]$ de points fixes. La proposition 2 (i), puis le changement d'indice $j = n-k$ et la symétrie $\binom{n}{n-j} = \binom{n}{j}$ donnent

$$\boxed{n! = \sum_{k=0}^{n} \binom{n}{k} D_{n-k} = \sum_{j=0}^{n} \binom{n}{j} D_j}$$

La relation détermine D_n à partir de D_0, …, D_{n-1}, puisque le coefficient de D_n vaut 1 :

$$D_4 = 24 - \big(1 + 4\times 0 + 6\times 1 + 4\times 2\big) = 9, \qquad D_5 = 120 - \big(1 + 0 + 10\times 1 + 10\times 2 + 5\times 9\big) = 44.$$

$$\boxed{D_4 = 9, \qquad D_5 = 44}$$

4. Posons $E_n = \displaystyle\sum_{k=0}^{n} (-1)^{n-k}\binom{n}{k} k!$ pour $n \in \mathbb{N}$, et montrons d'abord que la suite (E_n) vérifie la même relation que (D_n). Soit $n \in \mathbb{N}$. En intervertissant les sommes triangulaires (chapitre 3, proposition 8), puis en utilisant l'égalité de l'énoncé, qui se vérifie en écrivant les deux membres avec des factorielles,

$$\sum_{k=0}^{n}\binom{n}{k}E_k = \sum_{k=0}^{n}\sum_{j=0}^{k}(-1)^{k-j}\binom{n}{k}\binom{k}{j}j! = \sum_{j=0}^{n} j!\binom{n}{j}\sum_{k=j}^{n}(-1)^{k-j}\binom{n-j}{k-j}.$$

Avec $i = k - j$ et $m = n - j$, la somme intérieure vaut $\displaystyle\sum_{i=0}^{m}(-1)^i\binom{m}{i} = (1-1)^m$ par la formule du binôme : elle est nulle pour $m \geqslant 1$ et vaut 1 pour $m = 0$, c'est-à-dire $j = n$. Il reste

$$\sum_{k=0}^{n}\binom{n}{k}E_k = n!\binom{n}{n} = n!.$$

On montre alors par récurrence forte que $D_n = E_n$. On a $D_0 = 1 = E_0$. Si $D_k = E_k$ pour tout $k \leqslant n-1$, avec $n \geqslant 1$, les relations vérifiées par les deux suites donnent

$$D_n = n! - \sum_{k=0}^{n-1}\binom{n}{k}D_k = n! - \sum_{k=0}^{n-1}\binom{n}{k}E_k = E_n.$$

$$\boxed{D_n = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}k! = n!\sum_{i=0}^{n}\frac{(-1)^i}{i!}}$$

La seconde forme s'obtient en écrivant $\displaystyle\binom{n}{k}k! = \frac{n!}{(n-k)!}$ et en posant $i = n - k$. Pour $n = 4$, on retrouve $24 - 24 + 12 - 4 + 1 = 9$ ✓. La remarque de la section 21.6.3 obtient la même probabilité $\dfrac{D_n}{n!}$ par la formule du crible. Le passage d'une relation $a_n = \sum\binom{n}{k}b_k$ à $b_n = \sum(-1)^{n-k}\binom{n}{k}a_k$, que l'on vient de démontrer dans un cas particulier, s'appelle l'*inversion de Pascal*.

5. On numérote les hommes et les femmes de 1 à 5, de sorte que l'homme i et la femme i forment un couple légitime. Une association des danseurs est une permutation σ de $[\![1\,,5]\!]$, l'homme i dansant avec la femme $\sigma(i)$; on prend $\Omega = \mathfrak{S}_5$, muni de la probabilité uniforme, puisque toutes les associations sont également probables. Le nombre de couples légitimes reconstitués est le nombre de points fixes de σ, et la question 2 donne la probabilité qu'il vaille k : $\dfrac{1}{120}\binom{5}{k}D_{5-k}$. Avec k couples légitimes et $5-k$ illégitimes, il y a plus de couples illégitimes que de légitimes si et seulement si $5 - k > k$, c'est-à-dire $k \leqslant 2$.

$$\boxed{\mathrm{P}(\text{aucun couple légitime}) = \frac{D_5}{120} = \frac{11}{30}, \qquad \mathrm{P}(\text{un seul}) = \frac{5D_4}{120} = \frac{3}{8}}$$

$$\boxed{\mathrm{P}(\text{plus de couples illégitimes que de légitimes}) = \frac{D_5 + 5D_4 + 10D_3}{120} = \frac{44 + 45 + 20}{120} = \frac{109}{120}}$$

La probabilité $\dfrac{11}{30} \approx 0{,}367$ est déjà très proche de sa limite $\dfrac{1}{\mathrm{e}} \approx 0{,}368$, comme le notait le cours. Celle de l'événement contraire de la dernière question, $\dfrac{11}{120}$, se lit directement : il faut au moins trois couples légitimes, soit $\displaystyle\binom{5}{3}D_2 + \binom{5}{4}D_1 + D_0 = 10 + 0 + 1 = 11$ associations.

Piège. L'erreur la plus répandue, à la question 2, est d'écrire $\binom{n}{k}(n-k)!$: on choisit les k points fixes, puis on permute les autres éléments n'importe comment. Mais une permutation quelconque des $n-k$ autres éléments peut avoir des points fixes supplémentaires, et l'on compterait les permutations ayant *au moins* k points fixes, chacune plusieurs fois. Il faut un dérangement du complémentaire.

EXERCICE 19 ⋆⋆

Par l'exercice 7 de la feuille du chapitre 17, l'ordre d'une permutation différente de l'identité est le ppcm des longueurs des cycles de sa décomposition (chapitre 17, théorème 1), et l'identité est d'ordre 1. Une permutation $\sigma \in \mathfrak{S}_{2p}$ est donc d'ordre p si et seulement si $\sigma \neq \mathrm{id}$ et le ppcm des longueurs de ses cycles vaut p. Chaque longueur divise alors p et vaut au moins 2 : comme p est premier, tous les cycles sont de longueur p. Leurs supports étant disjoints dans un ensemble à $2p$ éléments, il y a un ou deux cycles. Réciproquement, une permutation formée d'un ou de deux p-cycles à supports disjoints est d'ordre p.

▷ *Nombre de p-cycles de support donné.* Soit X un ensemble à p éléments. L'application qui associe à une p-liste $(a_1, \dots, a_p)$ d'éléments distincts de X le cycle $(a_1\, a_2\, \cdots\, a_p)$ est surjective sur l'ensemble des p-cycles de support X. Deux listes donnent le même cycle c si et seulement si l'une se déduit de l'autre par une rotation, puisque c impose $a_{i+1} = c(a_i)$: une fois le premier terme choisi parmi les p éléments de X, toute la liste est déterminée. Les fibres ont p éléments, et le lemme des bergers donne $\dfrac{p!}{p} = (p-1)!$ cycles.

▷ *Un seul p-cycle.* On choisit son support, de $\binom{2p}{p}$ façons, puis le cycle, de $(p-1)!$ façons : $\binom{2p}{p}(p-1)!$ permutations.

▷ *Deux p-cycles.* Les deux supports forment une partition $\{X, \overline{X}\}$ de $[\![1\,,2p]\!]$ en deux parties à p éléments. L'application $X \mapsto \{X, \overline{X}\}$ a des fibres à deux éléments, X et $\overline{X}$, qui sont distincts : il y a $\dfrac{1}{2}\binom{2p}{p}$ telles partitions. Pour chacune, on choisit un p-cycle sur chaque partie, de $(p-1)!^2$ façons, et l'unicité de la décomposition en cycles (chapitre 17, théorème 1) assure que des choix différents donnent des permutations différentes.

$$\boxed{\binom{2p}{p}(p-1)! + \frac{1}{2}\binom{2p}{p}(p-1)!^2 = \frac{(2p)!}{p\cdot p!} + \frac{(2p)!}{2p^2} \text{ éléments d'ordre } p}$$

Pour $p=2$, on trouve $\dfrac{24}{4} + \dfrac{24}{8} = 6+3 = 9$: ce sont les 6 transpositions et les 3 doubles transpositions $(1\,2)(3\,4)$, $(1\,3)(2\,4)$, $(1\,4)(2\,3)$ de $\mathfrak{S}_4$ ✓. Pour $p=3$, on trouve $40+40=80$, et pour $p=5$, $6\,048 + 72\,576 = 78\,624$. Les valeurs pour $p=2$ et $p=3$ ont été confirmées par énumération de $\mathfrak{S}_4$ et de $\mathfrak{S}_6$.

Piège. L'erreur la plus répandue est de compter $\binom{2p}{p}(p-1)!^2$ permutations formées de deux p-cycles, en choisissant « le support du premier cycle » : les deux cycles commutent et jouent le même rôle, et chaque permutation est comptée deux fois. Pour $p=2$, on trouverait 6 doubles transpositions au lieu de 3.

EXERCICE 20 ⋆⋆⋆

1. La famille $\{\varnothing\}$, ou toute famille réduite à une partie, est une antichaîne de cardinal 1. La famille des singletons $\{\{1\},\dots,\{n\}\}$ en est une de cardinal n, car aucun singleton n'est inclus dans un autre. Plus généralement, pour $k \in [\![0\,,n]\!]$, la famille de toutes les parties à k éléments est une antichaîne : si $A \subset B$ avec $|A| = |B|$, alors $A = B$ (proposition 1 (ii)). Son cardinal est $\binom{n}{k}$, maximal pour $k = \lfloor \frac{n}{2} \rfloor$ (exercice 28 de la feuille du chapitre 3). Pour $n = 4$, la famille $\{\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}$ est une antichaîne à 6 éléments. On peut aussi mélanger les cardinaux, comme $\{\{1\},\{2,3\}\}$ pour $n = 3$.

2.(a) Si $C_{k-1} \subsetneq C_k$, alors $|C_{k-1}| < |C_k|$ (proposition 1 (ii)), donc $|C_k| \geqslant |C_{k-1}| + 1$, puisque ce sont des entiers. Par récurrence, $|C_k| \geqslant |C_0| + k \geqslant k$ et $|C_k| \leqslant |C_n| - (n-k) \leqslant k$.

$$\boxed{|C_k| = k \text{ pour tout } k \in [\![0\,,n]\!]}$$

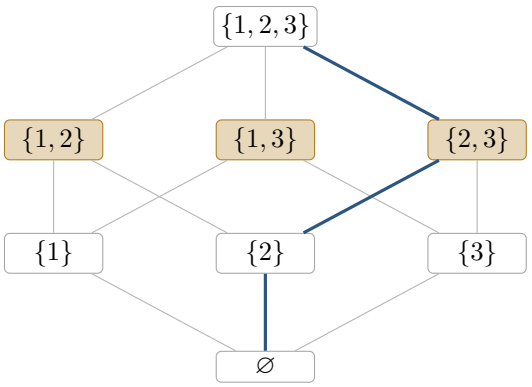
En particulier $C_0 = \varnothing$ et $C_n = [\![1\,,n]\!]$, et, pour $k \geqslant 1$, $C_k \setminus C_{k-1}$ est un singleton $\{x_k\}$. Les éléments x_1, …, x_n sont deux à deux distincts, puisque $x_k \notin C_{k-1}$ alors que $x_j \in C_{k-1}$ pour $j < k$, et $C_k = \{x_1,\dots,x_k\}$. L'application qui associe à une chaîne complète la liste $(x_1,\dots,x_n)$ est une bijection de l'ensemble des chaînes complètes sur l'ensemble des n-listes d'éléments distincts de $[\![1\,,n]\!]$: sa réciproque associe à une liste la chaîne des $C_k = \{x_1,\dots,x_k\}$. Par le théorème 1 :

$$\boxed{n! \text{ chaînes complètes}}$$

On construit une chaîne en ajoutant les éléments un par un, dans un ordre choisi.

2.(b) Soit X une partie, de cardinal k. Par la question (a), X ne peut être que le terme C_k d'une chaîne, et $C_k = X$ si et seulement si $\{x_1,\dots,x_k\} = X$: la liste commence par une énumération des éléments de X, en $k!$ ordres possibles, et se termine par une énumération de $\overline{X}$, en $(n-k)!$ ordres.

$$\boxed{|X|!\,(n-|X|)! \text{ chaînes complètes passent par } X}$$



La figure représente les parties de $[\![1\,,3]\!]$, chaque trait reliant une partie à une partie qui a un élément de plus. Une chaîne complète, en bleu, monte du bas vers le haut en ajoutant 2, puis 3, puis 1 : il y en a $3! = 6$, une par ordre d'ajout. Elle traverse l'antichaîne des parties à deux éléments, en ocre, en un seul point, $\{2,3\}$.

3. Pour $A \in \mathcal{A}$, notons $\mathcal{K}_A$ l'ensemble des chaînes complètes qui passent par A. Si une chaîne passait par deux éléments distincts A et A' de $\mathcal{A}$, l'un serait inclus dans l'autre, puisque deux termes d'une chaîne sont toujours comparables pour l'inclusion : c'est exclu dans une antichaîne. Les ensembles $\mathcal{K}_A$, pour $A \in \mathcal{A}$, sont donc deux à deux disjoints, et leur réunion est contenue dans l'ensemble des $n!$ chaînes complètes. Les propositions 2 (i) et 1 (ii), avec la question 2, donnent

$$\boxed{\sum_{A\in\mathcal{A}} |A|!\,(n-|A|)! = \sum_{A\in\mathcal{A}} |\mathcal{K}_A| \leqslant n!}$$

4. En divisant par $n!$, l'inégalité s'écrit $\sum_{A\in\mathcal{A}} \binom{n}{|A|}^{-1} \leqslant 1$. Posons $M = \binom{n}{\lfloor n/2\rfloor}$. Par l'exercice 28 de la feuille du chapitre 3, $\binom{n}{k} \leqslant M$ pour tout $k \in [\![0\,,n]\!]$, donc chaque terme de la somme est au moins égal à $\dfrac{1}{M}$, et $\dfrac{|\mathcal{A}|}{M} \leqslant 1$. La famille des parties à $\lfloor \dfrac{n}{2}\rfloor$ éléments est une antichaîne de cardinal M (question 1) : la borne est atteinte.

$$\boxed{\text{Le cardinal maximal d'une antichaîne de parties de } [\![1\,,n]\!] \text{ est } \binom{n}{\lfloor n/2\rfloor}}$$

Une recherche exhaustive parmi toutes les familles de parties confirme les maxima 1, 2, 3 et 6 pour $n = 1$, 2, 3, 4.

À retenir. La démonstration de Lubell, Yamamoto et Meshalkin est un double comptage : on compte les couples (chaîne, élément de l'antichaîne situé sur la chaîne), une fois selon la chaîne, chacune portant au plus un élément, une fois selon l'élément. On peut la lire en probabilités : si l'on choisit une chaîne complète au hasard, sous la probabilité uniforme, elle passe par A avec probabilité $\binom{n}{|A|}^{-1}$, et les événements correspondants sont deux à deux incompatibles ; leur somme est donc au plus 1.

EXERCICE 21 ⋆⋆⋆

1. Un chemin N/E de $(0,0)$ à (a,b) fait a pas E et b pas N, puisque chaque pas augmente de 1 l'une des coordonnées ; réciproquement, tout mot formé de a lettres E et b lettres N code un tel chemin. Un tel mot est déterminé par l'ensemble des positions de ses lettres E, partie à a éléments de $[\![1\,,a+b]\!]$.

$$\boxed{\binom{a+b}{a} \text{ chemins N/E de } (0,0) \text{ à } (a,b)}$$

2. Pour $n \in \mathbb{N}$, on écrit les coefficients avec des factorielles :

$$\frac{1}{2n+1}\binom{2n+1}{n} = \frac{(2n+1)!}{(2n+1)\,n!\,(n+1)!} = \frac{(2n)!}{n!\,(n+1)!} = \frac{1}{n+1}\cdot\frac{(2n)!}{n!\,n!}.$$

$$\boxed{\frac{1}{n+1}\binom{2n}{n} = \frac{1}{2n+1}\binom{2n+1}{n}}$$

3.(a) On étend la définition de ρ_i à tout $i \in \mathbb{Z}$, les indices étant lus modulo $2n+1$. Alors $\rho_i(\rho_j(w)) = \rho_{i+j}(w)$, puisque la lettre d'indice k des deux mots est w_{i+j+k}, et $\rho_i = \rho_j$ si $i \equiv j \pmod{2n+1}$. La classe de w est $\{\rho_i(w) \mid i \in [\![0\,,2n]\!]\}$; elle a $2n+1$ éléments si et seulement si les mots $\rho_i(w)$, pour $i \in [\![0\,,2n]\!]$, sont deux à deux distincts. Or $\rho_i(w) = \rho_j(w)$ entraîne $\rho_{j-i}(w) = w$ en appliquant ρ_{-i}.

Supposons qu'il existe $d \in [\![1\,,2n]\!]$ tel que $\rho_d(w) = w$, et choisissons le plus petit. Écrivons la division euclidienne $2n+1 = qd+r$, avec $0 \leqslant r < d$ (chapitre 12, théorème 1). Comme $\rho_{2n+1}(w) = w$ et $\rho_{qd}(w) = w$, on obtient $\rho_r(w) = w$, et la minimalité de d impose $r = 0$: d divise $2n+1$, et $q \geqslant 2$. La relation $w_{k+d} = w_k$ pour tout k montre que w est formé de q copies du bloc $(w_0, \ldots, w_{d-1})$. Si ce bloc contient e lettres E et ν lettres N, alors $n+1 = qe$ et $n = q\nu$: l'entier $q \geqslant 2$ divise $n+1$ et n, donc leur différence 1, ce qui est absurde.

$$\boxed{\text{Chaque classe possède exactement } 2n+1 \text{ éléments}}$$

C'est ici que sert le choix de $n+1$ lettres E et n lettres N : avec autant de lettres de chaque sorte, le mot $ENEN$ n'aurait que deux rotations distinctes.

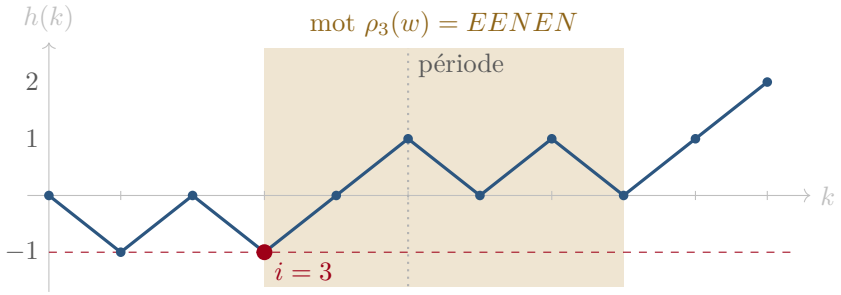
3.(b) Pour $k \in \mathbb{N}$, on note $h(k)$ le nombre de lettres E moins le nombre de lettres N parmi les k premières lettres du mot infini $W = www\cdots$ obtenu en répétant w ; ainsi $h(0) = 0$, $h(k+1) - h(k) = \pm 1$, et, comme w a une lettre E de plus que de lettres N, $h(k+2n+1) = h(k)+1$. Le préfixe de longueur $k+1$ du mot $\rho_i(w)$ est formé des lettres d'indices i à $i+k$ de W, et son nombre de E moins son nombre de N vaut $h(i+k+1) - h(i)$. Le mot $\rho_i(w)$ a donc la propriété demandée si et seulement si

$$h(j) > h(i) \quad \text{pour tout } j \in [\![i+1\,, i+2n+1]\!]. \tag{$*$}$$

▷ *Existence.* Soit m le minimum de h sur $[\![0\,, 2n]\!]$, et i le *plus grand* indice de $[\![0\,, 2n]\!]$ où il est atteint. Pour $j \in [\![i+1\,, 2n]\!]$, on a $h(j) \geqslant m$ et $h(j) \neq m$, donc $h(j) > h(i)$. Pour $j \in [\![2n+1\,, i+2n+1]\!]$, on a $h(j) = h(j-2n-1)+1 \geqslant m+1 > h(i)$. Ainsi $(*)$ est vérifiée, et $\rho_i(w)$ a la propriété.

▷ *Unicité.* Soit $i < i'$ deux indices de $[\![0\,, 2n]\!]$ vérifiant $(*)$. Comme $i' \in [\![i+1\,, i+2n]\!]$, la propriété en i donne $h(i') > h(i)$. Comme $i+2n+1 \in [\![i'+1\,, i'+2n+1]\!]$, la propriété en i' donne $h(i+2n+1) > h(i')$, c'est-à-dire $h(i)+1 > h(i')$. On aurait $h(i) < h(i') < h(i)+1$ entre entiers, ce qui est impossible. Par la question (a), les $2n+1$ rotations sont des mots distincts, et l'unicité de l'indice est celle du mot.

Chaque classe contient un unique mot dont tous les préfixes ont strictement plus de E que de N



La figure trace h pour $n = 2$ et $w = NENEE$, sur deux périodes séparées par les pointillés. Le minimum -1 sur $[\![0\,, 4]\!]$ est atteint en 1 et en 3 ; à partir du dernier, $i = 3$, la courbe reste strictement au-dessus de la droite en tirets sur toute la fenêtre ocre de longueur 5. On lit le mot $\rho_3(w) = EENEN$, dont tous les préfixes ont plus de E que de N. Partir de $i = 1$ échouerait, la courbe revenant à -1 en $k = 3$.

3.(c) Notons $\mathcal{G}$ l'ensemble des mots de $\mathcal{M}$ dont tous les préfixes ont strictement plus de E que de N. Les classes d'équivalence forment une partition de $\mathcal{M}$ (chapitre 2, théorème 4), chacune a $2n+1$ éléments, et $|\mathcal{M}| = \binom{2n+1}{n}$ par la question 1 : il y a $\dfrac{1}{2n+1}\binom{2n+1}{n}$ classes (lemme des bergers), et autant de mots dans $\mathcal{G}$ par la question (b).

Un mot $w \in \mathcal{G}$ commence par E, puisque son préfixe de longueur 1 doit contenir plus de E que de N ; on écrit $w = E\,v$, où v a n lettres E et n lettres N. Pour $k \in [\![1\,, 2n]\!]$, le préfixe de longueur $k+1$ de w a strictement plus de E que de N si et seulement si le préfixe de longueur k de v a au moins autant de E que de N, c'est-à-dire si le chemin codé par v est, après k pas, en un point (x, y) tel que $x \geqslant y$. L'application $w \mapsto v$ est donc une bijection de $\mathcal{G}$ sur l'ensemble des chemins N/E de $(0,0)$ à (n,n) qui restent dans S, de réciproque $v \mapsto Ev$. Avec la question 2 :

$$\boxed{C_n = |\mathcal{G}| = \frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{n+1}\binom{2n}{n}}$$

On trouve C_0, …, C_5 égaux à 1, 1, 2, 5, 14, 42, valeurs confirmées par énumération des chemins ; le lemme des rotations a lui aussi été vérifié classe par classe pour $n \leqslant 5$.

À retenir. L'argument des rotations est un lemme des bergers : on ne sait pas compter directement les bons chemins, mais on sait compter les classes, et chaque classe en contient exactement un. La division par $2n+1$ est légitime parce que toutes les classes ont la même taille, ce que garantit la question (a) ; c'est l'étape que l'on oublie le plus souvent.

EXERCICE 22 ⋆⋆⋆

1. Les orbites d'une permutation sont deux à deux disjointes (chapitre 17, théorème 1). Deux orbites distinctes à k éléments contiendraient ensemble $2k \geqslant 2n+2$ éléments de $[\![1\,,2n]\!]$, ce qui est impossible : une permutation a au plus une orbite à k éléments.

Une permutation σ qui a une orbite O à k éléments est déterminée par le triplet formé de O, de la restriction de σ à O, qui est un k-cycle de support O, et de la restriction de σ à $\overline{O}$, qui est une permutation de $\overline{O}$; réciproquement, tout tel triplet définit une permutation qui a O pour orbite. Il y a $\binom{2n}{k}$ choix de O, $(k-1)!$ cycles de support O (exercice 19) et $(2n-k)!$ permutations de $\overline{O}$.

$$\boxed{\binom{2n}{k}(k-1)!\,(2n-k)! = \frac{(2n)!}{k} \text{ permutations ont une orbite à } k \text{ éléments}}$$

2. Les ensembles de permutations ayant une orbite à k éléments, pour $k \in [\![n+1\,,2n]\!]$, sont deux à deux disjoints : deux orbites à k et k' éléments, avec $k, k' > n$, contiendraient plus de $2n$ éléments. Une permutation a toutes ses orbites à au plus n éléments si et seulement si elle n'appartient à aucun de ces ensembles, et la proposition 2 (i) et (ii) donne

$$a_n = (2n)! - \sum_{k=n+1}^{2n} \frac{(2n)!}{k}, \qquad \text{d'où} \qquad \boxed{\frac{a_n}{(2n)!} = 1 - \sum_{k=n+1}^{2n} \frac{1}{k} = 1 - (H_{2n} - H_n)}$$

où $H_m = \sum_{k=1}^{m} \frac{1}{k}$. Par la proposition 7 du chapitre 20, on a $H_m = \ln m + \gamma + o(1)$, donc

$$H_{2n} - H_n = \ln(2n) - \ln n + o(1) = \ln 2 + o(1),$$

comme le note la remarque qui suit cette proposition.

$$\boxed{\lim_{n\to+\infty} \frac{a_n}{(2n)!} = 1 - \ln 2 \approx 0{,}3069}$$

Pour $n = 1$, 2, 3, 4, la formule donne $\frac{1}{2}$, $\frac{5}{12}$, $\frac{23}{60}$ et $\frac{307}{840} \approx 0{,}365$, valeurs confirmées par énumération de $\mathfrak{S}_2$, $\mathfrak{S}_4$, $\mathfrak{S}_6$ et $\mathfrak{S}_8$. Pour $n = 1$, seule l'identité convient parmi les deux permutations de $[\![1\,,2]\!]$ ✓.

3. On choisit une permutation σ au hasard : $\Omega = \mathfrak{S}_{2n}$, muni de la probabilité uniforme. La probabilité que toutes les orbites de σ aient au plus n éléments vaut $\frac{a_n}{(2n)!}$, et la question 2 montre que la probabilité que σ ait une orbite à k éléments vaut $\frac{1}{k}$, pour chaque $k > n$.

$$\boxed{\mathrm{P}(\text{toutes les orbites de } \sigma \text{ ont au plus } n \text{ éléments}) \xrightarrow[n\to+\infty]{} 1 - \ln 2 \approx 0{,}31}$$

Quel que soit n, une permutation de $2n$ éléments choisie au hasard a donc, avec probabilité voisine de $\ln 2 \approx 0{,}69$, une orbite qui contient plus de la moitié des éléments : une « grande » orbite est la règle, et non l'exception. Pour $2n = 100$, la probabilité de n'en avoir aucune vaut $1-(H_{100}-H_{50}) \approx 0{,}312$. C'est le ressort de l'énigme dite des cent prisonniers, où une stratégie qui suit les orbites d'une permutation aléatoire réussit avec cette probabilité, alors que des choix indépendants n'auraient quasiment aucune chance. L'exercice 66 précisera la loi de la longueur de l'orbite d'un élément donné.

Piège. L'erreur la plus répandue est d'appliquer le même raisonnement aux orbites de taille $k \leqslant n$: une permutation peut en avoir plusieurs, les ensembles correspondants ne sont plus disjoints, et la somme des $\frac{(2n)!}{k}$ ne compte plus rien de simple. C'est l'hypothèse $k > n$ qui rend les événements incompatibles et le calcul exact.

EXERCICE 23 ⋆

1. La formule de la proposition 7 (iv) donne d'abord la probabilité de l'intersection :

$$\mathrm{P}(A\cap B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cup B)=\frac{7}{10}+\frac{5}{10}-\frac{9}{10}=\frac{3}{10}.$$

L'événement « l'un exactement des deux est réalisé » est $(A\cup B)\setminus(A\cap B)$, et $A\cap B\subset A\cup B$; la proposition 7 (ii) donne

$$\boxed{\mathrm{P}\big((A\cup B)\setminus(A\cap B)\big)=\mathrm{P}(A\cup B)-\mathrm{P}(A\cap B)=\frac{9}{10}-\frac{3}{10}=\frac{3}{5}}$$

On le retrouve en écrivant cet événement comme la réunion des événements incompatibles $A\cap\overline{B}$ et $\overline{A}\cap B$, de probabilités $\frac{7}{10}-\frac{3}{10}=\frac{2}{5}$ et $\frac{5}{10}-\frac{3}{10}=\frac{1}{5}$ ✓.

2. Les quatre événements $A\cap B$, $A\cap\overline{B}$, $\overline{A}\cap B$ et $\overline{A}\cap\overline{B}$ forment un système complet d'événements ; on note x, y, z, w leurs probabilités, de sorte que $x+y+z+w=1$ (proposition 7 (v)). Le contraire de chacune des quatre réunions est l'un de ces quatre « atomes », par les lois de De Morgan : par exemple, le contraire de $A\cup\overline{B}$ est $\overline{A}\cap B$. Par la proposition 7 (ii),

$$\mathrm{P}(A\cup B)+\mathrm{P}(A\cup\overline{B})+\mathrm{P}(\overline{A}\cup B)+\mathrm{P}(\overline{A}\cup\overline{B})=(1-w)+(1-z)+(1-y)+(1-x)=4-1,$$

$$\boxed{\mathrm{P}(A\cup B)+\mathrm{P}(A\cup\overline{B})+\mathrm{P}(\overline{A}\cup B)+\mathrm{P}(\overline{A}\cup\overline{B})=3}$$

Chaque atome est contenu dans exactement trois des quatre réunions : la somme compte trois fois la probabilité totale, ce qui est une autre façon de lire le résultat.

3. Comme $A\cap B$ est inclus dans A et dans B, la croissance (proposition 7 (iii)) donne $\mathrm{P}(A\cap B)\leqslant p$ et $\mathrm{P}(A\cap B)\leqslant q$. Une probabilité est positive, et $\mathrm{P}(A\cup B)\leqslant 1$, donc

$$\mathrm{P}(A\cap B)=p+q-\mathrm{P}(A\cup B)\geqslant p+q-1.$$

$$\boxed{\max(0,p+q-1)\leqslant\mathrm{P}(A\cap B)\leqslant\min(p,q)}$$

Réciproquement, soit t dans cet intervalle. On prend $\Omega=\{1,2,3,4\}$ et les quatre nombres

$$p_1=t,\qquad p_2=p-t,\qquad p_3=q-t,\qquad p_4=1-p-q+t.$$

Ils sont positifs exactement parce que $t\leqslant p$, $t\leqslant q$ et $t\geqslant p+q-1$, et leur somme vaut 1 : c'est une distribution de probabilités, et le théorème 4 fournit une probabilité P sur Ω telle que $\mathrm{P}(\{i\})=p_i$. Avec $A=\{1,2\}$ et $B=\{1,3\}$, on obtient $\mathrm{P}(A)=p$, $\mathrm{P}(B)=q$ et $\mathrm{P}(A\cap B)=t$.

$$\boxed{\text{toute valeur } t\in\big[\max(0,p+q-1)\,,\min(p,q)\big] \text{ est atteinte}}$$

	A, probabilité p	$\overline{A}$
B, probabilité q	$A\cap B:\ t$	$\overline{A}\cap B:\ q-t$
$\overline{B}$	$A\cap\overline{B}:\ p-t$	$\overline{A}\cap\overline{B}:\ 1-p-q+t$

Les quatre atomes déterminent tout : la probabilité t de la case pétrole fixe les trois autres, et la seule contrainte est que les quatre cases aient une probabilité positive. C'est exactement l'encadrement obtenu, que l'on retrouvera à l'exercice 55 pour la loi conjointe de deux variables de Bernoulli.

Piège. L'erreur la plus répandue est d'écrire $\mathrm{P}(A\cap B)=pq$: c'est l'indépendance, qui n'est qu'une valeur particulière de l'intervalle, et l'énoncé ne la suppose pas. Les bornes correspondent aux cas extrêmes : $t=\min(p,q)$ lorsque l'un des événements est (presque sûrement) inclus dans l'autre, $t=\max(0,p+q-1)$ lorsqu'ils se chevauchent le moins possible.

EXERCICE 24 ⋆

1. On cherche $\mathrm{P}(\{k\}) = ck$, avec $c > 0$. D'après le théorème 4, une telle famille définit une probabilité si et seulement si ses termes sont positifs et de somme 1 ; la somme vaut $c\dfrac{n(n+1)}{2}$ (chapitre 3), d'où l'unique valeur de c :

$$\boxed{\mathrm{P}(\{k\}) = \frac{2k}{n(n+1)} \qquad \text{pour } k \in [\![1\,,n]\!]}$$

Pour $n = 3$, on trouve les probabilités $\dfrac{1}{6}, \dfrac{2}{6}, \dfrac{3}{6}$, de somme 1 ✓.

2. On cherche $\mathrm{P}([\![1\,,k]\!]) = ck^2$ pour tout k. Avec $k = n$, il vient $c = \dfrac{1}{n^2}$. L'intervalle $[\![1\,,k]\!]$ est la réunion disjointe de $[\![1\,,k-1]\!]$ et de $\{k\}$, avec $[\![1\,,0]\!] = \varnothing$, de probabilité nulle, ce qui est cohérent avec la formule $c \times 0^2$. Donc, nécessairement,

$$\mathrm{P}(\{k\}) = \frac{k^2 - (k-1)^2}{n^2}, \qquad \boxed{\mathrm{P}(\{k\}) = \frac{2k-1}{n^2} \qquad \text{pour } k \in [\![1\,,n]\!]}$$

Réciproquement, ces nombres sont positifs et leur somme télescopique vaut $\dfrac{n^2}{n^2} = 1$; la probabilité qu'ils définissent (théorème 4) vérifie $\mathrm{P}([\![1\,,k]\!]) = \dfrac{k^2}{n^2}$ par le même télescopage. On a ainsi retrouvé que la somme des k premiers nombres impairs vaut k^2.

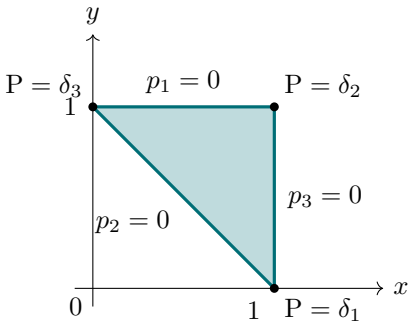
3. Notons $p_i = \mathrm{P}(\{i\})$. Par le théorème 4, une probabilité sur $\{1,2,3\}$ est la donnée de trois réels positifs de somme 1, et les conditions s'écrivent

$$p_1 + p_2 = x, \qquad p_2 + p_3 = y, \qquad p_1 + p_2 + p_3 = 1.$$

Ce système a pour unique solution $p_1 = 1 - y$, $p_3 = 1 - x$ et $p_2 = x + y - 1$. La probabilité cherchée existe si et seulement si ces trois nombres sont positifs :

$$\boxed{x \leqslant 1, \qquad y \leqslant 1, \qquad x + y \geqslant 1}$$

et elle est alors unique. Ces conditions entraînent $x \geqslant 1 - y \geqslant 0$ et de même $y \geqslant 0$: il n'est pas utile de les ajouter.



Les couples (x, y) admissibles remplissent le triangle pétrole. Sur chaque côté, l'une des trois probabilités s'annule, et aux sommets la probabilité est concentrée sur un seul point ; on a noté δ_i la probabilité qui donne la probabilité 1 au singleton $\{i\}$.

EXERCICE 25 ⋆

On prend $\Omega = [\![1\,,6]\!]^3$, ensemble des triplets de résultats des dés pris dans l'ordre des lancers, muni de la probabilité uniforme : $|\Omega| = 216$. On compte les cas favorables.

— *Au moins un six.* Le contraire, « aucun six », correspond à $[\![1\,,5]\!]^3$, qui a 125 éléments (proposition 2 (iv)) : la probabilité vaut $1 - \dfrac{125}{216}$.

— *Exactement un six.* On choisit le rang du six, de 3 façons, puis les deux autres résultats dans $[\![1\,,5]\!]$, de 25 façons : 75 issues.

— *Au moins deux résultats égaux.* Le contraire, « trois résultats distincts », compte $A_6^3 = 120$ issues (théorème 1).

— *Au moins deux égaux et somme paire.* Les triplets constants (a, a, a) ont une somme $3a$ paire si et seulement si a est pair : 3 issues. Les triplets ayant exactement deux résultats égaux s'écrivent, à l'ordre près, a, a, b avec $b \neq a$; leur somme $2a + b$ est paire si et seulement si b est pair. On choisit b pair (3 façons), $a \neq b$ (5 façons) et le rang de b (3 façons) : 45 issues, soit 48 en tout.

$$\boxed{\mathrm{P}(\text{au moins un six}) = \frac{91}{216}, \qquad \mathrm{P}(\text{un seul six}) = \frac{25}{72}}$$

$$\boxed{\mathrm{P}(\text{deux résultats égaux}) = \frac{4}{9}, \qquad \mathrm{P}(\text{deux résultats égaux, somme paire}) = \frac{2}{9}}$$

Une énumération des 216 issues confirme ces quatre valeurs ✓. Parmi les issues ayant deux résultats égaux, la moitié exactement ont une somme paire (48 sur 96).

Piège. Pour « au moins un six », l'erreur la plus répandue est d'additionner les probabilités $\dfrac{1}{6}$ des trois lancers, ce qui donne $\dfrac{1}{2}$: les événements « le i-ième lancer donne six » ne sont pas incompatibles, et $\dfrac{1}{2}$ n'est qu'un majorant (sous-additivité, proposition 7 (iv)). On passe par le contraire, comme le recommande la méthode de la section 21.2.6.

EXERCICE 26 ⋆

On note T l'événement « exactement deux trèfles et deux piques ». Dans les trois cas, l'univers est muni de la probabilité uniforme.

Tirage simultané. L'univers est l'ensemble des parties à 4 éléments du jeu, qui en compte $\dbinom{52}{4}$, soit 270 725. Une main favorable est formée de deux des treize trèfles et de deux des treize piques :

$$\mathrm{P}(T) = \frac{\dbinom{13}{2}^2}{\dbinom{52}{4}} = \frac{6\,084}{270\,725}, \qquad \boxed{\mathrm{P}(T) = \frac{468}{20\,825} \approx 0{,}0225}$$

Tirage successif sans remise. L'univers est l'ensemble des 4-listes de cartes distinctes, de cardinal $A_{52}^4 = 6\,497\,400$ (théorème 1). On choisit les deux rangs des trèfles, de $\dbinom{4}{2} = 6$ façons, puis la liste ordonnée des deux trèfles, de 13×12 façons, et celle des deux piques, de 13×12 façons :

$$\boxed{\mathrm{P}(T) = \frac{6 \times 156^2}{A_{52}^4} = \frac{146\,016}{6\,497\,400} = \frac{468}{20\,825}}$$

Tirage successif avec remise. L'univers est l'ensemble des 4-listes de cartes, de cardinal 52^4 (proposition 3). On choisit les rangs des trèfles, puis chacune des quatre cartes dans sa couleur :

$$\boxed{\mathrm{P}(T) = \frac{6 \times 13^4}{52^4} = \frac{6}{4^4} = \frac{3}{128} \approx 0{,}0234}$$

Pourquoi les deux premiers résultats sont égaux. L'application qui oublie l'ordre envoie chaque 4-liste de cartes distinctes sur la partie formée de ses termes, et chaque partie est l'image de exactement $4! = 24$ listes : c'est l'argument du lemme des bergers dans la démonstration du théorème 2. Elle envoie les listes favorables sur les mains favorables, puisque l'événement T ne dépend pas de l'ordre. Les cas favorables et les cas possibles sont donc tous deux multipliés par 24 : $146\,016 = 24 \times 6\,084$ et $6\,497\,400 = 24 \times 270\,725$ ✓. Avec remise, les tirages ne sont plus équivalents : une carte peut sortir deux fois, et les couleurs ne s'épuisent pas, ce qui rend T un peu plus probable.
À retenir. Pour un événement qui ne dépend pas de l'ordre, on peut toujours choisir le modèle ordonné ou le modèle non ordonné, à condition de rester cohérent : compter les cas favorables avec ordre et les cas possibles sans ordre est l'erreur la plus répandue, et donne ici un résultat 24 fois trop grand.

EXERCICE 27 ⋆

L'univers est l'ensemble des parties à deux éléments de l'ensemble des $n+5$ boules, supposées discernables, muni de la probabilité uniforme ; il a $\binom{n+5}{2} = \dfrac{(n+5)(n+4)}{2}$ éléments.

1. Deux boules de la même couleur forment une paire de rouges, de noires ou de blanches, ces trois événements étant incompatibles : il y a $\binom{3}{2} + \binom{2}{2} + \binom{n}{2} = 4 + \dfrac{n(n-1)}{2}$ cas favorables, et

$$\boxed{P_n = \frac{n^2-n+8}{(n+4)(n+5)}, \qquad Q_n = 1 - P_n = \frac{2(5n+6)}{(n+4)(n+5)}}$$

On contrôle Q_n directement : une paire bicolore est formée d'une rouge et d'une noire, d'une rouge et d'une blanche, ou d'une noire et d'une blanche, d'où $6 + 3n + 2n = 5n + 6$ cas favorables, à diviser par $\dfrac{(n+4)(n+5)}{2}$ ✓.

2. La probabilité de tirer deux blanches vaut $\dfrac{n(n-1)}{(n+4)(n+5)}$. Elle vaut $\dfrac{1}{6}$ si et seulement si $6n^2 - 6n = n^2 + 9n + 20$, soit $5n^2 - 15n - 20 = 0$, c'est-à-dire $(n-4)(n+1) = 0$. Seule la racine $n = 4$ est un entier naturel :

$$\boxed{n = 4, \qquad P_4 = \frac{16-4+8}{8\times 9} = \frac{5}{18}}$$

Avec 9 boules, il y a 36 paires, dont $3 + 1 + 6 = 10$ unicolores ✓. Lorsque n tend vers $+\infty$, P_n tend vers 1 : presque toutes les boules sont blanches.

EXERCICE 28 ⋆⋆

Une issue est un couple (R, H), où R est la partie à 10 éléments des cartes retirées et H la partie à 5 éléments des cartes tirées, avec $H \cap R = \varnothing$. Pour tout R, il y a $\binom{42}{5}$ choix de H ; les deux choix étant faits au hasard, on munit l'ensemble Ω de ces couples de la probabilité uniforme, avec $|\Omega| = \binom{52}{10}\binom{42}{5}$ (lemme des bergers).

1. Notons a l'as de trèfle. Un couple favorable s'obtient en choisissant R parmi les 51 autres cartes, puis H contenant a, c'est-à-dire ses quatre autres cartes parmi les 41 cartes restantes autres que a. À l'aide des relations $\binom{51}{10} = \dfrac{42}{52}\binom{52}{10}$ et $\binom{41}{4} = \dfrac{5}{42}\binom{42}{5}$,

$$\mathrm{P}(a \in H) = \frac{\binom{51}{10}\binom{41}{4}}{\binom{52}{10}\binom{42}{5}} = \frac{42}{52} \times \frac{5}{42}, \qquad \boxed{\mathrm{P}(a \in H) = \frac{5}{52}}$$

C'est la probabilité que l'as de trèfle figure dans une main de cinq cartes tirées directement dans le jeu de 52 cartes. On pouvait le prévoir : les dix cartes retirées sans être regardées n'apportent aucune information, et l'on ne fait que « tirer cinq cartes au hasard » de façon détournée. La question suivante rend cet argument rigoureux.

2. Soit H une partie à cinq éléments du jeu. Les couples (R, H) qui conduisent à H correspondent aux parties R à dix éléments du complémentaire de H, qui a 47 éléments : il y en a $\binom{47}{10}$, nombre qui ne dépend pas de H. Ainsi

$$\mathrm{P}(\text{la main tirée est } H) = \frac{\binom{47}{10}}{\binom{52}{10}\binom{42}{5}} = \frac{1}{\binom{52}{5}},$$

car, en simplifiant les factorielles, $\binom{52}{10}\binom{42}{5} = \dfrac{52!}{10!\,5!\,37!} = \binom{52}{5}\binom{47}{10}$.

la main tirée suit la loi uniforme sur les parties à 5 éléments du jeu de 52 cartes

Les deux membres de l'identité comptent de deux façons les couples formés de dix cartes et de cinq autres cartes. Une énumération exhaustive sur un jeu de 9 cartes, dont on retire 3 avant d'en tirer 2, confirme que chacune des 36 paires est obtenue par exactement $\binom{7}{3} = 35$ couples ✓.

3. On calcule donc dans le modèle de la main de cinq cartes parmi 52, avec $\binom{52}{5} = 2\,598\,960$ issues. Une hauteur ne comporte que quatre cartes, donc cinq cartes ont au moins deux hauteurs ; elles en ont exactement deux lorsque la main est un carré (quatre cartes d'une hauteur, une d'une autre) ou un full (trois et deux). On choisit la hauteur du carré, puis la cinquième carte parmi les 48 autres : $13 \times 48 = 624$ carrés ; on choisit la hauteur du brelan et ses trois cartes, puis celle de la paire et ses deux cartes : $13 \times 4 \times 12 \times 6 = 3\,744$ fulls, comme dans l'exemple de la section 21.2.6. Pour cinq hauteurs distinctes, on choisit les cinq hauteurs, puis une couleur pour chacune :

$$\mathrm{P}(\text{deux hauteurs}) = \frac{624 + 3\,744}{2\,598\,960} = \frac{1}{595}, \qquad \mathrm{P}(\text{cinq hauteurs}) = \frac{\binom{13}{5}4^5}{\binom{52}{5}} = \frac{2\,112}{4\,165} \approx 0{,}507$$

Une énumération des 2 598 960 mains confirme les nombres 4 368 et 1 317 888 de cas favorables ✓. Un peu plus d'une main sur deux ne contient aucune paire.

À retenir. L'information qu'on ne connaît pas ne change pas les probabilités : retirer des cartes sans les regarder, ou couper le paquet, ne modifie pas la loi de la main. La question 2 en donne une preuve par le lemme des bergers ; l'erreur la plus répandue est de calculer dans le modèle compliqué des couples (R, H), alors qu'une fois l'uniformité établie, le modèle simple suffit.

EXERCICE 29 ⋆⋆

L'univers est l'ensemble des ordres des 52 cartes, c'est-à-dire des bijections de $[\![1\,,52]\!]$ (les positions) sur l'ensemble des cartes, muni de la probabilité uniforme ; il a 52! éléments (théorème 1).

1. Le contraire de l'événement étudié est « deux cartes voisines sont toujours de couleurs différentes » : les couleurs alternent, et la suite des couleurs est rouge, noire, rouge, … ou noire, rouge, noire, … Pour chacun de ces deux motifs, on place les 26 cartes rouges dans les 26 positions qui leur sont réservées, de 26! façons, et les noires de même. Ainsi

$$\boxed{\mathrm{P}(\text{deux cartes de même couleur côte à côte}) = 1 - \frac{2\,(26!)^2}{52!} = 1 - \frac{2}{\binom{52}{26}}}$$

avec $2/\binom{52}{26} \approx 4{,}0 \times 10^{-15}$. L'événement est presque certain : une alternance parfaite des couleurs dans un jeu bien mélangé a une chance sur 250 milliers de milliards. Sur un modèle réduit de six cartes, trois rouges et trois noires, la formule donne $1 - \dfrac{72}{720} = \dfrac{9}{10}$, valeur confirmée par énumération des 720 ordres ✓.

2. À chaque ordre, on associe l'ensemble K des positions des quatre rois, partie à quatre éléments de $[\![1\,,52]\!]$. Pour une partie K donnée, il y a $4!\,48!$ ordres qui lui correspondent (on place les quatre rois dans les positions de K, puis les 48 autres cartes) : ce nombre ne dépend pas de K, et le lemme des bergers montre que K suit la loi uniforme sur les $\binom{52}{4}$ parties à quatre éléments. Il suffit donc de compter les parties sans deux entiers consécutifs. Pour $a_1 < a_2 < a_3 < a_4$, on pose $b_i = a_i - (i-1)$:

$$a_{i+1} - a_i \geqslant 2 \ \text{ pour tout } i \iff b_1 < b_2 < b_3 < b_4, \qquad \text{avec} \qquad 1 \leqslant b_1, \quad b_4 = a_4 - 3 \leqslant 49.$$

L'application $\{a_1, \ldots, a_4\} \mapsto \{b_1, \ldots, b_4\}$ est donc une bijection de l'ensemble des parties à quatre éléments de $[\![1\,,52]\!]$ sans entiers consécutifs sur l'ensemble des parties à quatre éléments de $[\![1\,,49]\!]$, de réciproque $b_i \mapsto b_i + i - 1$. Il y en a $\binom{49}{4}$, et

$$\boxed{\mathrm{P}(\text{deux rois côte à côte}) = 1 - \frac{\binom{49}{4}}{\binom{52}{4}} = 1 - \frac{211\,876}{270\,725} = \frac{1\,201}{5\,525} \approx 0{,}217}$$

Sur un jeu de 8 cartes contenant 3 rois, la même méthode donne $1 - \binom{6}{3}/\binom{8}{3} = \dfrac{9}{14}$, valeur confirmée par énumération des 8! ordres ✓. La même bijection montre qu'il y a $\binom{n-r+1}{r}$ parties à r éléments de $[\![1\,,n]\!]$ sans entiers consécutifs ; l'exercice 9 les avait dénombrées toutes tailles confondues.

Piège. Les deux questions se ressemblent, mais leurs réponses sont opposées : il est presque certain que deux cartes de même couleur se touchent, et peu probable que deux rois se touchent. L'erreur la plus répandue est de travailler directement sur les 52! ordres pour la question 2 ; le passage à l'ensemble des positions, dont la loi est uniforme, réduit le problème à un dénombrement de parties.

EXERCICE 30 ⋆⋆

L'univers est $\Omega = \{\mathrm{P},\mathrm{F}\}^{2n+1}$, ensemble des suites de résultats, muni de la probabilité uniforme puisque la pièce est équilibrée : $|\Omega| = 2^{2n+1}$.

1. L'application s qui échange pile et face dans chaque lancer est une bijection de Ω sur lui-même, égale à sa réciproque. Elle échange l'événement G « strictement plus de piles que de faces » et l'événement G' « strictement plus de faces que de piles », qui ont donc le même cardinal (proposition 1 (i)). Comme $2n+1$ est impair, il ne peut y avoir autant de piles que de faces : G et G' forment un système complet, et $\mathrm{P}(G) + \mathrm{P}(G') = 1$.

$$\boxed{\mathrm{P}(G) = \frac{1}{2}}$$

2. Le $(n+1)$-ième pile apparaît au k-ième lancer si et seulement si le k-ième lancer donne pile et si les $k-1$ premiers en donnent exactement n. On choisit les rangs de ces n piles parmi les $k-1$ premiers, de $\binom{k-1}{n}$ façons (théorème 2), puis les $2n+1-k$ derniers résultats librement :

$$\boxed{\mathrm{P}(\text{le } (n+1)\text{-ième pile apparaît au } k\text{-ième lancer}) = \frac{\binom{k-1}{n} 2^{2n+1-k}}{2^{2n+1}} = \frac{1}{2^k}\binom{k-1}{n}}$$

Cette probabilité est nulle pour $k \leqslant n$, ce qui est attendu : il faut au moins $n+1$ lancers pour obtenir $n+1$ piles.

3. L'événement G signifie « au moins $n+1$ piles », c'est-à-dire « le $(n+1)$-ième pile apparaît au cours des $2n+1$ lancers ». C'est la réunion des événements de la question 2 pour $k \in [\![n+1\,,2n+1]\!]$, qui sont deux à deux incompatibles puisque le rang du $(n+1)$-ième pile est unique. La proposition 7 (i) et la question 1 donnent

$$\boxed{\sum_{k=n+1}^{2n+1} \frac{1}{2^k}\binom{k-1}{n} = \mathrm{P}(G) = \frac{1}{2}}$$

Pour $n=1$, on trouve $\frac{1}{4} + \frac{2}{8} = \frac{1}{2}$ ✓ ; l'identité a été vérifiée par énumération jusqu'à $n=5$.

À retenir. C'est un exemple de démonstration probabiliste d'une identité : on calcule de deux façons la probabilité d'un même événement. Une démonstration purement calculatoire de cette identité n'est pas immédiate ; ici, la symétrie de la question 1 fait tout le travail.

EXERCICE 31 ⋆⋆

1. Notons E l'ensemble des $2n$ équipes et $\mathcal{R}$ l'ensemble des répartitions, c'est-à-dire des partitions de E en n paires. À toute bijection $\sigma : [\![1\,,2n]\!] \to E$, on associe la répartition formée des paires $\{\sigma(2i-1), \sigma(2i)\}$ pour $i \in [\![1\,,n]\!]$. Toute répartition est obtenue, et les bijections qui donnent une répartition fixée s'obtiennent en choisissant l'ordre des n matchs, de $n!$ façons, puis l'ordre des deux équipes dans chaque match, de 2^n façons. Par le lemme des bergers,

$$\boxed{|\mathcal{R}| = \frac{(2n)!}{2^n\, n!} = 1 \times 3 \times 5 \times \cdots \times (2n-1)}$$

On trouve 1, 3, 15, 105 répartitions pour $n = 1$, 2, 3, 4, valeurs confirmées par énumération ✓. L'univers est $\mathcal{R}$, muni de la probabilité uniforme, comme le fixe l'énoncé.

2. Dans une répartition favorable, chaque équipe de première division rencontre une équipe de deuxième division, et deux équipes de première division n'ont pas le même adversaire : une répartition favorable est déterminée par une bijection de l'ensemble des équipes de première division sur celui des équipes de deuxième division, et réciproquement. Il y en a $n!$ (théorème 1), d'où

$$\boxed{p_n = \frac{n!\, 2^n\, n!}{(2n)!} = \frac{2^n}{\binom{2n}{n}}}$$

On trouve $p_1 = 1$, $p_2 = \dfrac{2}{3}$, $p_3 = \dfrac{2}{5}$, $p_4 = \dfrac{8}{35}$ ✓.

3. Si tous les matchs opposent deux équipes de la même division, les n équipes de première division sont réparties en paires, ce qui impose n pair. Pour $n = 2m$, une répartition favorable est formée d'une répartition des $2m$ équipes de première division et d'une répartition de celles de deuxième division, d'où $\left(\dfrac{(2m)!}{2^m m!}\right)^2$ cas favorables, et, en simplifiant,

$$\boxed{q_n = 0 \text{ si } n \text{ est impair}, \qquad q_{2m} = \left(\frac{(2m)!}{2^m m!}\right)^2 \frac{2^{2m}(2m)!}{(4m)!} = \frac{(2m)!^3}{m!^2\,(4m)!} = \frac{\binom{2m}{m}}{\binom{4m}{2m}}}$$

On trouve $q_2 = \dfrac{2}{6} = \dfrac{1}{3}$ et $q_4 = \dfrac{6}{70} = \dfrac{3}{35}$, valeurs confirmées par énumération ✓.

4. Pour $0 \leqslant k \leqslant 2n-1$, on a $\dbinom{2n}{k+1} \Big/ \dbinom{2n}{k} = \dfrac{2n-k}{k+1}$, qui est supérieur ou égal à 1 si et seulement si $k \leqslant n - \dfrac{1}{2}$, c'est-à-dire $k \leqslant n-1$: les coefficients croissent jusqu'à $\dbinom{2n}{n}$, puis décroissent, et $\dbinom{2n}{n}$ est le plus grand. La somme des $2n+1$ coefficients vaut 2^{2n} (théorème 3), donc $\dbinom{2n}{n} \leqslant 2^{2n}$. Pour la minoration, on regroupe les deux coefficients extrêmes, $\dbinom{2n}{0} + \dbinom{2n}{2n} = 2 \leqslant \dbinom{2n}{n}$ pour $n \geqslant 1$, et l'on majore chacun des $2n-1$ autres par $\dbinom{2n}{n}$:

$$2^{2n} = 2 + \sum_{k=1}^{2n-1} \binom{2n}{k} \leqslant \binom{2n}{n} + (2n-1)\binom{2n}{n} = 2n\binom{2n}{n}, \qquad \boxed{\frac{2^{2n-1}}{n} \leqslant \binom{2n}{n} \leqslant 2^{2n}}$$

On en déduit $p_n = \dfrac{2^n}{\binom{2n}{n}} \leqslant \dfrac{2^n \times n}{2^{2n-1}} = \dfrac{2n}{2^n}$. De même, pour $n = 2m$, la minoration au rang $2m$ et la majoration au rang m donnent $q_{2m} \leqslant \dfrac{2^{2m} \times 2m}{2^{4m-1}} = \dfrac{2n}{2^n}$. Comme $\dfrac{2n}{2^n}$ tend vers 0 (croissances comparées),

$$\boxed{0 \leqslant p_n \leqslant \frac{2n}{2^n}, \quad 0 \leqslant q_n \leqslant \frac{2n}{2^n}, \qquad \lim_{n\to+\infty} p_n = \lim_{n\to+\infty} q_n = 0}$$

L'équivalent $\dbinom{2n}{n} \sim \dfrac{4^n}{\sqrt{\pi n}}$, tiré de la formule de Stirling (chapitre 10, théorème 6, et chapitre 18, théorème 10), précise la vitesse : $p_n \sim \dfrac{\sqrt{\pi n}}{2^n}$. Pour $n = 10$, on a déjà $p_{10} \approx 0{,}0055$ et $q_{10} \approx 0{,}0014$: un tirage intégral « première contre deuxième division », ou sans aucun match entre divisions, est un événement très rare.

EXERCICE 32 ⋆⋆

1. Pour $n = 2$, la proposition 7 (iv) et $\mathrm{P}(A_1 \cup A_2) \leqslant 1$ donnent

$$\mathrm{P}(A_1 \cap A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 \cup A_2) \geqslant \mathrm{P}(A_1) + \mathrm{P}(A_2) - 1.$$

Supposons l'inégalité vraie pour n événements quelconques, et soit A_1, …, A_{n+1}. On applique le cas $n = 2$ aux deux événements $B = A_1 \cap \cdots \cap A_n$ et A_{n+1}, puis l'hypothèse de récurrence à B :

$$\mathrm{P}(B \cap A_{n+1}) \geqslant \mathrm{P}(B) + \mathrm{P}(A_{n+1}) - 1 \geqslant \sum_{i=1}^{n} \mathrm{P}(A_i) - (n-1) + \mathrm{P}(A_{n+1}) - 1.$$

$$\boxed{\mathrm{P}(A_1 \cap \cdots \cap A_n) \geqslant \mathrm{P}(A_1) + \cdots + \mathrm{P}(A_n) - (n-1)}$$

C'est la sous-additivité appliquée aux contraires : $\mathrm{P}(\overline{A_1} \cup \cdots \cup \overline{A_n}) \leqslant \sum (1 - \mathrm{P}(A_i))$. Elle sert lorsque les $\mathrm{P}(A_i)$ sont proches de 1 : si chacun de 10 composants fonctionne avec la probabilité 0,99, tous fonctionnent avec une probabilité au moins 0,9, sans aucune hypothèse d'indépendance.

2. (a) Soit m le nombre d'indices i tels que $\omega \in A_i$. La première somme vaut m, et la seconde compte les paires $\{i, j\}$ d'indices tels que $\omega \in A_i \cap A_j$, c'est-à-dire les paires d'indices parmi les m précédents : elle vaut $\binom{m}{2}$. Le second membre vaut 0 si $m = 0$ et 1 si $m \geqslant 1$. Pour $m = 0$, les deux membres sont nuls. Pour $m \geqslant 1$,

$$1 - \left(m - \frac{m(m-1)}{2}\right) = \frac{m^2 - 3m + 2}{2} = \frac{(m-1)(m-2)}{2} \geqslant 0,$$

puisqu'aucun entier n'est strictement compris entre 1 et 2.

$$\boxed{\sum_{i=1}^{n} \mathbf{1}_{A_i}(\omega) - \sum_{1 \leqslant i < j \leqslant n} \mathbf{1}_{A_i \cap A_j}(\omega) \leqslant \mathbf{1}_{A_1 \cup \cdots \cup A_n}(\omega)}$$

avec égalité si et seulement si ω appartient à au plus deux des A_i.

(b) Pour tout événement A, le théorème 4 donne $\mathrm{P}(A) = \sum_{\omega \in A} \mathrm{P}(\{\omega\}) = \sum_{\omega \in \Omega} \mathrm{P}(\{\omega\})\, \mathbf{1}_A(\omega)$. On multiplie l'inégalité de (a) par le réel positif $\mathrm{P}(\{\omega\})$, on somme sur $\omega \in \Omega$ et l'on intervertit les sommes finies (chapitre 3) :

$$\boxed{\mathrm{P}(A_1 \cup \cdots \cup A_n) \geqslant \sum_{i=1}^{n} \mathrm{P}(A_i) - \sum_{1 \leqslant i < j \leqslant n} \mathrm{P}(A_i \cap A_j)}$$

3. On fixe j et l'on montre l'inégalité ponctuelle

$$\mathbf{1}_{A_1 \cup \cdots \cup A_n}(\omega) \leqslant \sum_{i=1}^{n} \mathbf{1}_{A_i}(\omega) - \sum_{i \neq j} \mathbf{1}_{A_i \cap A_j}(\omega).$$

Avec m défini comme en (a) : si $\omega \notin A_j$, la seconde somme est nulle et le second membre vaut m, qui est supérieur ou égal à $\mathbf{1}_{A_1 \cup \cdots \cup A_n}(\omega)$; si $\omega \in A_j$, la seconde somme vaut $m - 1$, le second membre vaut $m - (m-1) = 1$, et le premier aussi. On conclut comme en 2 (b) :

$$\boxed{\mathrm{P}(A_1 \cup \cdots \cup A_n) \leqslant \sum_{i=1}^{n} \mathrm{P}(A_i) - \sum_{i \neq j} \mathrm{P}(A_i \cap A_j)}$$

On peut aussi l'obtenir en écrivant la réunion comme celle de A_j et des $A_i \setminus A_j$, pour $i \neq j$, puis en appliquant la sous-additivité et $\mathrm{P}(A_i \setminus A_j) = \mathrm{P}(A_i) - \mathrm{P}(A_i \cap A_j)$. Des tests exacts sur 500 familles d'événements tirées au hasard confirment les trois inégalités ✓.

À retenir. Les questions 2 et 3 encadrent la probabilité d'une réunion à l'aide des probabilités des intersections deux à deux : tronquer la formule du crible (section 21.2.5) après le terme d'ordre 2 donne un minorant. L'outil est le passage des indicatrices aux probabilités : une inégalité vraie en chaque issue devient une inégalité entre probabilités, exactement comme l'exemple de la section 21.2.2 transformait une égalité de fonctions en égalité de cardinaux.

EXERCICE 33 ⋆⋆⋆

1. Une répartition est une application de l'ensemble des k objets dans l'ensemble des n tiroirs ; l'univers est l'ensemble de ces n^k applications (proposition 3), muni de la probabilité uniforme. Les répartitions favorables sont les injections, au nombre de A_n^k si $k \leqslant n$, et de 0 sinon (théorème 1) :

$$\boxed{q_{k,n} = \frac{A_n^k}{n^k} = \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right)}$$

la formule restant vraie pour $k > n$, puisque le produit contient alors le facteur nul $1 - \dfrac{n}{n}$.

2. La probabilité d'au moins une coïncidence vaut $1 - q_{k,365}$, comme dans l'exemple du cours (section 21.3.2). On la calcule en flottants, et l'on augmente k jusqu'à dépasser le seuil :

```
def q(k, n):
    p = 1.0
    for i in range(k):
        p *= (n - i) / n
    return p

def seuil(s, n=365):
    k = 1
    while 1 - q(k, n) <= s:
        k += 1
    return k

print(seuil(0.5), seuil(0.9))    # affiche 23 41
```

On obtient $1 - q_{22,365} \approx 0{,}476$, $1 - q_{23,365} \approx 0{,}507$, puis $1 - q_{40,365} \approx 0{,}891$ et $1 - q_{41,365} \approx 0{,}903$:

$$\boxed{k = 23 \text{ pour le seuil } 0{,}5, \qquad k = 41 \text{ pour le seuil } 0{,}9}$$

3. (a) On pose $f(x) = \ln(1-x) + x$ et $g(x) = \ln(1-x) + x + x^2$, dérivables sur $\left[0\,,\dfrac{1}{2}\right]$, avec

$$f'(x) = 1 - \frac{1}{1-x} = -\frac{x}{1-x} \leqslant 0, \qquad g'(x) = 1 + 2x - \frac{1}{1-x} = \frac{x(1-2x)}{1-x} \geqslant 0.$$

La fonction f décroît et g croît sur $\left[0\,,\dfrac{1}{2}\right]$, et $f(0) = g(0) = 0$, d'où $f \leqslant 0 \leqslant g$:

$$\boxed{-x - x^2 \leqslant \ln(1-x) \leqslant -x \qquad \text{pour tout } x \in \left[0\,,\frac{1}{2}\right]}$$

L'inégalité de droite est l'inégalité de concavité du logarithme (chapitre 9) ; celle de gauche cesse d'être vraie près de 1, où $\ln(1-x)$ tend vers $-\infty$, d'où la restriction à $\left[0\,,\dfrac{1}{2}\right]$.

(b) On note $k = k_n$. Comme $t\sqrt{n} - 1 < k \leqslant t\sqrt{n}$, on a $\dfrac{k}{\sqrt{n}} \to t$ par encadrement. Pour $n \geqslant 4t^2$ et $0 \leqslant i \leqslant k-1$, on a $\dfrac{i}{n} \leqslant \dfrac{t\sqrt{n}}{n} = \dfrac{t}{\sqrt{n}} \leqslant \dfrac{1}{2}$: la question (a) s'applique à chaque facteur, en particulier $k \leqslant n$ et $q_{k,n} > 0$. En sommant,

$$-\sum_{i=0}^{k-1} \frac{i}{n} - \sum_{i=0}^{k-1} \frac{i^2}{n^2} \leqslant \ln q_{k,n} \leqslant -\sum_{i=0}^{k-1} \frac{i}{n}, \qquad \text{avec} \qquad \sum_{i=0}^{k-1} \frac{i}{n} = \frac{k(k-1)}{2n} = \frac{1}{2} \cdot \frac{k}{\sqrt{n}} \cdot \frac{k-1}{\sqrt{n}}.$$

Cette dernière quantité tend vers $\dfrac{t^2}{2}$, et la somme des $\dfrac{i^2}{n^2}$, qui a k termes majorés par $\dfrac{k^2}{n^2}$, est majorée par $\dfrac{k^3}{n^2} \leqslant \dfrac{t^3}{\sqrt{n}}$, qui tend vers 0. Par encadrement, $\ln q_{k_n,n}$ tend vers $-\dfrac{t^2}{2}$, et, par continuité de l'exponentielle,

$$\boxed{\lim_{n\to+\infty} q_{k_n,n} = \mathrm{e}^{-t^2/2}}$$

Pour $n = 10^6$ et $t = 1$, on calcule $q_{1000,\,10^6} \approx 0{,}6067$, pour $\mathrm{e}^{-1/2} \approx 0{,}6065$ ✓.

Interprétation pour $n = 365$. Avec $k = t\sqrt{365}$, la probabilité que les k anniversaires soient distincts vaut environ $\mathrm{e}^{-k^2/730}$. Elle vaut $\dfrac{1}{2}$ pour $k = \sqrt{730\ln 2} \approx 22{,}5$, et $\dfrac{1}{10}$ pour $k = \sqrt{730\ln 10} \approx 41{,}0$: l'approximation retrouve les deux seuils de la question 2. Le nombre de personnes nécessaire pour une coïncidence probable est de l'ordre de $\sqrt{n}$, et non de n, parce que $\dfrac{k^2}{2n}$ est à peu près le nombre $\dbinom{k}{2}$ de paires de personnes, multiplié par la probabilité $\dfrac{1}{n}$ de coïncidence pour chaque paire.

Piège. L'erreur la plus répandue est d'écrire $\ln q_{k,n} \approx -\sum \dfrac{i}{n}$ sans contrôler l'erreur : chacun des k termes commet une erreur de l'ordre de $\dfrac{i^2}{n^2}$, et c'est leur somme, de l'ordre de $\dfrac{k^3}{n^2}$, qu'il faut rendre négligeable. Elle tend vers 0 parce que k est de l'ordre de $\sqrt{n}$; pour k de l'ordre de $n^{2/3}$, elle ne tendrait plus vers 0.

EXERCICE 34 ⋆⋆⋆

L'univers est l'ensemble E^E des applications de E dans E, de cardinal n^n (proposition 3), muni de la probabilité uniforme.

1. Supposons $f \circ f = f$ et soit $y \in f(E)$, de la forme $y = f(x)$. Alors $f(y) = f(f(x)) = f(x) = y$. Réciproquement, si $f(y) = y$ pour tout $y \in f(E)$, alors, pour tout $x \in E$, l'élément $f(x)$ appartient à $f(E)$, donc $f(f(x)) = f(x)$.

$$\boxed{f \circ f = f \iff f(y) = y \ \text{ pour tout } y \in f(E)}$$

C'est la question 2 de l'exercice 18 de la feuille du chapitre 2, avec sa réciproque.

2. On classe les applications f telles que $f \circ f = f$ selon leur image $F = f(E)$, qui est une partie non vide de E, puisque E ne l'est pas. Soit F une partie à $k \geqslant 1$ éléments. D'après la question 1, une application f telle que $f \circ f = f$ et $f(E) = F$ vaut l'identité sur F, et envoie chacun des $n - k$ éléments de $E \setminus F$ dans F. Réciproquement, une telle application a pour image F (elle contient F, puisque f fixe les points de F, et elle est contenue dans F), et vérifie donc $f \circ f = f$ par la question 1. Il y en a autant que d'applications de $E \setminus F$ dans F, soit k^{n-k} (proposition 3). Il y a $\dbinom{n}{k}$ parties F à k éléments (théorème 2), et la proposition 2 (i) donne

$$\boxed{\left|\{f \in E^E \mid f \circ f = f\}\right| = \sum_{k=1}^{n} \binom{n}{k} k^{n-k}, \qquad p_n = \frac{1}{n^n} \sum_{k=1}^{n} \binom{n}{k} k^{n-k}}$$

Pour $n = 1, \ldots, 6$, on trouve 1, 3, 10, 41, 196, 1 057 applications, nombres confirmés par une énumération des n^n applications ✓. Pour $n = 3$, ce sont l'identité, les 3 applications constantes, et les 6 applications d'image à deux éléments.

3. Soit $n \geqslant 2$. Comme $\dbinom{n}{k} = \dfrac{n(n-1)\cdots(n-k+1)}{k!} \leqslant \dfrac{n^k}{k!}$, le terme d'indice k de p_n vérifie

$$0 \leqslant u_k = \frac{1}{n^n}\binom{n}{k} k^{n-k} \leqslant \frac{n^k k^{n-k}}{k!\, n^n} = \frac{1}{k!}\left(\frac{k}{n}\right)^{n-k}.$$

— Si $k \leqslant \dfrac{n}{2}$, alors $\dfrac{k}{n} \leqslant \dfrac{1}{2}$ et $n-k \geqslant \dfrac{n}{2}$, donc $u_k \leqslant \dfrac{1}{k!}\,2^{-n/2}$. En sommant, et comme $\displaystyle\sum_{k\geqslant 1}\frac{1}{k!} = \mathrm{e}-1$ (chapitre 20, théorème 2), la somme de ces termes est majorée par $(\mathrm{e}-1)\,2^{-n/2}$.

— Si $k > \dfrac{n}{2}$, alors $k \geqslant m$, avec $m = \lfloor n/2\rfloor + 1$, et $\left(\dfrac{k}{n}\right)^{n-k} \leqslant 1$, donc $u_k \leqslant \dfrac{1}{m!}$. Il y a au plus $n \leqslant 2m$ tels termes, de somme majorée par $\dfrac{2m}{m!} = \dfrac{2}{(m-1)!}$.

$$\boxed{0 \leqslant p_n \leqslant (\mathrm{e}-1)\,2^{-n/2} + \frac{2}{\lfloor n/2\rfloor !} \xrightarrow[n\to+\infty]{} 0}$$

Numériquement, $p_5 \approx 0{,}063$, $p_{10} \approx 2{,}2\times 10^{-4}$ et $p_{20} \approx 2{,}6\times 10^{-10}$: la décroissance est bien plus rapide que ne le dit le majorant, qui ne sert qu'à établir la limite.

Piège. Majorer chaque terme par une quantité qui tend vers 0 ne suffit pas, puisque le nombre de termes augmente avec n : l'erreur la plus répandue est d'oublier ce nombre. La séparation en deux paquets répond à deux difficultés distinctes. Pour k petit, c'est le facteur $\left(\dfrac{k}{n}\right)^{n-k}$ qui est minuscule, et la série exponentielle contrôle la somme des $\dfrac{1}{k!}$; pour k grand, ce facteur peut être proche de 1, mais $\dfrac{1}{k!}$ écrase le nombre n de termes.

Probabilités conditionnelles et indépendance

EXERCICE 35 ⋆

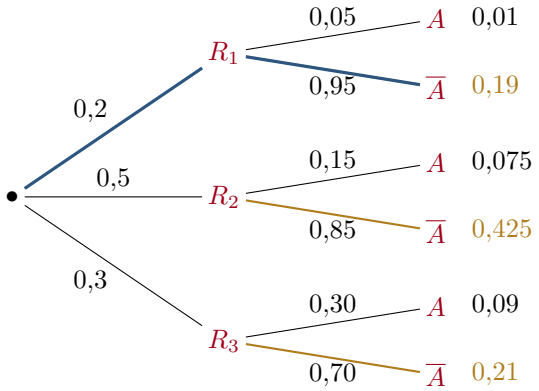
On choisit un client au hasard dans un espace probabilisé fini (Ω, P). Les événements R_1, R_2, R_3 forment un système complet, avec $\mathrm{P}(R_1) = 0{,}2$, $\mathrm{P}(R_2) = 0{,}5$ et $\mathrm{P}(R_3) = 0{,}3$. On note A l'événement « le client a un accident dans l'année » ; les données sont $\mathrm{P}_{R_1}(A) = 0{,}05$, $\mathrm{P}_{R_2}(A) = 0{,}15$ et $\mathrm{P}_{R_3}(A) = 0{,}30$.

1. La formule des probabilités totales (théorème 5) donne

$$\mathrm{P}(A) = 0{,}05\times 0{,}2 + 0{,}15\times 0{,}5 + 0{,}30\times 0{,}3 = 0{,}01 + 0{,}075 + 0{,}09, \qquad \boxed{\mathrm{P}(A) = 0{,}175 = \frac{7}{40}}$$

2. On cherche $\mathrm{P}_{\overline{A}}(R_1)$. Par la proposition 8, $\mathrm{P}_{R_1}(\overline{A}) = 1 - 0{,}05 = 0{,}95$, et $\mathrm{P}(\overline{A}) = 0{,}825$. La formule de Bayes (théorème 6) donne

$$\mathrm{P}_{\overline{A}}(R_1) = \frac{\mathrm{P}_{R_1}(\overline{A})\,\mathrm{P}(R_1)}{\mathrm{P}(\overline{A})} = \frac{0{,}95\times 0{,}2}{0{,}825} = \frac{0{,}19}{0{,}825}, \qquad \boxed{\mathrm{P}_{\overline{A}}(R_1) = \frac{38}{165} \approx 0{,}230}$$

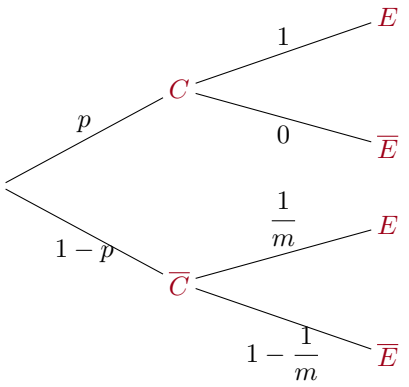


Les trois chemins qui mènent à $\overline{A}$ ont pour probabilités 0,19, 0,425 et 0,21, de somme 0,825 ; la formule de Bayes mesure la part du chemin bleu dans cette somme. Les deux autres parts donnent $\mathrm{P}_{\overline{A}}(R_2) = \dfrac{17}{33}$ et $\mathrm{P}_{\overline{A}}(R_3) = \dfrac{14}{55}$, et les trois probabilités totalisent $\dfrac{38+85+42}{165} = 1$ ✓.

L'absence d'accident fait passer la probabilité d'être un bon risque de 0,2 à 0,23 seulement : une année sans accident est le cas de loin le plus fréquent dans toutes les catégories, et elle n'apporte qu'une information faible. C'est pourquoi un assureur révise ses tarifs sur plusieurs années.

EXERCICE 36 ⋆

On note C l'événement « le candidat connaît la réponse » et E l'événement « il donne la réponse exacte ». Les données sont $\mathrm{P}(C) = p$, $\mathrm{P}_C(E) = 1$ et, s'il répond au hasard parmi les m réponses, $\mathrm{P}_{\overline{C}}(E) = \dfrac{1}{m}$.



On lit sur l'arbre les deux chemins qui mènent à E, de probabilités p et $\dfrac{1-p}{m}$; la réponse cherchée est la part du premier dans leur somme.

Par la formule des probabilités totales, $\mathrm{P}(E) = p + \dfrac{1-p}{m} > 0$, puis, par la formule de Bayes,

$$\mathrm{P}_E(C) = \frac{\mathrm{P}_C(E)\,\mathrm{P}(C)}{\mathrm{P}(E)} = \frac{p}{p + \frac{1-p}{m}}, \qquad \boxed{\mathrm{P}_E(C) = \frac{mp}{1 + (m-1)p}}$$

Pour $m = 4$ et $p = \dfrac{1}{2}$:

$$\boxed{\mathrm{P}_E(C) = \frac{2}{1 + \frac{3}{2}} = \frac{4}{5}}$$

À p fixé dans $]0\,,1[$, la fonction $m \mapsto \dfrac{mp}{1-p+mp}$ est croissante, de dérivée $\dfrac{p(1-p)}{(1-p+mp)^2} > 0$, et tend vers 1 quand m tend vers $+\infty$: pour $p = \dfrac{1}{2}$, elle vaut $\dfrac{2}{3}$ pour $m = 2$ et $\dfrac{10}{11}$ pour $m = 10$. Plus il y a de réponses proposées, moins une bonne réponse peut être due au hasard, et plus elle témoigne d'une connaissance réelle. Pour $p = 1$, la probabilité vaut 1 quel que soit m.
Piège. L'erreur la plus répandue est de confondre $\mathrm{P}_E(C)$, qui est demandée, et $\mathrm{P}_C(E) = 1$: un candidat qui sait répond juste, mais un candidat qui répond juste ne sait pas forcément. C'est la confusion signalée dans la remarque qui suit l'exemple du test de dépistage (section 21.4.1).

EXERCICE 37 ⋆

On note T l'événement « le dé choisi est pipé », de probabilité $\mathrm{P}(T) = \dfrac{25}{100} = \dfrac{1}{4}$, et S_n l'événement « les n premiers lancers donnent 6 ». L'hypothèse d'indépendance sachant le dé choisi signifie que, pour la probabilité P_T comme pour la probabilité $\mathrm{P}_{\overline{T}}$, les événements « le i-ième lancer donne 6 » sont mutuellement indépendants, de probabilités respectives $\dfrac{1}{2}$ et $\dfrac{1}{6}$. Ainsi

$$\mathrm{P}_T(S_n) = \frac{1}{2^n}, \qquad \mathrm{P}_{\overline{T}}(S_n) = \frac{1}{6^n}.$$

1. C'est le cas $n = 1$ de la question suivante : par la formule de Bayes,

$$\mathrm{P}_{S_1}(T) = \frac{\frac{1}{2} \times \frac{1}{4}}{\frac{1}{2} \times \frac{1}{4} + \frac{1}{6} \times \frac{3}{4}} = \frac{\frac{1}{8}}{\frac{1}{8} + \frac{1}{8}}, \qquad \boxed{\mathrm{P}_{S_1}(T) = \frac{1}{2}}$$

2. De même, par les théorèmes 5 et 6,

$$p_n = \mathrm{P}_{S_n}(T) = \frac{\frac{1}{4} \cdot \frac{1}{2^n}}{\frac{1}{4} \cdot \frac{1}{2^n} + \frac{3}{4} \cdot \frac{1}{6^n}} = \frac{1}{1 + 3 \cdot \frac{2^n}{6^n}}, \qquad \boxed{p_n = \frac{3^n}{3^n + 3} \xrightarrow[n \to +\infty]{} 1}$$

On trouve $p_0 = \dfrac{1}{4}$, $p_1 = \dfrac{1}{2}$, $p_2 = \dfrac{3}{4}$, $p_3 = \dfrac{9}{10}$, $p_4 = \dfrac{27}{28}$ ✓. Chaque nouveau 6 multiplie par 3 le rapport $\dfrac{p_n}{1 - p_n} = 3^{n-1}$, qui est le rapport des chances « pipé contre équilibré » : c'est le rapport $\dfrac{1/2}{1/6}$ des probabilités d'obtenir 6 avec les deux sortes de dés. Une longue série de 6 rend la culpabilité du dé presque certaine, bien qu'un dé équilibré puisse aussi la produire.

Piège. Les lancers sont indépendants sachant le dé, mais pas pour la probabilité P : l'événement « le premier lancer donne 6 », de probabilité $\dfrac{1}{4}$, rend le dé plus suspect et le second 6 plus probable. L'erreur la plus répandue est d'écrire $\mathrm{P}(S_n) = \mathrm{P}(S_1)^n$; on calcule $\mathrm{P}(S_n)$ par la formule des probabilités totales, comme le fera l'exercice 43.

EXERCICE 38 ⋆

1. On note $x = \mathrm{P}(A \cap B)$, $y = \mathrm{P}(A \cap \overline{B})$, $z = \mathrm{P}(\overline{A} \cap B)$, $w = \mathrm{P}(\overline{A} \cap \overline{B})$ les probabilités des quatre atomes de l'exercice 23, avec $x + y + z + w = 1$. Alors $\mathrm{P}(A) = x + y$ et $\mathrm{P}(B) = x + z$, d'où

$$\mathrm{P}(A)\,\mathrm{P}(B) - \mathrm{P}(A \cap B) = x^2 + xz + xy + yz - x = x(x + y + z - 1) + yz = yz - xw.$$

$$\boxed{A \text{ et } B \text{ indépendants} \iff \mathrm{P}(A \cap B)\,\mathrm{P}(\overline{A} \cap \overline{B}) = \mathrm{P}(A \cap \overline{B})\,\mathrm{P}(\overline{A} \cap B)}$$

La condition est symétrique en A et B, et inchangée si l'on remplace A par $\overline{A}$: on retrouve la proposition 10 (ii). Dans le tableau des quatre atomes, elle dit que le « déterminant » $xw - yz$ est nul ; l'exercice 62 généralisera cette remarque.

2. Si $A \subset B$, alors $A \cap B = A$, et l'indépendance s'écrit $\mathrm{P}(A) = \mathrm{P}(A)\,\mathrm{P}(B)$, c'est-à-dire

$$\mathrm{P}(A)(1 - \mathrm{P}(B)) = 0.$$

$$\boxed{A \subset B \text{ et } A, B \text{ indépendants} \implies \mathrm{P}(A) = 0 \text{ ou } \mathrm{P}(B) = 1}$$

Avec $B = A$, l'événement A est indépendant de lui-même si et seulement si $\mathrm{P}(A) = \mathrm{P}(A)^2$:

$$\boxed{A \text{ est indépendant de lui-même} \iff \mathrm{P}(A) \in \{0, 1\}}$$

Si toutes les issues ont une probabilité strictement positive, ce sont exactement $\varnothing$ et Ω ; sinon, un événement non vide peut être de probabilité nulle. Un événement indépendant de lui-même est « certain ou impossible » : savoir qu'il est réalisé n'apprend rien, parce qu'on le savait déjà.

3. Supposons $A_1, \ldots, A_n$ mutuellement indépendants, de probabilité $p < 1$, et de réunion Ω. Par la proposition 10 (iii), les contraires $\overline{A_1}, \ldots, \overline{A_n}$ sont mutuellement indépendants, et

$$\mathrm{P}\big(\overline{A_1} \cap \cdots \cap \overline{A_n}\big) = (1 - p)^n > 0.$$

Or cet événement est le contraire de $A_1 \cup \cdots \cup A_n = \Omega$, c'est-à-dire $\varnothing$, de probabilité nulle : contradiction.

$$\boxed{\text{il n'existe pas de tels événements}}$$

Des événements indépendants ne peuvent pas « se partager » l'univers : quel que soit leur nombre, il reste une probabilité $(1 - p)^n$ qu'aucun ne soit réalisé. Au contraire, des événements incompatibles de probabilité $\dfrac{1}{n}$ peuvent recouvrir Ω, ce qui illustre une fois de plus la différence entre indépendance et incompatibilité.

EXERCICE 39 ⋆

1. L'erreur i subsiste après la j-ième relecture si et seulement si elle subsistait avant cette relecture et si $C_{i,j}$ n'est pas réalisé. Par récurrence sur j, l'erreur i subsiste après n relectures si et seulement si l'événement $\overline{C_{i,1}} \cap \cdots \cap \overline{C_{i,n}}$ est réalisé. Ces contraires sont mutuellement indépendants (proposition 10 (iii)), de probabilité $\dfrac{2}{3}$:

$$\boxed{\mathrm{P}(\text{l'erreur 1 subsiste après } n \text{ relectures}) = \left(\frac{2}{3}\right)^n}$$

2. Notons $S_i = \overline{C_{i,1}} \cap \cdots \cap \overline{C_{i,n}}$ l'événement « l'erreur i subsiste ». Pour toute partie I de $[\![1\,,4]\!]$, l'intersection des S_i pour $i \in I$ est l'intersection des $\overline{C_{i,j}}$ pour $i \in I$ et $j \in [\![1\,,n]\!]$, d'où, par l'indépendance mutuelle des $\overline{C_{i,j}}$,

$$\mathrm{P}\Big(\bigcap_{i\in I} S_i\Big) = \left(\frac{2}{3}\right)^{n|I|} = \prod_{i\in I} \mathrm{P}(S_i) :$$

les événements S_1, …, S_4 sont mutuellement indépendants, donc aussi leurs contraires. Le livre ne contient plus d'erreur lorsque $\overline{S_1} \cap \cdots \cap \overline{S_4}$ est réalisé :

$$\boxed{\mathrm{P}(\text{plus aucune erreur après } n \text{ relectures}) = \left(1 - \left(\frac{2}{3}\right)^n\right)^4}$$

Cette probabilité croît avec n. Un calcul exact en nombres rationnels donne

$$\left(1 - \left(\frac{2}{3}\right)^9\right)^4 \approx 0{,}89994 < 0{,}9 < \left(1 - \left(\frac{2}{3}\right)^{10}\right)^4 \approx 0{,}93242, \qquad \boxed{\text{il faut 10 relectures}}$$

Piège. La valeur obtenue pour $n = 9$ est à moins de 10^{-4} du seuil : en résolvant $(\frac{2}{3})^n < 1 - 0{,}9^{1/4}$ par logarithmes, on trouve $n > 9{,}0015$, et une valeur approchée trop grossière de $0{,}9^{1/4}$ conduirait à la réponse fausse $n = 9$. Dans un tel cas, on contrôle les deux entiers voisins par un calcul exact. L'erreur la plus répandue, sur le fond, est de répondre $(\frac{2}{3})^{4n}$ ou $1 - (\frac{2}{3})^{4n}$ en mélangeant les quatre erreurs et les n relectures.

EXERCICE 40 ⋆⋆

Comme $A \subset A \cup B$, on a $\mathrm{P}(A \cup B) \geqslant \mathrm{P}(A) > 0$, et les deux probabilités conditionnelles sont définies. Comme $(A \cap B) \cap (A \cup B) = A \cap B$ et $(A \cap B) \cap A = A \cap B$,

$$\mathrm{P}_{A\cup B}(A \cap B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A \cup B)} \leqslant \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} = \mathrm{P}_A(A \cap B), \qquad \boxed{\mathrm{P}_{A\cup B}(A \cap B) \leqslant \mathrm{P}_A(A \cap B)}$$

avec égalité si et seulement si $\mathrm{P}(A \cap B) = 0$ ou $\mathrm{P}(A \cup B) = \mathrm{P}(A)$, c'est-à-dire $\mathrm{P}(B \setminus A) = 0$.

Interprétation. Les informations « A est réalisé » et « A ou B est réalisé » ne sont pas aussi riches : la première est plus précise, et elle est plus favorable à la réalisation simultanée de A et de B. Si $\mathrm{P}(B) > 0$, l'inégalité vaut aussi en échangeant A et B, et

$$\mathrm{P}_A(A \cap B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)}, \qquad \mathrm{P}_B(A \cap B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} :$$

parmi les trois informations « A », « B » et « A ou B », la plus favorable est donc la réalisation de celui des deux événements qui est le *moins* probable, et la moins favorable est « A ou B ».

C'est exactement l'exemple qui suit la proposition 8 : avec A « l'aîné est un garçon » et B « le cadet est un garçon », on a $A \cup B = U$ et $A \cap B = D$, et l'on y a trouvé $\mathrm{P}_U(D) = \dfrac{1}{3}$, contre $\mathrm{P}_A(D) = \dfrac{1}{2}$.
Piège. L'erreur la plus répandue est de penser qu'en apprenant « A ou B », on apprend « au moins autant » qu'en apprenant « A » : c'est l'inverse, puisque $A \cup B$ est un événement plus grand, donc une information plus pauvre. On retrouvera ce phénomène à l'exercice 42.

EXERCICE 41 ⋆⋆

1. Le contraire de $A_1 \cup \cdots \cup A_n$ est $\overline{A_1} \cap \cdots \cap \overline{A_n}$, et les $\overline{A_i}$ sont mutuellement indépendants (proposition 10 (iii)) :

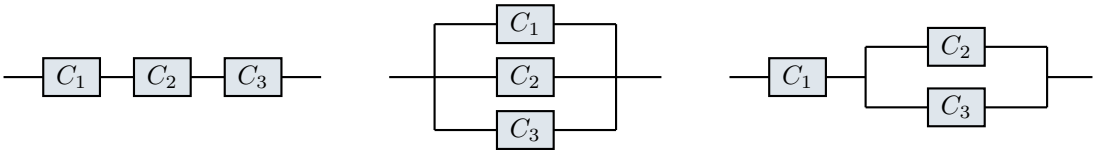
$$\boxed{\mathrm{P}(A_1 \cup \cdots \cup A_n) = 1 - \prod_{i=1}^{n}(1-p_i)}$$

Pour n expériences provoquant chacune un accident avec la probabilité p, il vient

$$\boxed{\mathrm{P}(\text{au moins un accident}) = 1-(1-p)^n \xrightarrow[n\to+\infty]{} 1 \quad \text{si } p>0}$$

Pour $p = 0{,}01$ et $n = 100$, la probabilité vaut déjà $1-0{,}99^{100} \approx 0{,}634$. Un risque faible, répété souvent, finit presque sûrement par se réaliser.

2. On note F_i l'événement « le composant C_i fonctionne », ces trois événements étant mutuellement indépendants.



en série — en parallèle — montage mixte

Le courant passe à travers un montage en série si tous les composants fonctionnent, à travers un montage en parallèle si l'un au moins fonctionne : la série est une intersection, le parallèle une réunion.

- *En série*, le circuit fonctionne si $F_1 \cap F_2 \cap F_3$ est réalisé, avec la probabilité $p_1p_2p_3$.
- *En parallèle*, il fonctionne si $F_1 \cup F_2 \cup F_3$ est réalisé, avec la probabilité $1-(1-p_1)(1-p_2)(1-p_3)$ par la question 1.
- *Montage mixte.* Il fonctionne si $F_1 \cap (F_2 \cup F_3) = (F_1 \cap F_2) \cup (F_1 \cap F_3)$ est réalisé. Par la proposition 7 (iv) et l'indépendance mutuelle, sa probabilité vaut $p_1p_2 + p_1p_3 - p_1p_2p_3$.

$$\boxed{\mathrm{P}_{\text{série}} = p_1p_2p_3, \qquad \mathrm{P}_{\text{parallèle}} = 1-(1-p_1)(1-p_2)(1-p_3), \qquad \mathrm{P}_{\text{mixte}} = p_1(p_2+p_3-p_2p_3)}$$

Le résultat du montage mixte est le produit de p_1 et de la probabilité $1-(1-p_2)(1-p_3)$ que le bloc parallèle fonctionne : le calcul par le crible a évité d'avoir à justifier l'indépendance de F_1 et de $F_2 \cup F_3$, que le lemme des coalitions (théorème 7), appliqué aux indicatrices, donnerait directement.

3. On note $M_1, \ldots, M_4$ les événements « le moteur i tombe en panne », mutuellement indépendants de probabilité p. L'avion A arrive si au plus un moteur tombe en panne, c'est-à-dire si l'un des cinq événements deux à deux incompatibles $\overline{M_1} \cap \cdots \cap \overline{M_4}$ et $M_i \cap \bigcap_{j\neq i} \overline{M_j}$, pour $i \in [\![1\,,4]\!]$, est réalisé ; par la proposition 10 (iii), ils ont pour probabilités $(1-p)^4$ et $p(1-p)^3$. L'avion B arrive si aucun de ses deux moteurs ne tombe en panne. Ainsi

$$\mathrm{P}(A \text{ arrive}) = (1-p)^4 + 4p(1-p)^3, \qquad \mathrm{P}(B \text{ arrive}) = (1-p)^2,$$

et, en factorisant par $(1-p)^2$,

$$\mathrm{P}(A \text{ arrive}) - \mathrm{P}(B \text{ arrive}) = (1-p)^2\big[(1-p)^2 + 4p(1-p) - 1\big] = (1-p)^2(2p-3p^2),$$

$$\boxed{\mathrm{P}(A \text{ arrive}) - \mathrm{P}(B \text{ arrive}) = p(1-p)^2(2-3p)}$$

$$\boxed{\text{on choisit l'avion } A \text{ si } p<\frac{2}{3}, \quad \text{l'avion } B \text{ si } p>\frac{2}{3}, \quad \text{indifféremment si } p=\frac{2}{3}}$$

Pour $p = \dfrac{1}{4}$, l'avion A arrive avec la probabilité $\dfrac{189}{256} \approx 0{,}74$ contre $\dfrac{9}{16} \approx 0{,}56$ pour B ; pour $p = \dfrac{2}{3}$, les deux probabilités valent $\dfrac{1}{9}$ ✓. Lorsque les moteurs sont fiables, avoir quatre moteurs, dont un peut lâcher, protège mieux que d'en avoir deux, dont aucun ne peut lâcher ; lorsqu'ils sont très peu fiables, les deux avions sont dangereux, et il vaut mieux exposer moins de moteurs à la panne.

EXERCICE 42 ⋆⋆

1. (a) On note $R_\heartsuit$, $R_\diamondsuit$, $D_\heartsuit$, $D_\diamondsuit$ les quatre cartes. On prend pour univers l'ensemble Ω des 2-listes de cartes distinctes (première carte, seconde carte), muni de la probabilité uniforme : les tirages successifs au hasard rendent ces $A_4^2 = 12$ issues également probables.

$$\boxed{\Omega = \{(c_1, c_2) \mid c_1, c_2 \in \{R_\heartsuit, R_\diamondsuit, D_\heartsuit, D_\diamondsuit\},\ c_1 \neq c_2\}, \qquad |\Omega| = 12, \qquad \text{P uniforme}}$$

(b) Soit U l'événement « au moins une dame » et DD l'événement « deux dames ». Le contraire de U est « deux rois », qui contient les 2 issues $(R_\heartsuit, R_\diamondsuit)$ et $(R_\diamondsuit, R_\heartsuit)$, donc $|U| = 10$; l'événement DD, inclus dans U, contient 2 issues. Ainsi

$$\boxed{\mathrm{P}_U(DD) = \frac{|DD \cap U|}{|U|} = \frac{2}{10} = \frac{1}{5}}$$

(c) Soit H l'événement « l'une des deux cartes est la dame de cœur ». Il contient 3 issues où $D_\heartsuit$ sort en premier et 3 où elle sort en second, soit 6 issues, dont 2 réalisent DD :

$$\boxed{\mathrm{P}_H(DD) = \frac{2}{6} = \frac{1}{3}}$$

Savoir *quelle* dame a été tirée fait passer la probabilité de $\frac{1}{5}$ à $\frac{1}{3}$: l'événement H est plus petit que U, donc plus informatif, et il élimine davantage d'issues défavorables, ce qui est le phénomène de l'exercice 40. Sachant H, l'autre carte est l'une des trois autres cartes, prise au hasard, et une seule est une dame.

2. Les hypothèses de modélisation sont les suivantes : les sexes des deux enfants sont indépendants et chacun vaut G ou F avec la probabilité $\frac{1}{2}$, et l'enfant qui accompagne le parent est choisi au hasard, avec la probabilité $\frac{1}{2}$ pour chacun, indépendamment des sexes. On prend donc

$$\Omega = \{G, F\}^2 \times \{1, 2\},$$

muni de la probabilité uniforme, une issue (s_1, s_2, c) donnant les sexes de l'aîné et du cadet et le rang c de l'enfant rencontré. L'événement V « l'enfant rencontré est un garçon » est formé des issues telles que $s_c = G$:

$$V = \{(G, G, 1), (G, G, 2), (G, F, 1), (F, G, 2)\}, \qquad \mathrm{P}(V) = \frac{4}{8} = \frac{1}{2}.$$

Parmi elles, l'autre enfant est une fille pour $(G, F, 1)$ et $(F, G, 2)$:

$$\boxed{\mathrm{P}_V(\text{l'autre enfant est une fille}) = \frac{2}{4} = \frac{1}{2}}$$

Dans l'exemple qui suit la proposition 8, l'information était « la famille compte au moins un garçon », et la probabilité que la famille compte aussi une fille valait $1 - \frac{1}{3} = \frac{2}{3}$. Ici, on a vu un garçon *déterminé*, comme si l'on savait que l'aîné est un garçon ; la probabilité vaut alors $\frac{1}{2}$.

Piège. La réponse dépend de la façon dont l'information a été obtenue, et pas seulement de son contenu apparent : l'erreur la plus répandue est d'oublier de modéliser ce mécanisme. Si l'on suppose au contraire que le parent se montre toujours avec un garçon lorsqu'il en a un, l'événement « on le croise avec un garçon » coïncide avec « au moins un garçon », et la probabilité que l'autre soit une fille redevient $\frac{2}{3}$, valeur confirmée par énumération. Les deux réponses sont justes, chacune dans son modèle.

EXERCICE 43 ⋆⋆

On note Δ_i l'événement « le dé choisi est D_i ». Les deux lancers éventuels du dé A étant indépendants, la famille $(\Delta_1, \dots, \Delta_7)$ est un système complet, avec

$$\mathrm{P}(\Delta_i) = \frac{1}{6} \text{ pour } 2 \leqslant i \leqslant 6, \qquad \mathrm{P}(\Delta_1) = \mathrm{P}(\Delta_7) = \frac{1}{6} \times \frac{1}{2} = \frac{1}{12},$$

de somme $\dfrac{5}{6} + \dfrac{1}{6} = 1$ ✓. Le dé D_i a $7-i$ faces noires sur 6 : le dé D_1 est entièrement noir, le dé D_7 entièrement blanc. L'hypothèse dit que, pour chaque probabilité P_{Δ_i}, les événements N_k sont mutuellement indépendants, de probabilité $\dfrac{7-i}{6}$. Pour $m \geqslant 1$, la formule des probabilités totales donne donc

$$a_m = \mathrm{P}(N_1 \cap \dots \cap N_m) = \sum_{i=1}^{7} \left(\frac{7-i}{6}\right)^m \mathrm{P}(\Delta_i) = \frac{1}{12} + \frac{1}{6} \sum_{j=1}^{5} \left(\frac{j}{6}\right)^m,$$

le dé D_1 contribuant $\dfrac{1}{12}$ et le dé D_7 rien.

1. Avec $m = 1$ et $1 + 2 + \dots + 5 = 15$:

$$\mathrm{P}(N_1) = a_1 = \frac{1}{12} + \frac{1}{6} \times \frac{15}{6} = \frac{3}{36} + \frac{15}{36}, \qquad \boxed{\mathrm{P}(N_1) = \frac{1}{2}}$$

Ce résultat se prévoit par symétrie : l'échange des couleurs transforme le dé D_i en le dé D_{8-i}, qui a la même probabilité d'être choisi.

2. Avec $1^2 + \dots + 5^2 = 55$ et $1^3 + \dots + 5^3 = 225$,

$$a_2 = \frac{1}{12} + \frac{55}{216} = \frac{73}{216}, \qquad a_3 = \frac{1}{12} + \frac{225}{1\,296} = \frac{333}{1\,296},$$

$$\mathrm{P}(N_3 \mid N_1 \cap N_2) = \frac{\mathrm{P}(N_1 \cap N_2 \cap N_3)}{\mathrm{P}(N_1 \cap N_2)} = \frac{a_3}{a_2} = \frac{333}{6 \times 73}, \qquad \boxed{\mathrm{P}(N_3 \mid N_1 \cap N_2) = \frac{111}{146} \approx 0{,}760}$$

3. De même, $P_2 = \dfrac{a_2}{a_1}$ et, pour $n \geqslant 3$, $P_n = \dfrac{a_n}{a_{n-1}}$, ces quotients étant définis puisque $a_m \geqslant \dfrac{1}{12} > 0$:

$$\boxed{P_2 = \mathrm{P}_{N_1}(N_2) = \frac{73}{108} \approx 0{,}676, \qquad P_n = \frac{a_n}{a_{n-1}} = \frac{\frac{1}{12} + \frac{1}{6}\sum_{j=1}^{5} \left(\frac{j}{6}\right)^n}{\frac{1}{12} + \frac{1}{6}\sum_{j=1}^{5} \left(\frac{j}{6}\right)^{n-1}}}$$

La formule vaut aussi pour $n = 2$, et une énumération exacte de toutes les issues (deux lancers du dé A, trois lancers du dé choisi) confirme $\mathrm{P}(N_1) = \dfrac{1}{2}$ et $\dfrac{111}{146}$ ✓.

4. Pour $1 \leqslant j \leqslant 5$, la suite géométrique $\left(\dfrac{j}{6}\right)^n$ tend vers 0, donc a_n et a_{n-1} tendent vers $\dfrac{1}{12}$:

$$\boxed{\lim_{n \to +\infty} P_n = 1}$$

Numériquement, $P_3 \approx 0{,}760$, $P_5 \approx 0{,}852$, $P_{10} \approx 0{,}942$ et $P_{20} \approx 0{,}990$. Après une longue suite de faces noires, il devient presque certain que le dé choisi est le dé D_1, entièrement noir : la formule de Bayes donne $\mathrm{P}(\Delta_1 \mid N_1 \cap \dots \cap N_n) = \dfrac{1/12}{a_n} \to 1$, et la face suivante est alors noire à coup sûr.

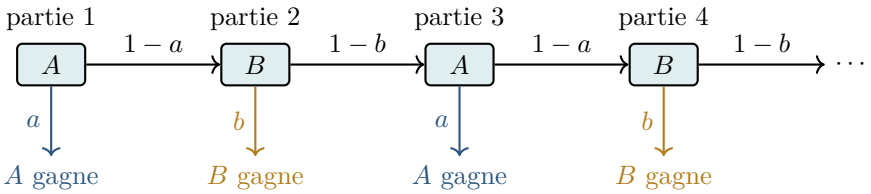
Piège. L'erreur la plus répandue est de croire les lancers indépendants pour la probabilité P, et de répondre

$$\mathrm{P}(N_3 \mid N_1 \cap N_2) = \mathrm{P}(N_3) = \frac{1}{2}.$$

Ils ne le sont que sachant le dé : on a $\mathrm{P}(N_1 \cap N_2) = \dfrac{73}{216} \neq \dfrac{1}{4} = \mathrm{P}(N_1)\mathrm{P}(N_2)$, parce que chaque face noire renseigne sur le dé choisi, comme les 6 de l'exercice 37.

EXERCICE 44 ⋆⋆

On modélise les parties par des événements W_1, …, W_{2n} mutuellement indépendants, où W_k signifie « le joueur dont c'est le tour gagnerait la k-ième partie, si elle avait lieu », avec $\mathrm{P}(W_k) = a$ pour k impair (tour de A) et $\mathrm{P}(W_k) = b$ pour k pair (tour de B). Un tel modèle existe : la proposition 17 (section 21.5.6) le fournit sur $\Omega = \{0,1\}^{2n}$. On note $r = (1-a)(1-b) \in]0\,,1[$, probabilité qu'un tour complet, formé d'une partie de A et d'une partie de B, ne donne aucun vainqueur.



Le jeu se lit le long de la chaîne : à chaque partie, le joueur gagne et le jeu s'arrête, ou il perd et la main passe. Chaque tour complet sans vainqueur multiplie par r la probabilité du chemin.

1. Le jeu est nul si aucune des $2n$ parties n'est gagnée, c'est-à-dire si $\overline{W_1} \cap \cdots \cap \overline{W_{2n}}$ est réalisé. Par la proposition 10 (iii),

$$\boxed{\mathrm{P}(\text{jeu nul}) = (1-a)^n(1-b)^n = r^n}$$

2. Le joueur A remporte le jeu si la première partie gagnée est une partie d'indice impair $2j+1$, avec $0 \leqslant j \leqslant n-1$, c'est-à-dire si l'un des événements deux à deux incompatibles $\overline{W_1} \cap \cdots \cap \overline{W_{2j}} \cap W_{2j+1}$ est réalisé ; ce dernier a pour probabilité $r^j a$. De même, B gagne à la partie $2j+2$ avec la probabilité $r^j(1-a)b$. En sommant les suites géométriques de raison $r \neq 1$,

$$\boxed{\mathrm{P}(A \text{ gagne}) = a\,\frac{1-r^n}{1-r}, \qquad \mathrm{P}(B \text{ gagne}) = (1-a)b\,\frac{1-r^n}{1-r}}$$

On contrôle que les trois probabilités totalisent 1 : comme $1-r = a+b-ab = a+(1-a)b$, la somme des deux premières vaut $1-r^n$ ✓. Une énumération exacte pour $n=2$, $a=\dfrac{1}{3}$, $b=\dfrac{1}{5}$ confirme les trois formules.

3. Le jeu est équitable si et seulement si $\mathrm{P}(A \text{ gagne}) = \mathrm{P}(B \text{ gagne})$, c'est-à-dire, puisque $1-r^n > 0$,

$$\boxed{a = (1-a)\,b, \qquad \text{soit} \qquad b = \frac{a}{1-a}}$$

Cette condition ne dépend pas de n, et elle impose $\dfrac{a}{1-a} < 1$, c'est-à-dire $a < \dfrac{1}{2}$; par exemple, $a = \dfrac{1}{3}$ et $b = \dfrac{1}{2}$. Elle dit que A et B ont la même probabilité de gagner lors du premier tour : A gagne d'emblée avec la probabilité a, tandis que B ne joue que si A a perdu, et gagne alors avec la probabilité $(1-a)b$. Chaque tour suivant reproduit la même situation, à un facteur r près, ce qui explique que n n'intervienne pas. L'avantage de jouer en premier doit être compensé par une meilleure probabilité de succès pour le second joueur, $b > a$; si $a \geqslant \dfrac{1}{2}$, aucune valeur de b ne compense cet avantage.

À retenir. Pour un jeu « au premier qui réussit », on découpe l'événement selon le rang du premier succès : ces événements sont incompatibles, et chacun est une intersection d'événements indépendants. C'est la même décomposition qu'à la question 3 de l'exercice 30, où l'on découpait selon le rang du $(n+1)$-ième pile.

EXERCICE 45 ⋆⋆

1. Par définition de P_C (section 21.4.1), on a $\mathrm{P}_C(A) = \dfrac{\mathrm{P}(A\cap C)}{\mathrm{P}(C)}$, et de même pour B et pour $A\cap B$. L'inégalité $\mathrm{P}_C(A\cap B) > \mathrm{P}_C(A)\,\mathrm{P}_C(B)$ s'écrit donc

$$\frac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(C)} > \frac{\mathrm{P}(A\cap C)\,\mathrm{P}(B\cap C)}{\mathrm{P}(C)^2},$$

soit, en multipliant par $\mathrm{P}(C)^2 > 0$,

$$\boxed{\mathrm{P}(A\cap B\cap C)\,\mathrm{P}(C) > \mathrm{P}(A\cap C)\,\mathrm{P}(B\cap C)}$$

La corrélation négative sachant C s'écrit de même avec l'inégalité stricte dans l'autre sens. Pour $C=\Omega$, on retrouve la définition de l'énoncé.

2. On prend $\Omega = [\![1\,,9]\!]$, muni de la probabilité uniforme, et

$$\boxed{A=\{1,2,3,4\},\qquad B=\{1,2,5,6\},\qquad C=\{3,5,7,8,9\}}$$

Sans conditionnement. On a $\mathrm{P}(A)=\mathrm{P}(B)=\dfrac{4}{9}$ et $A\cap B=\{1,2\}$, d'où

$$\mathrm{P}(A\cap B)=\frac{2}{9}=\frac{18}{81}>\frac{16}{81}=\mathrm{P}(A)\,\mathrm{P}(B)\,:\qquad A \text{ et } B \text{ sont positivement corrélés.}$$

Sachant C. On a $\mathrm{P}(C)=\dfrac{5}{9}\in\,]0\,,1[$, $A\cap C=\{3\}$, $B\cap C=\{5\}$ et $A\cap B\cap C=\varnothing$. Le critère de la question 1 donne $0\times\dfrac{5}{9}<\dfrac{1}{9}\times\dfrac{1}{9}$: la corrélation est négative. En probabilités conditionnelles, $\mathrm{P}_C(A)=\mathrm{P}_C(B)=\dfrac{1}{5}$ et $\mathrm{P}_C(A\cap B)=0$, et l'écart vaut $0-\dfrac{1}{25}=-\dfrac{1}{25}$.

Sachant $\overline{C}$. On a $\overline{C}=\{1,2,4,6\}$, $A\cap\overline{C}=\{1,2,4\}$, $B\cap\overline{C}=\{1,2,6\}$ et $A\cap B\cap\overline{C}=\{1,2\}$, d'où

$$\mathrm{P}_{\overline{C}}(A\cap B)=\frac{1}{2}=\frac{8}{16}<\frac{9}{16}=\frac{3}{4}\times\frac{3}{4}=\mathrm{P}_{\overline{C}}(A)\,\mathrm{P}_{\overline{C}}(B).$$

$$\boxed{\begin{aligned}&\mathrm{P}(A\cap B)-\mathrm{P}(A)\mathrm{P}(B)=\frac{2}{81}>0,\\&\mathrm{P}_C(A\cap B)-\mathrm{P}_C(A)\mathrm{P}_C(B)=-\frac{1}{25}<0,\\&\mathrm{P}_{\overline{C}}(A\cap B)-\mathrm{P}_{\overline{C}}(A)\mathrm{P}_{\overline{C}}(B)=-\frac{1}{16}<0\end{aligned}}$$

Le mécanisme se lit sur les chiffres : A et B sont tous deux rares sachant C (probabilité $\dfrac{1}{5}$) et fréquents sachant $\overline{C}$ (probabilité $\dfrac{3}{4}$). Savoir que A est réalisé rend $\overline{C}$ plus probable, donc B aussi : c'est par l'intermédiaire de C que A et B semblent liés positivement, alors qu'à C fixé ils se repoussent.

3. Considérons une population de malades, C désignant l'événement « la forme de la maladie est grave », A « le patient a reçu le traitement » et B « le patient guérit ». Si les médecins réservent le traitement aux formes bénignes, qui guérissent souvent d'elles-mêmes, la situation peut être exactement celle de la question 2 : sur l'ensemble des patients, les traités guérissent plus souvent que les autres, et l'on conclurait à l'efficacité du traitement, alors que, parmi les formes graves comme parmi les formes bénignes, il diminue les chances de guérison.

Piège. L'erreur la plus répandue est de croire qu'une corrélation observée sur toute une population se retrouve dans chacune des sous-populations d'un système complet $(C,\overline{C})$, ou l'inverse. La formule des probabilités totales (théorème 5) écrit bien $\mathrm{P}(A\cap B)$, $\mathrm{P}(A)$ et $\mathrm{P}(B)$ comme des moyennes pondérées par $\mathrm{P}(C)$ et $\mathrm{P}(\overline{C})$, mais le produit $\mathrm{P}(A)\,\mathrm{P}(B)$ n'est pas la moyenne des produits : c'est ce défaut de linéarité qui autorise le renversement.

EXERCICE 46 ⋆⋆

On modélise les n lancers par $\Omega = \{0,1\}^n$, muni de la probabilité uniforme, 1 codant pile : c'est le modèle de la proposition 17, où chaque coordonnée X_k suit la loi $\mathcal{B}(\frac{1}{2})$, et les X_k sont mutuellement indépendantes. On a $A_k = (X_k = 1)$ pour $k \geqslant 1$, et $A_0 = (X_1 + \cdots + X_n \text{ pair})$.

Probabilités. On a $\mathrm{P}(A_k) = \dfrac{1}{2}$ pour $k \geqslant 1$. L'application qui change la première coordonnée d'une issue est une bijection de Ω qui change la parité du nombre de piles : elle échange A_0 et $\overline{A_0}$, qui ont donc le même cardinal, et $\mathrm{P}(A_0) = \dfrac{1}{2}$.

La famille entière n'est pas indépendante. L'événement $A_1 \cap \cdots \cap A_n$ est réduit à l'issue $(1,\dots,1)$, qui compte n piles. Il est contenu dans A_0 si n est pair, et disjoint de A_0 si n est impair, d'où

$$\boxed{\mathrm{P}(A_0 \cap A_1 \cap \cdots \cap A_n) = \begin{cases} \dfrac{1}{2^n} & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases} \neq \frac{1}{2^{n+1}} = \mathrm{P}(A_0)\,\mathrm{P}(A_1)\cdots\mathrm{P}(A_n)}$$

Les événements A_0, …, A_n ne sont donc pas mutuellement indépendants : les résultats des n lancers déterminent la parité du nombre de piles.

Les sous-familles de n événements. Une sous-famille de n événements s'obtient en retirant l'un des A_j.

— Si l'on retire A_0, il reste A_1, …, A_n, qui sont mutuellement indépendants par la proposition 15, puisque les X_k le sont.

— Si l'on retire A_j avec $j \geqslant 1$, soit J une partie de $[\![0\,,n]\!] \setminus \{j\}$. Si $0 \notin J$, l'égalité demandée par la définition de l'indépendance mutuelle résulte encore de l'indépendance des X_k. Sinon, on écrit $J = \{0\} \cup J'$ avec $J' \subset [\![1\,,n]\!] \setminus \{j\}$, et l'on note $m = |J'|$. Les issues de $\bigcap_{i\in J'} A_i$ ont leurs m coordonnées d'indices dans J' égales à 1, les $n-m$ autres étant libres ; il y en a 2^{n-m}. Parmi elles, celles de A_0 sont celles dont les coordonnées libres comptent un nombre de 1 de la parité de m. Comme j est l'un des indices libres, changer la coordonnée d'indice j échange ces issues et les autres, qui sont donc en nombre égal : il y en a 2^{n-m-1}, et

$$\mathrm{P}\Big(A_0 \cap \bigcap_{i\in J'} A_i\Big) = \frac{2^{n-m-1}}{2^n} = \frac{1}{2^{m+1}} = \mathrm{P}(A_0)\prod_{i\in J'} \mathrm{P}(A_i).$$

$$\boxed{\text{toute sous-famille de } n \text{ événements parmi } A_0,\dots,A_n \text{ est mutuellement indépendante}}$$

Pour $n = 2$, l'événement A_0, « deux piles ou deux faces », est l'événement C « les deux pièces donnent le même côté » du second exemple qui suit la proposition 10 : on retrouve exactement cet exemple, avec trois événements deux à deux indépendants et non mutuellement indépendants. L'exercice le prolonge à $n+1$ événements dont toutes les sous-familles strictes, et pas seulement les paires, sont indépendantes.

À retenir. Contrôler l'indépendance de toutes les sous-familles strictes ne dispense pas de contrôler la famille entière : la définition de l'indépendance mutuelle porte sur toutes les parties J de l'ensemble des indices, et une seule égalité manquante suffit à la mettre en défaut.

EXERCICE 47 ⋆⋆

Pour deux événements A et B de Ω_n, muni de la probabilité uniforme, on a

$$A \text{ et } B \text{ indépendants} \iff \frac{|A\cap B|}{n} = \frac{|A|}{n}\cdot\frac{|B|}{n} \iff \boxed{n\,|A\cap B| = |A|\,|B|}$$

et A est non trivial si et seulement si $1 \leqslant |A| \leqslant n-1$.

1. Supposons n premier, et soit A, B deux événements non triviaux. Si A et B étaient indépendants, n diviserait le produit $|A|\,|B|$, donc l'un des facteurs par le lemme d'Euclide (chapitre 12, proposition 9 (ii)). C'est impossible, puisque $|A|$ et $|B|$ appartiennent à $[\![1\,,n-1]\!]$.

$$\boxed{n \text{ premier} \implies \text{deux événements non triviaux de } \Omega_n \text{ ne sont jamais indépendants}}$$

2. Supposons $n = ab$ avec a, $b \geqslant 2$. On pose

$$\boxed{A = \{b, 2b, \dots, ab\} \quad \text{(les multiples de } b\text{)}, \qquad B = [\![1\,,b]\!]}$$

On a $|A| = a$ et $|B| = b$, tous deux dans $[\![1\,,n-1]\!]$ puisque a, $b \geqslant 2$, et $A\cap B = \{b\}$. Ainsi $n\,|A\cap B| = n = ab = |A|\,|B|$: les événements A et B, de probabilités $\dfrac{1}{b}$ et $\dfrac{1}{a}$, sont non triviaux et indépendants. Pour $n = 6$, avec $a = 3$ et $b = 2$, on obtient $A = \{2,4,6\}$ et $B = \{1,2\}$: « être pair » est indépendant de « être au plus 2 ».

L'idée de la construction est d'écrire chaque $k \in [\![1\,,n]\!]$ de façon unique sous la forme $k = qb + s$, avec $q \in [\![0\,,a-1]\!]$ et $s \in [\![1\,,b]\!]$ (division euclidienne de $k-1$ par b, chapitre 12, théorème 1) : sous la probabilité uniforme, le couple (q,s) est uniforme sur $[\![0\,,a-1]\!]\times[\![1\,,b]\!]$, et deux événements qui ne portent l'un que sur s, l'autre que sur q, sont indépendants. L'événement A est $(s=b)$ et l'événement B est $(q=0)$.

À retenir. L'indépendance est une propriété numérique, $\mathrm{P}(A\cap B) = \mathrm{P}(A)\mathrm{P}(B)$, et elle dépend de l'arithmétique des cardinaux : sur un univers uniforme de cardinal premier, aucune paire d'événements non triviaux n'y satisfait. C'est la même dépendance que dans le premier exemple qui suit la proposition 10, où l'indépendance de « divisible par 2 » et « divisible par 3 » vaut sur $[\![1\,,60]\!]$ et pas sur $[\![1\,,100]\!]$.

EXERCICE 48 ⋆⋆

1. Par la proposition 10 (iii), la famille $(B_1,\dots,B_n)$, obtenue en remplaçant certains A_k par leurs contraires, est formée d'événements mutuellement indépendants. Chaque $\mathrm{P}(B_k)$ vaut $\mathrm{P}(A_k)$ ou $1-\mathrm{P}(A_k)$, qui appartiennent à $]0\,,1[$. Donc

$$\mathrm{P}(B_1\cap\cdots\cap B_n) = \mathrm{P}(B_1)\cdots\mathrm{P}(B_n) > 0, \qquad \text{d'où} \qquad \boxed{B_1\cap\cdots\cap B_n \neq \varnothing}$$

puisque $\mathrm{P}(\varnothing) = 0$.

2. Soit $(B_1,\dots,B_n)$ et $(B'_1,\dots,B'_n)$ deux choix différents : il existe k tel que $B_k \neq B'_k$, c'est-à-dire, quitte à échanger les rôles, $B_k = A_k$ et $B'_k = \overline{A_k}$. Les deux intersections sont contenues respectivement dans A_k et dans $\overline{A_k}$, qui sont disjoints :

$$\boxed{(B_1\cap\cdots\cap B_n)\cap(B'_1\cap\cdots\cap B'_n) = \varnothing}$$

3. Un choix est une application de $[\![1\,,n]\!]$ dans un ensemble à deux éléments ; il y en a 2^n (proposition 3). Les 2^n intersections correspondantes sont des parties de Ω non vides (question 1) et deux à deux disjointes (question 2). En choisissant un élément dans chacune, on obtient 2^n éléments distincts de Ω :

$$\boxed{|\Omega| \geqslant 2^n}$$

Cas d'égalité. On prend $\Omega = \{0,1\}^n$, muni de la probabilité uniforme, et $A_k = \{\omega\in\Omega \mid \omega_k = 1\}$. Par la proposition 17, les coordonnées sont mutuellement indépendantes, de loi $\mathcal{B}(\dfrac{1}{2})$, donc les A_k

sont mutuellement indépendants (proposition 15), de probabilité $\frac{1}{2} \in]0\,,1[$, et $|\Omega| = 2^n$. Chaque intersection $B_1 \cap \cdots \cap B_n$ y est réduite à une issue.
À retenir. Les questions 1 et 2 montrent que les 2^n intersections forment un système complet d'événements, tous non vides : n événements indépendants non triviaux « découpent » l'univers en au moins 2^n morceaux. Modéliser n épreuves indépendantes et non déterministes demande donc un univers d'au moins 2^n issues ; c'est l'analogue probabiliste de l'exercice 13, où m parties séparant n points imposaient $n \leqslant 2^m$.

EXERCICE 49 ⋆⋆

On modélise les n lancers par des variables $X_1, \ldots, X_n$ mutuellement indépendantes de loi uniforme sur $[\![0\,,9]\!]$, ce que permet la proposition 17.

1. Les événements $(X_k \neq 6)$ sont mutuellement indépendants (proposition 15), de probabilité $\frac{9}{10}$. La probabilité qu'aucun 6 n'apparaisse vaut donc $\left(\frac{9}{10}\right)^n$, qui tend vers 0 (suite géométrique de raison appartenant à $]-1\,,1[$, chapitre 7).

$$\boxed{\mathrm{P}(\text{au moins un 6}) = 1 - \left(\frac{9}{10}\right)^n \xrightarrow[n\to+\infty]{} 1}$$

Pour $n = 100$, la probabilité de n'obtenir aucun 6 vaut déjà $0{,}9^{100} \approx 2{,}7 \cdot 10^{-5}$.

2. On note $m = \lfloor \frac{n}{2} \rfloor$ et, pour $k \in [\![1\,,m]\!]$, $B_k = (X_{2k-1} = 6) \cap (X_{2k} = 6)$. Par indépendance, $\mathrm{P}(B_k) = \frac{1}{100}$. L'événement B_k s'écrit $(Z_k = 1)$, où Z_k est l'indicatrice de $\{(6,6)\}$ évaluée en (X_{2k-1}, X_{2k}) : c'est une fonction du paquet (X_{2k-1}, X_{2k}), et ces paquets sont disjoints. Par le lemme des coalitions (théorème 7), les variables $Z_1, \ldots, Z_m$ sont mutuellement indépendantes, donc les événements $B_1, \ldots, B_m$ aussi (proposition 15), ainsi que leurs contraires (proposition 10 (iii)).

Si deux 6 consécutifs n'apparaissent jamais, aucun B_k n'est réalisé :

$$\mathrm{P}(\text{jamais deux 6 consécutifs}) \leqslant \mathrm{P}\left(\overline{B_1} \cap \cdots \cap \overline{B_m}\right) = \left(\frac{99}{100}\right)^m,$$

$$\boxed{\mathrm{P}(\text{deux 6 consécutifs au moins une fois}) \geqslant 1 - \left(\frac{99}{100}\right)^{\lfloor n/2 \rfloor} \xrightarrow[n\to+\infty]{} 1}$$

puisque $\lfloor \frac{n}{2} \rfloor$ tend vers $+\infty$. Pour $n = 1\,000$, la minoration vaut $1 - 0{,}99^{500} \approx 0{,}9934$, et la valeur exacte, calculée en suivant à chaque rang si le dernier lancer est un 6, vaut 0,9999 : la minoration est grossière, mais elle suffit pour la limite.
Piège. Les événements « les lancers j et $j+1$ donnent 6 », pour $j \in [\![1\,,n-1]\!]$, ne sont pas indépendants : deux d'entre eux d'indices consécutifs partagent un lancer, et la formule $\left(\frac{99}{100}\right)^{n-1}$ serait fausse. C'est pour obtenir des événements indépendants que l'on se restreint aux paquets disjoints $\{2k-1, 2k\}$, en perdant de l'information, mais sans perdre la limite.

3. Le même raisonnement s'applique à n'importe quelle suite de chiffres fixée à l'avance, de longueur ℓ : en découpant les lancers en $\lfloor \frac{n}{\ell} \rfloor$ paquets disjoints de ℓ lancers consécutifs, la probabilité qu'elle n'apparaisse jamais est au plus $\left(1 - 10^{-\ell}\right)^{\lfloor n/\ell \rfloor}$, qui tend vers 0. Un singe tapant au hasard sur un clavier, chaque touche étant indépendante des précédentes, finit donc presque certainement par écrire n'importe quel texte donné, par exemple une œuvre de Shakespeare, codée en chiffres. L'image a été employée par Émile Borel en 1913, dans un article du *Journal de Physique* sur la mécanique statistique. Le calcul montre aussi pourquoi c'est sans conséquence pratique : il faut un nombre de lancers de l'ordre de 10^ℓ, astronomique dès que le texte a quelques dizaines de caractères.

EXERCICE 50 ⋆⋆⋆

Modèle. L'énoncé fixe les données : $\mathrm{P}(A_1) = \dfrac{1}{2}$ et, pour tout $k \geqslant 1$ et tout événement H ne portant que sur les $k-1$ premières parties, $\mathrm{P}(G_k \mid A_k \cap H) = \dfrac{1}{5}$ et $\mathrm{P}(G_k \mid \overline{A_k} \cap H) = \dfrac{1}{10}$ lorsque ces probabilités sont définies ; en particulier $\mathrm{P}_{A_k}(G_k) = \dfrac{1}{5}$ et $\mathrm{P}_{\overline{A_k}}(G_k) = \dfrac{1}{10}$. La règle de changement donne $A_{k+1} = (A_k \cap G_k) \cup (\overline{A_k} \cap \overline{G_k})$.

1. On tire la première machine, puis, à chaque partie, on simule un succès de probabilité 0,2 ou 0,1 avec `random.random()`, qui suit la loi uniforme sur $[0\,,1[$; on change de machine en cas de perte.

```
import random

def jouer(n):
    machine = random.choice("AB")           # machine au hasard
    gains = 0
    for _ in range(n):
        p = 0.2 if machine == "A" else 0.1
        if random.random() < p:             # partie gagnee : on reste
            gains += 1
        else:                               # partie perdue : on change
            machine = "B" if machine == "A" else "A"
    return gains / n
```

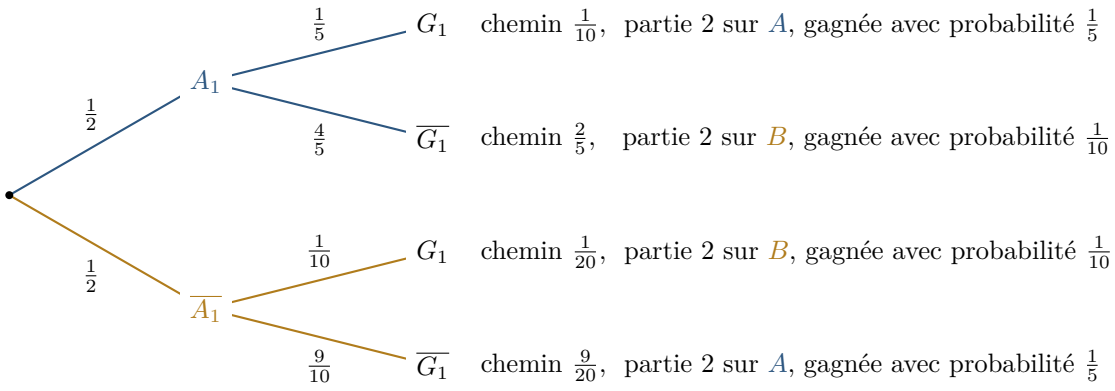
2. La formule des probabilités totales (théorème 5), appliquée au système complet $(A_1, \overline{A_1})$, donne

$$\mathrm{P}(G_1) = \mathrm{P}_{A_1}(G_1)\mathrm{P}(A_1) + \mathrm{P}_{\overline{A_1}}(G_1)\mathrm{P}(\overline{A_1}) = \frac{1}{5}\cdot\frac{1}{2} + \frac{1}{10}\cdot\frac{1}{2}, \qquad \boxed{\mathrm{P}(G_1) = \frac{3}{20}}$$

Pour la deuxième partie, on calcule d'abord la probabilité de jouer sur A : la deuxième partie a lieu sur A si la première a été gagnée sur A ou perdue sur B, d'où, par les probabilités composées,

$$\mathrm{P}(A_2) = \mathrm{P}(A_1 \cap G_1) + \mathrm{P}(\overline{A_1} \cap \overline{G_1}) = \frac{1}{2}\cdot\frac{1}{5} + \frac{1}{2}\cdot\frac{9}{10} = \frac{11}{20},$$

$$\mathrm{P}(G_2) = \frac{1}{5}\cdot\frac{11}{20} + \frac{1}{10}\cdot\frac{9}{20} = \frac{22+9}{200}, \qquad \boxed{\mathrm{P}(G_2) = \frac{31}{200}}$$



L'arbre décrit la première partie, et chaque feuille porte la probabilité de son chemin ; les branches bleues correspondent à la machine A et les branches ocre à la machine B. Les quatre feuilles totalisent $\dfrac{2+8+1+9}{20} = 1$, et la couleur de la machine utilisée pour la deuxième partie se lit au bout de chaque chemin : elle est conservée après un gain et change après une perte.

3. On cherche $\mathrm{P}_{G_2}(A_1) = \dfrac{\mathrm{P}(A_1 \cap G_2)}{\mathrm{P}(G_2)}$. Les deux premières feuilles de l'arbre donnent, par les probabilités composées,

$$\mathrm{P}(A_1 \cap G_2) = \mathrm{P}(A_1 \cap G_1 \cap G_2) + \mathrm{P}(A_1 \cap \overline{G_1} \cap G_2) = \frac{1}{10} \cdot \frac{1}{5} + \frac{2}{5} \cdot \frac{1}{10} = \frac{3}{50},$$

$$\text{d'où} \qquad \boxed{\mathrm{P}_{G_2}(A_1) = \frac{3/50}{31/200} = \frac{12}{31} \approx 0{,}387}$$

Ce résultat est inférieur à $\dfrac{1}{2}$: un gain à la deuxième partie rend plus probable un départ sur B, parce qu'un départ sur B, suivi d'une perte probable, conduit sur la bonne machine.

4.(a) La formule des probabilités totales, avec le système complet $(A_k, \overline{A_k})$, donne

$$\mathrm{P}(G_k) = \frac{1}{5}\,\mathrm{P}(A_k) + \frac{1}{10}(1 - \mathrm{P}(A_k)), \qquad \boxed{\mathrm{P}(G_k) = \frac{1 + \mathrm{P}(A_k)}{10}}$$

(b) L'égalité $A_{k+1} = (A_k \cap G_k) \cup (\overline{A_k} \cap \overline{G_k})$ est une réunion d'événements incompatibles, et chaque terme se calcule par les probabilités composées :

$$\mathrm{P}(A_{k+1}) = \frac{1}{5}\,\mathrm{P}(A_k) + \frac{9}{10}(1 - \mathrm{P}(A_k)), \qquad \boxed{\mathrm{P}(A_{k+1}) = -\frac{7}{10}\,\mathrm{P}(A_k) + \frac{9}{10}}$$

(c) La suite $a_k = \mathrm{P}(A_k)$ est arithmético-géométrique (chapitre 7, proposition 15), de raison $-\dfrac{7}{10}$ et de point fixe c solution de $c = -\dfrac{7}{10}c + \dfrac{9}{10}$, soit $c = \dfrac{9}{17}$. La suite $(a_k - c)$ est géométrique de raison $-\dfrac{7}{10}$, et $a_1 - c = \dfrac{1}{2} - \dfrac{9}{17} = -\dfrac{1}{34}$:

$$\boxed{\mathrm{P}(A_k) = \frac{9}{17} - \frac{1}{34}\Big(-\frac{7}{10}\Big)^{k-1}}, \qquad \boxed{\mathrm{P}(G_k) = \frac{13}{85} - \frac{1}{340}\Big(-\frac{7}{10}\Big)^{k-1}}$$

On contrôle : $k = 1$ donne $\dfrac{52 - 1}{340} = \dfrac{3}{20}$, et $k = 2$ donne $\dfrac{13}{85} + \dfrac{7}{3\,400} = \dfrac{527}{3\,400} = \dfrac{31}{200}$ ✓.

5. On somme la constante et la progression géométrique de raison $-\dfrac{7}{10} \neq 1$:

$$S_n = \frac{13n}{85} - \frac{1}{340} \cdot \frac{1 - (-7/10)^n}{1 + 7/10}, \qquad \boxed{S_n = \frac{13n}{85} - \frac{1 - (-7/10)^n}{578}, \qquad \frac{S_n}{n} \xrightarrow[n\to+\infty]{} \frac{13}{85} \approx 0{,}1529}$$

puisque $\big|1 - (-\dfrac{7}{10})^n\big| \leqslant 2$. Trois appels à `jouer(10**6)` ont renvoyé 0,1534, 0,1523 et 0,1528 : la proportion de parties gagnées est proche de $\dfrac{13}{85}$, avec un écart de l'ordre de 10^{-3}.

À retenir. Le nombre S_n est la somme des probabilités de gagner chaque partie ; la section 21.6 l'interprétera comme l'espérance du nombre de parties gagnées, et la loi faible des grands nombres (section 21.8) expliquera pourquoi une simulation en donne une valeur approchée. La limite $\dfrac{13}{85} = \dfrac{1}{10}(1 + \dfrac{9}{17})$ s'interprète : à long terme, le joueur joue sur A avec la probabilité $\dfrac{9}{17}$, soit environ neuf parties sur dix-sept, et la stratégie « on reste après un gain » fait un peu mieux que le choix d'une machine au hasard à chaque partie, qui donnerait $\dfrac{3}{20} = \dfrac{12{,}75}{85}$.

EXERCICE 51 ⋆⋆⋆

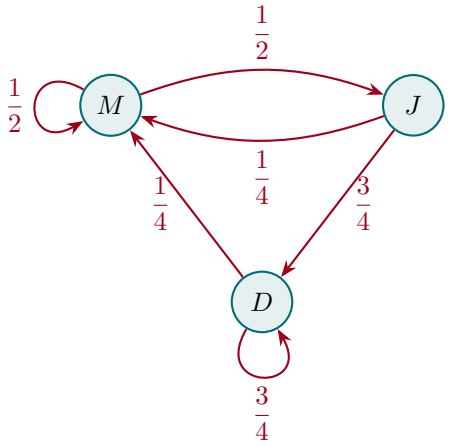
Modèle. On note $X_n \in \{M, J, D\}$ l'activité du chat dans la tranche $[5n\,, 5n+5]$, de sorte que $m_n = \mathrm{P}(X_n = M)$, et de même pour j_n et d_n. L'énoncé donne $X_0 = M$, soit $\mu_0 = (1,0,0)$, et les probabilités de transition, par exemple $\mathrm{P}_{(X_n=M)}(X_{n+1} = J) = \dfrac{1}{2}$, qui ne dépendent pas des tranches antérieures.

1. Les événements $(X_n = M)$, $(X_n = J)$ et $(X_n = D)$ forment un système complet (proposition 11). La formule des probabilités totales (théorème 5), avec la convention $\mathrm{P}_B(A)\mathrm{P}(B) = 0$ lorsque $\mathrm{P}(B) = 0$, donne

$$m_{n+1} = \sum_{s\in\{M,J,D\}} \mathrm{P}_{(X_n=s)}(X_{n+1} = M)\,\mathrm{P}(X_n = s) = \frac{1}{2}m_n + \frac{1}{4}j_n + \frac{1}{4}d_n,$$

et de même $j_{n+1} = \dfrac{1}{2}m_n$ et $d_{n+1} = \dfrac{3}{4}j_n + \dfrac{3}{4}d_n$. Ces trois égalités s'écrivent $\mu_{n+1} = \mu_n T$, la ligne d'indice s de T contenant les probabilités de passer de l'état s aux états M, J, D :

$$\boxed{T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}}, \qquad \mu_n = \mu_0 T^n \text{ pour tout } n \in \mathbb{N}.$$



Le graphe de la chaîne porte sur chaque flèche une probabilité de transition, comme celui de la chaîne à deux états de la section 21.5.7 ; les flèches issues d'un même état totalisent 1, ce qui traduit la somme 1 de chaque ligne de T.

2. On cherche les vecteurs colonnes V tels que $TV = \lambda V$. Pour $\lambda = 1$, la colonne $(1,1,1)^T$ convient, puisque chaque ligne de T a pour somme 1. Les deux dernières lignes de T sont égales : T n'est pas inversible, et la résolution de $TV = 0$ donne $V = (3,-3,-1)^T$. Enfin, la première ligne du système $TV = \dfrac{1}{4}V$ s'écrit $\dfrac{1}{4}v_1 + \dfrac{1}{2}v_2 = 0$, la troisième $\dfrac{1}{4}v_1 + \dfrac{1}{2}v_3 = 0$, et la deuxième en découle : le système équivaut à $v_2 = v_3 = -\dfrac{v_1}{2}$, et admet la solution $V = (2,-1,-1)^T$. Avec ces trois colonnes,

$$P = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & -3 \\ 1 & -1 & -1 \end{pmatrix}, \qquad \det P = -6, \qquad P^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} & -1 \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad \boxed{P^{-1}TP = \Delta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 \end{pmatrix}}$$

Ainsi $T = P\Delta P^{-1}$ et $T^n = P\Delta^n P^{-1}$ (chapitre 16, matrices semblables), avec $\Delta^n = \mathrm{diag}(1, 4^{-n}, 0)$ pour $n \geqslant 1$. Comme $\mu_0 = (1,0,0)$, la ligne μ_n est la première ligne de T^n, c'est-à-dire la première ligne de P, multipliée par Δ^n, puis par P^{-1} : pour $n \geqslant 1$,

$$\mu_n = (1, 2\cdot 4^{-n}, 0)P^{-1}, \qquad \boxed{m_n = \frac{1}{3} + \frac{2}{3}\cdot\frac{1}{4^n}, \quad j_n = \frac{1}{6} + \frac{4}{3}\cdot\frac{1}{4^n}, \quad d_n = \frac{1}{2} - \frac{2}{4^n} \qquad (n \geqslant 1)}$$

On contrôle pour $n=1$: $\mu_1 = (\frac{1}{2}, \frac{1}{2}, 0)$, première ligne de T ✓.

Piège. Ces formules sont fausses pour $n=0$: elles donneraient $m_0 = 1$, mais $j_0 = \dfrac{3}{2}$ et $d_0 = -\dfrac{3}{2}$. La raison est la valeur propre 0 : 0^n vaut 0 pour $n \geqslant 1$, mais 1 pour $n=0$, et la composante de μ_0 selon le troisième vecteur, qui disparaît dès le premier pas, n'a pas été comptée. L'erreur la plus répandue est d'écrire $\Delta^n = \mathrm{diag}(1, 4^{-n}, 0)$ sans préciser $n \geqslant 1$.

3. Comme 4^{-n} tend vers 0,

$$\boxed{(m_n, j_n, d_n) \xrightarrow[n\to+\infty]{} \Big(\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\Big)}$$

et l'on vérifie que la ligne $\mu = (\frac{1}{3}, \frac{1}{6}, \frac{1}{2})$ satisfait $\mu T = \mu$: c'est la loi stationnaire de la chaîne. À long terme, le chat mange avec la probabilité $\dfrac{1}{3}$, joue avec la probabilité $\dfrac{1}{6}$ et dort avec la probabilité $\dfrac{1}{2}$, ce que l'on interprète comme les fractions de son temps consacrées à chaque activité. La convergence est très rapide : au bout d'une heure, soit douze tranches, l'écart à la limite est au plus $2\cdot 4^{-12} \approx 1{,}2\cdot 10^{-7}$, et le chat a oublié qu'il avait commencé par manger, comme la chaîne à deux états de la section 21.5.7 oublie le temps du premier jour.

EXERCICE 52 ⋆⋆⋆

On note $\alpha_1 = 1$, $\alpha_2 = \dfrac{1}{2}$ et $\alpha_3 = 0$: pour un parent de génotype i, l'allèle transmis, choisi au hasard parmi ses deux allèles, est A avec la probabilité α_i (deux allèles A, un seul, ou aucun), et B avec la probabilité $1-\alpha_i$.

1. L'enfant est de génotype 1 si et seulement s'il reçoit l'allèle A de chacun de ses parents. Sachant $(P, M) = (i, j)$, les deux choix sont indépendants, d'où $\mathrm{P}_{((P,M)=(i,j))}(E=1) = \alpha_i\alpha_j$:

$\mathrm{P}_{((P,M)=(i,j))}(E=1)$	$j=1$	$j=2$	$j=3$
$i=1$	1	$\frac{1}{2}$	0
$i=2$	$\frac{1}{2}$	$\frac{1}{4}$	0
$i=3$	0	0	0

De même, l'enfant est de génotype 3 s'il reçoit B de chacun, avec la probabilité $(1-\alpha_i)(1-\alpha_j)$, et de génotype 2 dans les autres cas, avec la probabilité $\alpha_i(1-\alpha_j) + (1-\alpha_i)\alpha_j$.

2. Les événements $((P,M)=(i,j))$, pour $(i,j) \in \{1,2,3\}^2$, forment un système complet (proposition 11), et $\mathrm{P}((P,M)=(i,j)) = u_i(n)u_j(n)$ par indépendance de P et M. On abrège $u_i(n)$ en u_i. La formule des probabilités totales donne, en reconnaissant le développement d'un carré,

$$\mathrm{P}(E=1) = \sum_{i=1}^{3}\sum_{j=1}^{3} \alpha_i\alpha_j u_i u_j = \Big(\sum_{i=1}^{3} \alpha_i u_i\Big)^2,$$

et de même pour les autres valeurs, avec $\displaystyle\sum_i (1-\alpha_i)u_i = u_3 + \frac{1}{2}u_2$:

$$\boxed{\mathrm{P}(E=1) = \Big(u_1 + \frac{u_2}{2}\Big)^2, \qquad \mathrm{P}(E=2) = 2\Big(u_1 + \frac{u_2}{2}\Big)\Big(u_3 + \frac{u_2}{2}\Big), \qquad \mathrm{P}(E=3) = \Big(u_3 + \frac{u_2}{2}\Big)^2}$$

La somme vaut $(u_1+u_2+u_3)^2 = 1$ ✓.

3. Comme $u_1+u_2+u_3 = 1$, on a $u_3 + \dfrac{1}{2}u_2 = 1-\theta(n)$. En notant $\theta = \theta(n)$,

$$\boxed{u_1(n+1) = \theta^2, \qquad u_2(n+1) = 2\theta(1-\theta), \qquad u_3(n+1) = (1-\theta)^2}$$

Ce sont les probabilités de la loi $\mathcal{B}(2,\theta)$ pour le nombre d'allèles A de l'enfant, qui vaut 2, 1 ou 0 selon que son génotype est 1, 2 ou 3 : tout se passe comme si l'enfant recevait deux allèles tirés indépendamment dans le stock de la génération n, où la proportion de A est θ.

4. On calcule

$$\theta(n+1) = u_1(n+1) + \frac{1}{2}u_2(n+1) = \theta^2 + \theta(1-\theta), \qquad \boxed{\theta(n+1) = \theta(n)}$$

La suite $(\theta(n))$ est donc constante, égale à $\theta(0)$, et la question 3 donne, pour tout $n \geqslant 1$,

$$\boxed{u(n) = \Big(\theta(0)^2,\ 2\theta(0)(1-\theta(0)),\ \big(1-\theta(0)\big)^2\Big), \qquad \text{proportion de malades } u_3(n) = \big(1-\theta(0)\big)^2}$$

Les proportions sont les mêmes à toutes les générations $n \geqslant 1$, mais pas forcément à la génération 0. Par exemple, si $u(0) = (\frac{1}{2}, 0, \frac{1}{2})$, alors $\theta(0) = \frac{1}{2}$ et $u(n) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ pour tout $n \geqslant 1$: la moitié de la population initiale est malade, mais un quart seulement des générations suivantes.

À retenir. Ce résultat, publié indépendamment par G. H. Hardy et W. Weinberg en 1908, est connu sous le nom d'équilibre de Hardy-Weinberg : dans ce modèle, le hasard des unions ne fait ni disparaître ni progresser une maladie récessive, et l'équilibre est atteint dès la première génération. Le calcul repose sur l'indépendance des deux parents et des deux transmissions ; il cesse de valoir si les malades ont moins d'enfants, ou si les unions ne se font pas au hasard.

EXERCICE 53 ⋆⋆⋆

Modèle. On note U le numéro de l'urne choisie, de loi uniforme sur $[\![0\,,r]\!]$. Sachant $(U=k)$, les tirages sont indépendants et chacun donne une boule rouge avec la probabilité $\dfrac{k}{r}$; pour $m \in [\![1\,,n+1]\!]$, on a donc $\mathrm{P}_{(U=k)}(R_1 \cap \cdots \cap R_m) = (\frac{k}{r})^m$. La formule des probabilités totales, avec le système complet $\big((U=k)\big)_{0\leqslant k\leqslant r}$, donne

$$\pi_m = \mathrm{P}(R_1 \cap \cdots \cap R_m) = \frac{1}{r+1}\sum_{k=0}^{r}\Big(\frac{k}{r}\Big)^m = \frac{s_m}{(r+1)\,r^m}, \qquad s_m = \sum_{k=0}^{r} k^m.$$

Ces probabilités sont strictement positives, grâce au terme $k=r$.

1. Avec $s_1 = \dfrac{r(r+1)}{2}$ et $s_2 = \dfrac{r(r+1)(2r+1)}{6}$ (chapitre 3), on obtient $\pi_1 = \dfrac{1}{2}$ et $\pi_2 = \dfrac{2r+1}{6r}$, d'où

$$\boxed{\mathrm{P}(R_1) = \frac{1}{2}, \qquad \mathrm{P}_{R_1}(R_2) = \frac{\mathrm{P}(R_1\cap R_2)}{\mathrm{P}(R_1)} = \frac{2r+1}{3r}}$$

Pour $n \geqslant 2$, la définition de la probabilité conditionnelle donne de même

$$\boxed{p_{n,r} = \mathrm{P}(R_{n+1} \mid R_1 \cap \cdots \cap R_n) = \frac{\pi_{n+1}}{\pi_n} = \frac{s_{n+1}}{r\,s_n} = \frac{\sum_{k=0}^{r} k^{n+1}}{r\sum_{k=0}^{r} k^n}}$$

formule qui redonne $\mathrm{P}_{R_1}(R_2)$ pour $n=1$. Pour $r=3$ et $n=2$, on trouve $p_{2,3} = \dfrac{36}{3\times 14} = \dfrac{6}{7}$.

2. Pour $m \geqslant 1$, le terme d'indice $k=0$ est nul, et

$$\frac{s_m}{r^{m+1}} = \frac{1}{r}\sum_{k=1}^{r}\Big(\frac{k}{r}\Big)^m$$

est la somme de Riemann $R'_r(f)$ de la fonction continue $f : t \mapsto t^m$ sur $[0\,,1]$, pour la subdivision régulière en r intervalles. Par le théorème 9 du chapitre 18, elle tend vers $\displaystyle\int_0^1 t^m\,\mathrm{d}t = \frac{1}{m+1}$ quand r tend vers $+\infty$. On écrit alors

$$p_{n,r} = \frac{s_{n+1}/r^{n+2}}{s_n/r^{n+1}} \xrightarrow[r\to+\infty]{} \frac{1/(n+2)}{1/(n+1)}, \qquad \boxed{\lim_{r\to+\infty} p_{n,r} = \frac{n+1}{n+2}}$$

Pour $r = 1\,000$ et $n = 3$, on trouve $p_{3,1000} \approx 0{,}8004$, contre $\dfrac{4}{5}$ ✓.

Comparaison. On a $\mathrm{P}(R_1) = \dfrac{1}{2}$, alors que la probabilité de tirer une boule rouge après n boules rouges vaut à peu près $\dfrac{n+1}{n+2}$, qui croît vers 1. Les tirages ne sont pas indépendants : ils le sont sachant l'urne, mais l'urne est inconnue, et chaque boule rouge rend plus probables les urnes riches en boules rouges. C'est la loi de succession de Laplace : n'ayant aucune information sur la proportion de boules rouges, sinon qu'elle est « uniforme », on estime la probabilité d'un nouveau succès après n succès par $\dfrac{n+1}{n+2}$, et non par 1.

Piège. L'erreur la plus répandue est de croire que les R_j sont indépendants, puisque les tirages ont lieu avec remise, et d'en conclure que $p_{n,r} = \mathrm{P}(R_1) = \dfrac{1}{2}$. L'indépendance sachant $(U = k)$, pour chaque k, n'entraîne pas l'indépendance sous P : déjà pour $n = 1$, $\mathrm{P}_{R_1}(R_2) = \dfrac{2r+1}{3r} > \dfrac{1}{2}$. L'exercice 65 retrouvera exactement le nombre $\dfrac{n+1}{n+2}$ avec l'urne de Pólya.

EXERCICE 54 ⋆⋆⋆

Modèle. Un tournoi est déterminé par le vainqueur de chacun des $\binom{n}{2}$ matchs ; on munit l'ensemble des $2^{\binom{n}{2}}$ tournois de la probabilité uniforme, ce qui revient, par la proposition 17, à considérer des matchs mutuellement indépendants, chacun gagné par l'une ou l'autre équipe avec la probabilité $\dfrac{1}{2}$. Un événement qui fixe l'issue de m matchs donnés a donc la probabilité 2^{-m}.

1.(a) Pour $i \in [\![1\,,n]\!]$, l'événement E_i « l'équipe i a battu toutes les autres » fixe l'issue de $n-1$ matchs : $\mathrm{P}(E_i) = 2^{-(n-1)}$. Les E_i sont deux à deux incompatibles, puisque deux équipes ne peuvent pas s'être battues mutuellement. Par additivité,

$$\boxed{\mathrm{P}(\text{une équipe a battu toutes les autres}) = \frac{n}{2^{n-1}}}$$

(b) Un classement est une permutation $\sigma \in \mathfrak{S}_n$, l'équipe $\sigma(1)$ étant classée première. L'événement F_σ « pour $a < b$, l'équipe $\sigma(a)$ a battu l'équipe $\sigma(b)$ » fixe l'issue des $\binom{n}{2}$ matchs, et $\mathrm{P}(F_\sigma) = 2^{-\binom{n}{2}}$. Si $\sigma \neq \sigma'$, deux équipes sont classées dans des ordres différents par σ et σ', et leur match ne peut pas avoir les deux issues : les F_σ sont deux à deux incompatibles. Par le théorème 1 et l'additivité,

$$\boxed{\mathrm{P}(\text{les équipes peuvent être classées}) = \frac{n!}{2^{\binom{n}{2}}}}$$

Pour $n = 3$, on trouve $\dfrac{6}{8} = \dfrac{3}{4}$: les deux tournois exclus sont ceux où 1 bat 2, 2 bat 3 et 3 bat 1, ou l'inverse. Pour $n = 10$, la probabilité vaut environ $1{,}0 \cdot 10^{-7}$.

(c) Pour $j \notin K$, l'événement H_j « l'équipe j a battu toutes les équipes de K » fixe l'issue de k matchs, et $\mathrm{P}(H_j) = 2^{-k}$. Les événements H_j, pour $j \notin K$, portent sur des ensembles de matchs deux à deux disjoints, les matchs entre j et les équipes de K : ils sont mutuellement indépendants par le lemme des coalitions (théorème 7) et la proposition 15, ainsi que leurs contraires (proposition 10 (iii)). La probabilité qu'aucune des $n-k$ équipes hors de K ne convienne vaut donc $(1-2^{-k})^{n-k}$, et

$$\boxed{\mathrm{P}(\text{une équipe hors de } K \text{ a battu toutes celles de } K) = 1 - (1-2^{-k})^{n-k}}$$

2. On suppose $n > k$. Le tournoi n'est pas k-indécis si et seulement s'il existe un ensemble K de k équipes qu'aucune autre équipe n'a toutes battues. Par la sous-additivité (proposition 7 (iv)) et la

question 1(c), en sommant sur les $\binom{n}{k}$ ensembles K,

$$1 - \mathrm{P}(\text{le tournoi est } k\text{-indécis}) \leqslant \binom{n}{k}(1-2^{-k})^{n-k} \leqslant \frac{n^k q^n}{q^k}, \qquad q = 1-2^{-k} \in]0\,,1[\,,$$

puisque $\binom{n}{k} = \dfrac{n(n-1)\cdots(n-k+1)}{k!} \leqslant n^k$. En écrivant $q^n = \mathrm{e}^{-\gamma n}$ avec $\gamma = -\ln q > 0$, les croissances comparées (chapitre 10, proposition 2) donnent $n^k q^n \to 0$, d'où

$$\boxed{\mathrm{P}(\text{un tournoi aléatoire de taille } n \text{ est } k\text{-indécis}) \xrightarrow[n\to+\infty]{} 1}$$

3. Pour n assez grand, cette probabilité est strictement positive, donc l'événement correspondant n'est pas vide : il existe un tournoi k-indécis à n équipes. C'est la *méthode probabiliste* : on prouve l'existence d'un objet sans le construire, en montrant qu'un objet tiré au hasard a la propriété voulue avec une probabilité non nulle. Pour $k = 2$, il suffit que la majoration de la question 2 soit strictement inférieure à 1 ; le programme suivant cherche le plus petit tel n.

```
from math import comb

def taille(k):
    n = k + 1
    while comb(n, k) * (1 - 2**-k)**(n - k) >= 1:
        n += 1
    return n

print(taille(2), taille(3))       # affiche 21 91
```

$$\boxed{\text{pour } n = 21,\ \binom{21}{2}\left(\frac{3}{4}\right)^{19} \approx 0{,}888 < 1 : \text{il existe un tournoi 2-indécis à 21 équipes}}$$

tandis que $\binom{20}{2}(\frac{3}{4})^{18} \approx 1{,}071$ ne permet pas de conclure pour $n = 20$. Le calcul en fractions exactes donne les mêmes seuils, 21 pour $k = 2$ et 91 pour $k = 3$.
À retenir. La majoration ne donne qu'une taille suffisante. Par une recherche exhaustive, aucun des $2^{15} = 32\,768$ tournois à 6 équipes n'est 2-indécis, et le tournoi à 7 équipes où i bat j lorsque $j - i$ est congru à 1, 2 ou 4 modulo 7 l'est : la plus petite taille est 7. La méthode probabiliste est grossière, mais elle s'applique à tout k, alors qu'une construction explicite est délicate.

Variables aléatoires et lois

EXERCICE 55 ⋆

1. Les variables X_1 et X_2 prennent leurs valeurs dans $\{0,1\}$.

— Comme $0^2 = 0$ et $1^2 = 1$, on a $X_1^2 = X_1$, et $\boxed{X_1^2 \sim \mathcal{B}(p_1)}$.

— La variable $1 - X_1$ est à valeurs dans $\{0,1\}$, et $\mathrm{P}(1 - X_1 = 1) = \mathrm{P}(X_1 = 0) = 1 - p_1$: $\boxed{1 - X_1 \sim \mathcal{B}(1-p_1)}$.

— Le produit X_1X_2 est à valeurs dans $\{0,1\}$, et il vaut 1 si et seulement si $X_1 = X_2 = 1$; par indépendance, $\mathrm{P}(X_1X_2 = 1) = \mathrm{P}(X_1 = 1)\mathrm{P}(X_2 = 1)$, et $\boxed{X_1X_2 \sim \mathcal{B}(p_1p_2)}$.

En termes d'événements, ces trois variables sont les indicatrices de A_1, de $\overline{A_1}$ et de $A_1 \cap A_2$, où $A_i = (X_i = 1)$.

2. Soit $t = \mathrm{P}(X = 1, Y = 1)$. Par la proposition 14, les lois marginales s'obtiennent en sommant les lignes et les colonnes de la loi conjointe, ce qui détermine les trois autres cases :

	$Y=0$	$Y=1$	
$X=0$	$1-p-q+t$	$q-t$	$1-p$
$X=1$	$p-t$	t	p
	$1-q$	q	1

Ces quatre nombres doivent être positifs, ce qui donne $t \geqslant 0$, $t \leqslant p$, $t \leqslant q$ et $t \geqslant p+q-1$:

$$\boxed{\max(0, p+q-1) \leqslant t \leqslant \min(p,q)}$$

C'est l'encadrement de l'exercice 23, appliqué aux événements $(X = 1)$ et $(Y = 1)$. Réciproquement, pour tout t de cet intervalle, les quatre nombres du tableau forment une distribution de probabilités sur $\{0,1\}^2$, de somme 1 ; en munissant $\Omega = \{0,1\}^2$ de la probabilité qu'elle définit (théorème 4) et en prenant pour X et Y les deux projections, on obtient un couple qui a cette loi conjointe. Toutes ces lois conjointes sont donc possibles.

Par la proposition 15, X et Y sont indépendantes si et seulement si chaque case est le produit des marges, et la case $(1,1)$ impose $t = pq$. Réciproquement, pour $t = pq$, les trois autres cases valent $p(1-q)$, $(1-p)q$ et $(1-p)(1-q)$, produits des marges, et

$$\boxed{X \perp\!\!\!\perp Y \iff t = pq}$$

Cette valeur appartient bien à l'intervalle : $pq \leqslant \min(p,q)$, et $pq - (p+q-1) = (1-p)(1-q) \geqslant 0$. L'exemple illustre la remarque de la section 21.5.4 : les lois marginales $\mathcal{B}(p)$ et $\mathcal{B}(q)$ sont compatibles avec toute une famille de lois conjointes, à un paramètre, dont une seule rend les variables indépendantes.

EXERCICE 56 ⋆⋆

1. On prend $\Omega = [\![1\,,6]\!]^2$, muni de la probabilité uniforme. L'événement « résultats égaux » contient les six couples (i,i) :

$$\boxed{\mathrm{P}(D_1 = D_2) = \frac{6}{36} = \frac{1}{6}}$$

2. Les résultats D_1 et D_2 sont indépendants et de même loi (p_i). L'événement $(D_1 = D_2)$ est la réunion disjointe des événements $(D_1 = i, D_2 = i)$, d'où

$$\mathrm{P}(D_1 = D_2) = \sum_{i=1}^{6} \mathrm{P}(D_1 = i)\mathrm{P}(D_2 = i) = \sum_{i=1}^{6} p_i^2.$$

En développant les carrés et en utilisant $\displaystyle\sum_i p_i = 1$,

$$\sum_{i=1}^{6}\left(p_i - \frac{1}{6}\right)^2 = \sum_{i=1}^{6} p_i^2 - \frac{2}{6}\sum_{i=1}^{6} p_i + \frac{6}{36} = \sum_{i=1}^{6} p_i^2 - \frac{1}{6}, \qquad \text{d'où} \qquad \boxed{\mathrm{P}(D_1 = D_2) = \frac{1}{6} + \sum_{i=1}^{6}\left(p_i - \frac{1}{6}\right)^2 \geqslant \frac{1}{6}}$$

Il y a égalité si et seulement si tous les p_i valent $\dfrac{1}{6}$, c'est-à-dire si les dés sont équilibrés. On peut aussi appliquer l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^6$ muni du produit scalaire canonique (chapitre 19, théorème 1) aux vecteurs $(p_1, \ldots, p_6)$ et $(1, \ldots, 1)$: elle donne $1 = \left(\sum p_i\right)^2 \leqslant 6\sum p_i^2$, avec égalité si et seulement si les deux vecteurs sont colinéaires, ce qui conduit au même cas d'égalité.

À retenir. Piper deux dés de la même façon augmente toujours la probabilité d'un double. L'identité obtenue dit plus : l'excès de probabilité est exactement la somme des carrés des écarts à la loi uniforme, et il est d'autant plus grand que les dés sont fortement pipés. Si les dés sont pipés différemment, le résultat tombe : un dé qui ne donne que des 1 et un dé qui ne donne que des 2 ne font jamais de double.

EXERCICE 57 ⋆⋆

L'hypothèse $X \sim -X$ donne, pour tout réel x, $\mathrm{P}(X=-x)=\mathrm{P}(-X=x)=\mathrm{P}(X=x)$.

1. La variable $|X|$ prend ses valeurs dans une partie finie F de $\mathbb{R}_+$. On munit $\Omega'=F\times\{-1,1\}$ de la probabilité définie par la distribution $p_{(y,e)}=\dfrac{1}{2}\mathrm{P}(|X|=y)$ (proposition 17, avec la loi de $|X|$ sur F et la loi uniforme sur $\{-1,1\}$), et l'on note Y et ε les deux projections. Par la proposition 17, Y et ε sont indépendantes, $Y\sim|X|$ et ε suit la loi uniforme sur $\{-1,1\}$. Comparons les lois de εY et de X.

— Pour $x\neq 0$, on a $(\varepsilon Y=x)=(Y=|x|)\cap(\varepsilon=\mathrm{signe}(x))$, puisque $Y\geqslant 0$. Par indépendance, et grâce à la symétrie de la loi de X,

$$\mathrm{P}'(\varepsilon Y=x)=\frac{1}{2}\mathrm{P}(|X|=|x|)=\frac{1}{2}(\mathrm{P}(X=x)+\mathrm{P}(X=-x))=\mathrm{P}(X=x).$$

— Pour $x=0$, on a $\mathrm{P}'(\varepsilon Y=0)=\mathrm{P}'(Y=0)=\mathrm{P}(|X|=0)=\mathrm{P}(X=0)$.

Les deux variables ont les mêmes probabilités $\mathrm{P}(\,\cdot=x)$ pour tout réel x, donc la même loi (proposition 12) :

$$\boxed{(\Omega',\mathrm{P}')=(F\times\{-1,1\},\ \mathrm{P}'),\qquad Y=|X|\text{ en loi},\quad \varepsilon\text{ uniforme sur }\{-1,1\},\quad Y\perp\!\!\!\perp\varepsilon,\qquad \varepsilon Y\sim X}$$

Une variable de loi symétrique s'obtient donc en tirant sa valeur absolue, puis un signe au hasard, indépendamment.

2. La réponse est non. Soit $\Omega=\{\omega_1,\omega_2,\omega_3\}$, muni de la probabilité uniforme, et X définie par $X(\omega_1)=0$, $X(\omega_2)=1$, $X(\omega_3)=-1$: X suit la loi uniforme sur $\{-1,0,1\}$, et $X\sim -X$. Supposons qu'il existe $Y:\Omega\to\mathbb{R}_+$ et $\varepsilon:\Omega\to\{-1,1\}$ indépendantes telles que $\varepsilon Y\sim X$. Comme $Y\geqslant 0$, on a $(\varepsilon Y=1)=(\varepsilon=1,Y=1)$ et $(\varepsilon Y=-1)=(\varepsilon=-1,Y=1)$, d'où, par indépendance,

$$\mathrm{P}(\varepsilon=1)\,\mathrm{P}(Y=1)=\frac{1}{3}=\mathrm{P}(\varepsilon=-1)\,\mathrm{P}(Y=1).$$

Ainsi $\mathrm{P}(Y=1)\neq 0$, et $\mathrm{P}(\varepsilon=1)=\mathrm{P}(\varepsilon=-1)=\dfrac{1}{2}$. Or toute probabilité d'un événement de Ω est un multiple de $\dfrac{1}{3}$: c'est impossible.

$$\boxed{\text{on ne peut pas toujours choisir }(\Omega',\mathrm{P}')=(\Omega,\mathrm{P})}$$

Une recherche exhaustive sur les $2^3\times 3^3$ couples (ε,Y), avec Y à valeurs dans $\{0,1,2\}$, confirme qu'aucun ne convient ✓.

Piège. L'erreur la plus répandue est de prendre $Y=|X|$ et $\varepsilon=\mathrm{signe}(X)$ sur Ω lui-même : on a bien $X=\varepsilon|X|$ là où $X\neq 0$, mais ε et $|X|$ ne sont pas indépendantes en général. Dans l'exemple de la question 2, en posant par exemple $\varepsilon(\omega_1)=1$, on obtient $\mathrm{P}(\varepsilon=1,|X|=0)=\dfrac{1}{3}$ alors que $\mathrm{P}(\varepsilon=1)\mathrm{P}(|X|=0)=\dfrac{2}{3}\cdot\dfrac{1}{3}$; plus profondément, l'univers est trop pauvre pour porter un signe de probabilité $\dfrac{1}{2}$. C'est pour cette raison que la question 1 agrandit l'univers, comme le permet la proposition 17.

EXERCICE 58 ⋆⋆

Modèle. On numérote les boules et l'on munit l'ensemble des $(b+r)!$ ordres de sortie de la probabilité uniforme. L'ensemble des rangs de sortie des boules rouges est une partie à r éléments de $[\![1\,,b+r]\!]$, et chacune de ces parties correspond au même nombre $r!\,b!$ d'ordres (on permute les rouges entre elles et les blanches entre elles) : par le lemme des bergers, cette partie suit la loi uniforme sur les $\binom{b+r}{r}$ parties à r éléments. Le rang N est le plus petit élément de cette partie, et $N \in [\![1\,,b+1]\!]$, puisque les $b+1$ premiers tirages ne peuvent pas être tous blancs.

Loi de N. Soit $k \in [\![1\,,b+1]\!]$. L'événement $(N=k)$ est formé des parties qui contiennent k, ne contiennent aucun élément de $[\![1\,,k-1]\!]$, et contiennent $r-1$ éléments de $[\![k+1\,,b+r]\!]$, ensemble à $b+r-k$ éléments :

$$\boxed{\mathrm{P}(N=k) = \frac{\binom{b+r-k}{r-1}}{\binom{b+r}{r}}, \qquad k \in [\![1\,,b+1]\!]}$$

On retrouve ce résultat par les probabilités composées (proposition 9), en tirant d'abord $k-1$ boules blanches, puis une rouge :

$$\mathrm{P}(N=k) = \frac{b}{b+r} \cdot \frac{b-1}{b+r-1} \cdots \frac{b-k+2}{b+r-k+2} \cdot \frac{r}{b+r-k+1},$$

et les deux formules coïncident, comme on le vérifie en simplifiant les factorielles. Par exemple, pour $b=3$ et $r=2$, la loi de N est $(\frac{4}{10}, \frac{3}{10}, \frac{2}{10}, \frac{1}{10})$ sur $[\![1\,,4]\!]$: elle décroît, puisqu'il est d'autant moins probable d'attendre longtemps.

Identité. Les probabilités $\mathrm{P}(N=k)$, pour $k \in [\![1\,,b+1]\!]$, ont pour somme 1. En posant $j = b+r-k$, qui parcourt $[\![r-1\,,b+r-1]\!]$, on obtient

$$\boxed{\sum_{k=1}^{b+1} \binom{b+r-k}{r-1} = \sum_{j=r-1}^{b+r-1} \binom{j}{r-1} = \binom{b+r}{r}}$$

C'est la formule de la crosse de hockey, utilisée dans l'exercice 35 de la feuille du chapitre 3, avec $p = r-1$ et $n = b+r-1$. Elle reçoit ici une démonstration probabiliste : on classe les parties à r éléments de $[\![1\,,b+r]\!]$ selon leur plus petit élément.

EXERCICE 59 ⋆⋆

On note $q = 1-p$ et $\pi_k = \mathrm{P}(X=k) = \binom{n}{k} p^k q^{n-k} > 0$ pour $k \in [\![0\,,n]\!]$. Pour $k \in [\![0\,,n-1]\!]$, la relation $\binom{n}{k+1} = \frac{n-k}{k+1}\binom{n}{k}$ donne

$$\boxed{\frac{\pi_{k+1}}{\pi_k} = \frac{(n-k)p}{(k+1)q}} \qquad \text{et} \qquad \frac{\pi_{k+1}}{\pi_k} > 1 \iff (n-k)p > (k+1)(1-p) \iff k < (n+1)p - 1,$$

avec égalité du quotient à 1 si et seulement si $k = (n+1)p-1$. On note $m = \lfloor (n+1)p \rfloor$, qui appartient à $[\![0\,,n]\!]$ puisque $0 < (n+1)p < n+1$.

- Pour $k \geqslant m$, on a $k > (n+1)p - 1$, donc $\pi_{k+1} < \pi_k$: la suite (π_k) est strictement décroissante sur $[\![m\,,n]\!]$.
- Pour $k \leqslant m-1$, on a $k \leqslant (n+1)p - 1$, donc $\pi_{k+1} \geqslant \pi_k$, avec égalité seulement si $k = m-1$ et $(n+1)p = m$.

La suite (π_k) croît donc jusqu'au rang m, puis décroît strictement :

$$\boxed{\begin{array}{l} \text{si } (n+1)p \notin \mathbb{N}, \text{ la valeur la plus probable est } \lfloor (n+1)p \rfloor ; \\ \text{si } (n+1)p \in \mathbb{N}, \text{ il y en a deux, } (n+1)p - 1 \text{ et } (n+1)p \end{array}}$$

Dans le second cas, $(n+1)p$ appartient à $[\![1\,,n]\!]$, et les deux valeurs sont dans $[\![0\,,n]\!]$. Pour $\mathcal{B}(10,0{,}3)$, on a $(n+1)p = 3{,}3$ et la valeur la plus probable est 3, comme sur la figure de la section 21.5.3 ; pour $\mathcal{B}(9,\frac{1}{2})$, on a $(n+1)p = 5$, et 4 et 5 sont également probables.

Pour $p = \dfrac{1}{2}$, on a $\pi_k = \dfrac{1}{2^n}\binom{n}{k}$, et l'on retrouve le résultat de l'exercice 28 de la feuille du chapitre 3 : les coefficients $\binom{n}{k}$ croissent puis décroissent, avec un maximum en $\dfrac{n}{2}$ si n est pair, et deux maxima égaux en $\dfrac{n-1}{2}$ et $\dfrac{n+1}{2}$ si n est impair. Le facteur $(\frac{p}{q})^k$ décale le maximum de $\dfrac{n}{2}$ vers np : de $(n+1)p-1 \leqslant m \leqslant (n+1)p$, on tire $np-q \leqslant m \leqslant np+p$, et la valeur la plus probable est à moins de 1 de np, nombre que la section 21.6 identifiera à l'espérance.

EXERCICE 60 ⋆⋆

Couplage. Soit $n \geqslant 1$. Par la proposition 17, il existe un espace probabilisé fini, par exemple $\{0,1\}^{n+1}$ muni de la probabilité produit, portant des variables B_1, …, B_{n+1} mutuellement indépendantes de loi $\mathcal{B}(p)$. On pose

$$Y_n = B_1 + \cdots + B_n, \qquad Y_{n+1} = Y_n + B_{n+1}.$$

Par le théorème 8, $Y_n \sim \mathcal{B}(n,p)$ et $Y_{n+1} \sim \mathcal{B}(n+1,p)$. Comme $B_{n+1} \geqslant 0$, on a $Y_n \leqslant Y_{n+1}$ en chaque issue, donc $(Y_n \geqslant k) \subset (Y_{n+1} \geqslant k)$, et, par croissance de la probabilité (proposition 7 (iii)), $\mathrm{P}(Y_n \geqslant k) \leqslant \mathrm{P}(Y_{n+1} \geqslant k)$.

Retour aux X_n. La probabilité $\mathrm{P}(X_n \geqslant k)$ ne dépend que de la loi de X_n, qui est celle de Y_n ; de même pour X_{n+1} et Y_{n+1}. Donc

$$\boxed{\mathrm{P}(X_n \geqslant k) = \mathrm{P}(Y_n \geqslant k) \leqslant \mathrm{P}(Y_{n+1} \geqslant k) = \mathrm{P}(X_{n+1} \geqslant k)}$$

et la suite $\big(\mathrm{P}(X_n \geqslant k)\big)_{n\geqslant 1}$ est croissante.

Le couplage donne même l'accroissement. On a $(Y_{n+1} \geqslant k) \setminus (Y_n \geqslant k) = (Y_n = k-1) \cap (B_{n+1} = 1)$, et Y_n est indépendante de B_{n+1} par le lemme des coalitions (théorème 7), d'où, pour $k \geqslant 1$,

$$\boxed{\mathrm{P}(X_{n+1} \geqslant k) - \mathrm{P}(X_n \geqslant k) = \binom{n}{k-1} p^k (1-p)^{n-k+1} \geqslant 0}$$

avec la convention $\binom{n}{k-1} = 0$ si $k-1 > n$; pour $k = 0$, la suite est constante, égale à 1.

Piège. Les variables X_n de l'énoncé sont données séparément, peut-être sur des espaces probabilisés différents, et l'inégalité $X_n \leqslant X_{n+1}$ n'a aucun sens. L'erreur la plus répandue est de l'écrire quand même. Le couplage contourne la difficulté : on construit, sur un même univers, des variables ayant les lois voulues et comparables issue par issue, ce qui suffit pour conclure, puisque l'énoncé ne porte que sur des lois. La comparaison directe des sommes $\displaystyle\sum_{j\geqslant k} \binom{n}{j} p^j q^{n-j}$ serait nettement plus pénible.

EXERCICE 61 ⋆⋆

Modèle. On note $E = [\![1\,,n]\!]$. Par indépendance (proposition 15), le r-uplet $(X_1, \dots, X_r)$ suit la loi uniforme sur $\mathcal{P}(E)^r$, chaque r-uplet ayant la probabilité $(\frac{1}{2^n})^r$.

1. Le couple (X_1, X_2) suit la loi uniforme sur les 4^n couples de parties de E. L'exercice 39 de la feuille du chapitre 1 a dénombré 3^n couples (X, Y) tels que $X \subset Y$, et 3^n tels que $X \cap Y = \varnothing$; la question 3 de l'exercice 10 en a dénombré $n\,3^{n-1}$ tels que $X \cap Y$ soit un singleton. D'où

$$\boxed{\mathrm{P}(X_1 \subset X_2) = \Big(\frac{3}{4}\Big)^n, \qquad \mathrm{P}(X_1 \cap X_2 = \varnothing) = \Big(\frac{3}{4}\Big)^n, \qquad \mathrm{P}(|X_1 \cap X_2| = 1) = \frac{n\,3^{n-1}}{4^n}}$$

Les deux premières probabilités sont égales, ce qui s'explique : $X_2 \mapsto E \setminus X_2$ est une bijection de $\mathcal{P}(E)$, qui conserve la loi uniforme, et $X_1 \subset X_2$ équivaut à $X_1 \cap (E \setminus X_2) = \varnothing$. Toutes trois tendent vers 0 : deux parties prises au hasard sont rarement emboîtées.

2. L'application $X \mapsto (\mathbf{1}_X(1), \dots, \mathbf{1}_X(n))$ est une bijection de $\mathcal{P}(E)$ sur $\{0,1\}^n$ (chapitre 2). La loi uniforme sur $\mathcal{P}(E)^r$ correspond donc à la loi uniforme sur $\{0,1\}^{nr}$, et, par la proposition 17, les nr variables $\mathbf{1}_{X_j}(i)$, pour $i \in E$ et $j \in [\![1\,,r]\!]$, sont mutuellement indépendantes de loi $\mathcal{B}(\frac{1}{2})$. Pour $i \in E$, on pose

$$Z_i = \mathbf{1}_{X_1 \cap \dots \cap X_r}(i) = \prod_{j=1}^{r} \mathbf{1}_{X_j}(i).$$

Chaque Z_i vaut 1 si et seulement si les r facteurs valent 1, avec la probabilité 2^{-r} par indépendance : $Z_i \sim \mathcal{B}(2^{-r})$. Les variables Z_1, …, Z_n sont des fonctions des paquets disjoints $\big(\mathbf{1}_{X_j}(i)\big)_{1 \leqslant j \leqslant r}$, donc mutuellement indépendantes par le lemme des coalitions (théorème 7). Enfin $|X_1 \cap \dots \cap X_r| = Z_1 + \dots + Z_n$, et le théorème 8 conclut :

$$\boxed{|X_1 \cap X_2 \cap \dots \cap X_r| \sim \mathcal{B}(n, 2^{-r})}$$

Pour $r = 2$, on retrouve $\mathrm{P}(|X_1 \cap X_2| = 0) = (\frac{3}{4})^n$ et $\mathrm{P}(|X_1 \cap X_2| = 1) = n \cdot \frac{1}{4}(\frac{3}{4})^{n-1}$, conformément à la question 1 ✓. Avec $r = 1$, le même raisonnement montre que le cardinal d'une partie uniforme suit la loi $\mathcal{B}(n, \frac{1}{2})$.

3. Il y a $4^n - 2^n$ couples de parties distinctes, et, parmi les 3^n couples tels que $A \subset B$, les 2^n couples (A, A) sont exclus :

$$\boxed{\mathrm{P}(A \subset B) = \frac{3^n - 2^n}{4^n - 2^n}}$$

Pour $n = 1$, on obtient $\frac{1}{2}$, contre $\frac{3}{4}$ à la question 1, et pour $n = 4$, $\frac{13}{48} \approx 0{,}271$ contre $\frac{81}{256} \approx 0{,}316$. Cette probabilité est toujours strictement inférieure à $(\frac{3}{4})^n$, puisque

$$4^n(3^n - 2^n) - 3^n(4^n - 2^n) = 6^n - 8^n < 0 :$$

le tirage sans remise exclut les couples (A, A), qui réalisent tous l'inclusion. Le rapport des deux probabilités tend vers 1, puisque $\frac{2^n}{3^n}$ et $\frac{2^n}{4^n}$ tendent vers 0 : pour n grand, tirer avec ou sans remise revient presque au même, car deux tirages avec remise coïncident rarement.

EXERCICE 62 ⋆⋆

On note $a_i = \mathrm{P}(X = i)$ et $b_j = \mathrm{P}(Y = j)$ pour i, $j \in [\![1\,,n]\!]$, et a, b les colonnes correspondantes. La proposition 15, appliquée aux valeurs $(i, j) \in [\![1\,,n]\!]^2$, donne

$$X \perp\!\!\!\perp Y \iff m_{i,j} = a_i b_j \text{ pour tous } i, j \iff \boxed{M = a\,b^T}$$

On peut remplacer $X(\Omega) \times Y(\Omega)$ par $[\![1\,,n]\!]^2$: si $a_i = 0$, alors $0 \leqslant m_{i,j} \leqslant a_i$ donne $m_{i,j} = 0 = a_i b_j$, et de même si $b_j = 0$.

Condition nécessaire. Si X et Y sont indépendantes, $M = ab^T$, et les colonnes a et b sont non nulles, puisque leurs coefficients ont pour somme 1. Par l'exercice 43 de la feuille du chapitre 16, M est de rang 1.

Condition suffisante. Si M est de rang 1, l'exercice 43 fournit deux colonnes non nulles U et V telles que $M = UV^T$, soit $m_{i,j} = u_i v_j$. On note $\sigma_U = \sum_i u_i$ et $\sigma_V = \sum_j v_j$. Par la proposition 14, les lois marginales sont les sommes des lignes et des colonnes de M :

$$a_i = \sum_{j=1}^{n} u_i v_j = \sigma_V\, u_i, \qquad b_j = \sigma_U\, v_j, \qquad 1 = \sum_{i,j} m_{i,j} = \sigma_U \sigma_V.$$

Donc $a_i b_j = \sigma_U \sigma_V\, u_i v_j = m_{i,j}$ pour tous i, j, et X et Y sont indépendantes.

$$\boxed{\mathrm{rg}(M) = 1 \iff X \perp\!\!\!\perp Y}$$

Le rang de M n'est jamais nul, puisque la somme de ses coefficients vaut 1 : la dépendance se traduit donc par un rang au moins égal à 2. Dans l'exemple de la section 21.5.4, la loi conjointe du tirage avec remise est la matrice dont tous les coefficients valent $\dfrac{1}{9}$, de rang 1, et celle du tirage sans remise, de coefficients 0 sur la diagonale et $\dfrac{1}{6}$ ailleurs, est de rang 3 ✓.

Piège. Les colonnes U et V fournies par l'exercice 43 ne sont pas les lois marginales : elles ne sont définies qu'à un facteur près, puisque $UV^T = (\lambda U)(\lambda^{-1}V)^T$, et peuvent avoir des coefficients négatifs. C'est la normalisation $\sigma_U \sigma_V = 1$, imposée par la somme des probabilités, qui permet de revenir aux marges.

EXERCICE 63 ⋆⋆

Modèle. On munit $\Omega = [\![-n\,,n]\!]^4$ de la probabilité uniforme, chaque quadruplet (a, b, c, d) ayant la probabilité $\dfrac{1}{(2n+1)^4}$; c'est le modèle de la proposition 17 pour quatre tirages indépendants uniformes. La matrice n'est pas inversible si et seulement si son déterminant $ad - bc$ est nul (chapitre 17).

Majoration. L'événement $(ad = bc)$ est contenu dans la réunion des événements $(a = 0)$ et $(a \neq 0,\ ad = bc)$. Le premier a la probabilité $\dfrac{(2n+1)^3}{(2n+1)^4} = \dfrac{1}{2n+1}$. Pour le second, on fixe (a, b, c) avec $a \neq 0$: l'égalité $ad = bc$ impose $d = \dfrac{bc}{a}$, et il y a au plus un d convenable. Il y a $2n(2n+1)^2$ triplets (a, b, c) avec $a \neq 0$, donc l'événement contient au plus $2n(2n+1)^2$ issues, et

$$\mathrm{P}(a \neq 0,\ ad = bc) \leqslant \frac{2n(2n+1)^2}{(2n+1)^4} = \frac{2n}{(2n+1)^2} \leqslant \frac{1}{2n+1}.$$

Par la sous-additivité (proposition 7 (iv)),

$$\boxed{1 - p(n) = \mathrm{P}(ad - bc = 0) \leqslant \frac{2}{2n+1}} \qquad \text{d'où} \qquad 1 - \frac{2}{2n+1} \leqslant p(n) \leqslant 1, \qquad \boxed{\lim_{n\to+\infty} p(n) = 1}$$

par encadrement. Les valeurs exactes, obtenues par énumération, confirment la majoration : $1 - p(n)$ vaut $\dfrac{11}{27} \approx 0{,}407$ pour $n = 1$ (majorant $\dfrac{2}{3}$), $\dfrac{129}{625} \approx 0{,}206$ pour $n = 2$ (majorant 0,4) et $\dfrac{833}{14\,641} \approx 0{,}057$ pour $n = 5$ (majorant $\dfrac{2}{11} \approx 0{,}18$).

À retenir. Une matrice à coefficients entiers pris au hasard dans un grand intervalle est presque sûrement inversible : être non inversible impose une relation entre les coefficients, qui détermine l'un d'eux en fonction des autres. La distinction selon a sert à pouvoir diviser par a ; sans elle, l'équation $0 \cdot d = bc$ laisserait d libre, et la majoration par « au plus une valeur de d » tomberait.

EXERCICE 64 ⋆⋆

1. Supposons $ab = cd = \lambda$. Comme $(ac - bd)^2 \geqslant 0$, on a

$$(ac + bd)^2 = (ac - bd)^2 + 4abcd \geqslant 4(ab)(cd) = 4\lambda^2, \qquad \text{d'où} \qquad \boxed{ac + bd \geqslant 2\lambda > \lambda}$$

puisque $ac + bd > 0$ et $\lambda > 0$. Les trois conditions sont donc incompatibles.

2. Supposons que X et Y, indépendantes et à valeurs dans $[\![0\,,n]\!]$, aient une somme $S = X + Y$ de loi uniforme sur $[\![0\,,2n]\!]$, et notons $\lambda = \dfrac{1}{2n+1}$. On pose

$$a = \mathrm{P}(X = 0), \qquad d = \mathrm{P}(X = n), \qquad b = \mathrm{P}(Y = 0), \qquad c = \mathrm{P}(Y = n).$$

— La somme vaut 0 si et seulement si $X = Y = 0$, d'où $\lambda = \mathrm{P}(S = 0) = ab$ par indépendance.

— La somme vaut $2n$ si et seulement si $X = Y = n$, d'où $\lambda = \mathrm{P}(S = 2n) = dc$.

— Les événements $(X = 0, Y = n)$ et $(X = n, Y = 0)$ sont incompatibles, puisque $n \geqslant 1$, et contenus dans $(S = n)$: $\lambda = \mathrm{P}(S = n) \geqslant ac + db$.

Les deux premières égalités montrent que a, b, c, d sont strictement positifs, et les trois conditions contredisent la question 1.

$$\boxed{\text{il n'existe pas de variables indépendantes à valeurs dans } [\![0\,,n]\!] \text{ de somme uniforme sur } [\![0\,,2n]\!]}$$

3. Si deux dés pipés donnaient des résultats D_1 et D_2, indépendants, de somme uniforme sur $[\![2\,,12]\!]$, les variables $X = D_1 - 1$ et $Y = D_2 - 1$ seraient indépendantes (proposition 16), à valeurs dans $[\![0\,,5]\!]$, et leur somme $D_1 + D_2 - 2$ suivrait la loi uniforme sur $[\![0\,,10]\!]$, ce que la question 2, avec $n = 5$, interdit.

$$\boxed{\text{on ne peut pas piper deux dés pour que la somme suive la loi uniforme sur } [\![2\,,12]\!]}$$

Le résultat vaut que les deux dés soient pipés de la même façon ou non. Une recherche numérique sur 20 000 couples de lois tirés au hasard le confirme : pour chacun, l'une des onze probabilités de la somme s'écarte de $\dfrac{1}{11}$ de plus de 0,05 ✓.

À retenir. L'argument n'utilise que les trois valeurs extrêmes et centrale : les valeurs 0 et $2n$ ne s'obtiennent que d'une façon, ce qui force ab et cd à être grands, et rend alors trop probable la valeur n. Il dépend de ce que X et Y ont le même intervalle de valeurs : si X suit la loi uniforme sur $\{0, 1\}$ et Y la loi uniforme sur $\{0, 2\}$, indépendantes, la somme suit la loi uniforme sur $[\![0\,,3]\!]$.

EXERCICE 65 ⋆⋆

Modèle. Après k tirages, l'urne contient $k + 2$ boules, dont N_k blanches. L'énoncé se traduit par $N_0 = 1$ et, pour $j \in [\![1\,,k+1]\!]$ tel que $\mathrm{P}(N_k = j) > 0$,

$$\mathrm{P}_{(N_k=j)}(N_{k+1} = j + 1) = \frac{j}{k+2}, \qquad \mathrm{P}_{(N_k=j)}(N_{k+1} = j) = \frac{k+2-j}{k+2},$$

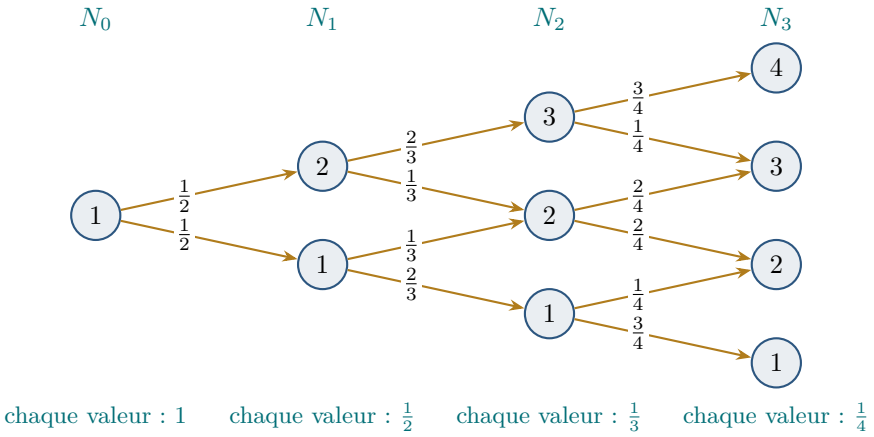
selon que la boule tirée est blanche ou noire.

Récurrence. Pour $k = 0$, $N_0 = 1$ suit la loi uniforme sur $[\![1\,,1]\!] = \{1\}$. Supposons que N_k suive la loi uniforme sur $[\![1\,,k+1]\!]$. Comme $N_{k+1} \in \{N_k, N_k + 1\}$, la variable N_{k+1} prend ses valeurs dans $[\![1\,,k+2]\!]$. Soit $j \in [\![1\,,k+2]\!]$. La formule des probabilités totales, appliquée au système complet $((N_k = i))_{1 \leqslant i \leqslant k+1}$, ne fait intervenir que $i = j - 1$ et $i = j$:

$$\mathrm{P}(N_{k+1} = j) = \frac{j-1}{k+2} \cdot \frac{1}{k+1} + \frac{k+2-j}{k+2} \cdot \frac{1}{k+1} = \frac{k+1}{(k+2)(k+1)}, \qquad \boxed{\mathrm{P}(N_{k+1} = j) = \frac{1}{k+2}}$$

Aux bords, la formule reste exacte : pour $j = 1$, le terme $i = 0$ n'existe pas, mais son coefficient $\dfrac{j-1}{k+2}$ est nul ; pour $j = k+2$, le terme $i = k+1$ n'existe pas, et son coefficient $\dfrac{k+2-j}{k+2}$ est nul. La récurrence est établie :

$$\boxed{N_k \sim \mathcal{U}(\llbracket 1\,,k+1\rrbracket) \quad \text{pour tout } k \in \mathbb{N}}$$



Chaque colonne représente les valeurs possibles de N_k, et chaque flèche ocre porte la probabilité de transition ; on lit la récurrence sur un nœud, par exemple $N_3 = 2$, atteint depuis $N_2 = 1$ avec la probabilité $\dfrac{1}{4}$ et depuis $N_2 = 2$ avec la probabilité $\dfrac{2}{4}$, d'où $\dfrac{1}{3}\left(\dfrac{1}{4} + \dfrac{2}{4}\right) = \dfrac{1}{4}$.

À retenir. Le résultat contredit l'intuition : on s'attendrait à ce que la proportion de boules blanches se stabilise autour de $\dfrac{1}{2}$, puisque l'urne part équilibrée. Au contraire, la proportion $\dfrac{N_k}{k+2}$ suit la loi uniforme sur l'ensemble des fractions $\dfrac{j}{k+2}$, pour $j \in \llbracket 1\,,k+1\rrbracket$: les premiers tirages, amplifiés par les ajouts, décident d'une proportion qui peut prendre presque n'importe quelle valeur. Après n tirages tous blancs, l'urne contient $n+1$ blanches parmi $n+2$ boules, et le tirage suivant est blanc avec la probabilité $\dfrac{n+1}{n+2}$: c'est exactement la limite trouvée dans l'exercice 53 pour la loi de succession de Laplace. Les deux expériences sont en effet très proches : dans l'urne de Pólya comme dans le choix d'une urne au hasard, la proportion de boules blanches est inconnue et uniforme, et chaque tirage en informe.

EXERCICE 66 ⋆⋆

Comme dans le troisième exemple de la section 21.5.7, Σ est une variable aléatoire de loi uniforme sur $\mathfrak{S}_n$: chaque permutation a la probabilité $\dfrac{1}{n!}$, et calculer une probabilité revient à dénombrer des permutations.

1.(a) L'application $\sigma \mapsto \sigma \circ (1\ 2)$ est une bijection de $\mathfrak{S}_n$ sur lui-même, qui échange les valeurs $\sigma(1)$ et $\sigma(2)$; elle échange donc les ensembles $\{\sigma \mid \sigma(1) < \sigma(2)\}$ et $\{\sigma \mid \sigma(1) > \sigma(2)\}$, qui ont le même cardinal et forment une partition de $\mathfrak{S}_n$, puisque $\sigma(1) \neq \sigma(2)$.

$$\boxed{\mathrm{P}(\Sigma(1) < \Sigma(2)) = \frac{1}{2}}$$

(b) L'exemple du cours donne $\mathrm{P}(\Sigma(1) = i) = \dfrac{1}{n}$: $\Sigma(1)$ et $\Sigma(2)$ suivent la loi uniforme sur $\llbracket 1\,,n\rrbracket$. Pour $i \neq j$, les permutations telles que $\sigma(1) = i$ et $\sigma(2) = j$ sont en bijection avec les bijections de $\llbracket 3\,,n\rrbracket$ sur $\llbracket 1\,,n\rrbracket \setminus \{i,j\}$, qui sont au nombre de $(n-2)!$ (théorème 1) ; pour $i = j$, il n'y en a aucune, par injectivité. Donc

$$\boxed{\mathrm{P}(\Sigma(1) = i, \Sigma(2) = j) = \begin{cases} \dfrac{1}{n(n-1)} & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases}}$$

Comme $\mathrm{P}(\Sigma(1)=1,\Sigma(2)=1)=0\neq\dfrac{1}{n^2}=\mathrm{P}(\Sigma(1)=1)\mathrm{P}(\Sigma(2)=1)$, les variables $\Sigma(1)$ et $\Sigma(2)$ ne sont pas indépendantes. C'est la loi conjointe de deux tirages sans remise dans une urne de n boules numérotées, dont le cas $n=3$ figure dans la section 21.5.4.

(c) Les couples (i,j) d'entiers consécutifs de $[\![1\,,n]\!]$, c'est-à-dire tels que $|i-j|=1$, sont au nombre de $2(n-1)$: $(i,i+1)$ et $(i+1,i)$ pour $i\in[\![1\,,n-1]\!]$. Par la question précédente,

$$\boxed{\mathrm{P}\big(|\Sigma(1)-\Sigma(2)|=1\big)=\frac{2(n-1)}{n(n-1)}=\frac{2}{n}}$$

Pour $n=2$, on trouve 1, ce qui est attendu, puisque $\{\Sigma(1),\Sigma(2)\}=\{1,2\}$.

2.(a) Rappelons que les orbites de σ sont les classes de la relation « $j=\sigma^p(i)$ pour un $p\in\mathbb{Z}$ » (chapitre 17, démonstration du théorème 1), et que σ agit sur une orbite à $k\geqslant 2$ éléments comme un k-cycle. La partie C est une orbite de σ si et seulement si $\sigma(C)\subset C$, ce qui entraîne $\sigma(C)=C$ par injectivité et $\sigma(E\setminus C)=E\setminus C$, où $E=[\![1\,,n]\!]$, et si la restriction $\sigma_{|C}$, qui est une permutation de C, a pour seule orbite C.

L'application $\sigma\mapsto(\sigma_{|C},\sigma_{|E\setminus C})$ est une bijection de l'ensemble des permutations qui stabilisent C sur $\mathfrak{S}_C\times\mathfrak{S}_{E\setminus C}$. Il reste à dénombrer les permutations de C ayant une seule orbite. Soit c le plus petit élément de C ; une telle permutation τ est déterminée par la suite $\big(\tau(c),\tau^2(c),\dots,\tau^{k-1}(c)\big)$, qui est une liste ordonnée des $k-1$ éléments de $C\setminus\{c\}$, et toute liste de ce type provient d'une unique telle permutation, le cycle $\big(c\ \tau(c)\ \cdots\ \tau^{k-1}(c)\big)$. Il y en a donc $(k-1)!$, y compris pour $k=1$, où seule l'identité de C convient. Ainsi C est une orbite de $(k-1)!\,(n-k)!$ permutations :

$$\boxed{\mathrm{P}(C\text{ est une orbite de }\Sigma)=\frac{(k-1)!\,(n-k)!}{n!}=\frac{1}{k\binom{n}{k}}}$$

Pour $k=n$, on obtient la probabilité $\dfrac{1}{n}$ que Σ soit un n-cycle ; pour $k=1$, la probabilité $\dfrac{1}{n}$ que l'élément de C soit un point fixe.

(b) Les entiers 1 et 2 sont dans la même orbite si et seulement si celle-ci est l'une des parties C contenant 1 et 2 ; pour une permutation donnée, une seule partie convient, donc les événements « C est une orbite », pour ces parties C, sont deux à deux incompatibles. Il y a $\dbinom{n-2}{k-2}$ parties à k éléments contenant 1 et 2, pour $k\in[\![2\,,n]\!]$, et

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

En sommant sur k, avec $\displaystyle\sum_{k=2}^{n}(k-1)=\frac{n(n-1)}{2}$,

$$\boxed{\mathrm{P}(1\text{ et }2\text{ appartiennent à la même orbite de }\Sigma)=\frac{1}{2}}$$

(c) Soit L le nombre d'éléments de l'orbite de Σ qui contient 1, et $k\in[\![1\,,n]\!]$. L'événement $(L=k)$ est la réunion disjointe, sur les $\dbinom{n-1}{k-1}$ parties C à k éléments contenant 1, des événements « C est une orbite de Σ » :

$$\mathrm{P}(L=k)=\binom{n-1}{k-1}\frac{(k-1)!\,(n-k)!}{n!}=\frac{(n-1)!}{n!},\qquad\boxed{L\sim\mathcal{U}([\![1\,,n]\!])}$$

Le calcul a été contrôlé par énumération de $\mathfrak{S}_n$ pour $n\leqslant 6$ ✓.

À retenir. Ces résultats surprennent : on s'attendrait à ce que les petites orbites soient plus probables, alors que l'orbite de 1 a une longueur uniforme, et qu'il y a une chance sur deux que 1 et 2 soient dans la même orbite, quel que soit n. Le cas $k=n$, où Σ est un n-cycle avec la probabilité $\dfrac{1}{n}$, sera retrouvé dans l'exercice 69, qui utilise aussi les parties stables, et la question 2(a) prolonge l'exercice 17, où les parties stables par une permutation étaient les réunions d'orbites.

EXERCICE 67 ⋆⋆⋆

Les variables $X_1, \dots, X_n$ sont données par l'énoncé ; on n'a pas besoin de décrire l'univers (section 21.5.6). On note $q = 1 - p$.

1. Soit $k \in [\![1\,, n-1]\!]$. Comme $\Pi_{k+1} = \Pi_k X_{k+1}$ et que ces deux variables sont à valeurs dans $\{-1, 1\}$, le produit vaut 1 si et seulement si les deux facteurs sont égaux :

$$(\Pi_{k+1} = 1) = (\Pi_k = 1,\ X_{k+1} = 1) \cup (\Pi_k = -1,\ X_{k+1} = -1),$$

réunion de deux événements incompatibles. La variable Π_k est une fonction de $(X_1, \dots, X_k)$, donc elle est indépendante de X_{k+1} par le lemme des coalitions (théorème 7). Par la proposition 7 (i), puis l'indépendance,

$$\boxed{u_{k+1} = p\,u_k + q\,v_k} \qquad \text{et, de même,} \qquad \boxed{v_{k+1} = q\,u_k + p\,v_k}$$

la seconde formule venant de $(\Pi_{k+1} = -1) = (\Pi_k = 1, X_{k+1} = -1) \cup (\Pi_k = -1, X_{k+1} = 1)$.

2. Comme Π_k prend ses valeurs dans $\{-1, 1\}$, les événements $(\Pi_k = 1)$ et $(\Pi_k = -1)$ forment un système complet (proposition 11), et $u_k + v_k = 1$. En soustrayant les deux relations de la question 1, on obtient

$$u_{k+1} - v_{k+1} = (p - q)(u_k - v_k) = (2p - 1)(u_k - v_k).$$

La suite $(u_k - v_k)$ est géométrique de raison $2p - 1$, de premier terme $u_1 - v_1 = p - q = 2p - 1$, donc $u_k - v_k = (2p - 1)^k$. Avec $u_k + v_k = 1$,

$$\boxed{u_k = \frac{1 + (2p-1)^k}{2}, \qquad v_k = \frac{1 - (2p-1)^k}{2}}$$

Comme $p \in]0\,, 1[$, on a $|2p - 1| < 1$, et $(2p - 1)^k$ tend vers 0 : lorsque n, puis k, deviennent grands, $\boxed{u_k \text{ et } v_k \text{ tendent vers } \frac{1}{2}}$. Le signe d'un long produit de signes aléatoires est presque équiprobable, quel que soit p ; la convergence est géométrique, d'autant plus lente que p est proche de 0 ou de 1. Pour p proche de 0, le signe change très probablement d'un rang au suivant, et u_k oscille autour de $\dfrac{1}{2}$, puisque $2p - 1 < 0$, avant de s'en approcher.

On pouvait aussi remarquer que $u_{k+1} = (2p - 1)u_k + q$: la suite (u_k) est arithmético-géométrique (chapitre 7), de point fixe $\dfrac{1}{2}$. Matriciellement, la colonne $(u_k, v_k)^T$ est multipliée à chaque étape par la matrice $\begin{pmatrix} p & q \\ q & p \end{pmatrix}$, comme dans l'exemple de la chaîne de Markov à deux états de la section 21.5.7.

3. On suppose $n \geqslant 2$. Par la proposition 15, $\Pi_1 = X_1$ et $\Pi_2 = X_1 X_2$ sont indépendantes si et seulement si $\mathrm{P}(\Pi_1 = a, \Pi_2 = b) = \mathrm{P}(\Pi_1 = a)\mathrm{P}(\Pi_2 = b)$ pour tous a, $b \in \{-1, 1\}$.

Condition nécessaire. Pour $a = b = 1$, on a $(\Pi_1 = 1, \Pi_2 = 1) = (X_1 = 1, X_2 = 1)$, de probabilité p^2, tandis que $\mathrm{P}(\Pi_1 = 1)\mathrm{P}(\Pi_2 = 1) = p\,u_2 = p(2p^2 - 2p + 1)$. Comme $p \neq 0$, l'égalité équivaut à $p = 2p^2 - 2p + 1$, c'est-à-dire à $2p^2 - 3p + 1 = (2p - 1)(p - 1) = 0$, soit $p = \dfrac{1}{2}$, puisque $p \neq 1$.

Condition suffisante. Si $p = \dfrac{1}{2}$, pour tous a, $b \in \{-1, 1\}$, l'événement $(\Pi_1 = a, \Pi_2 = b)$ est égal à $(X_1 = a, X_2 = ab)$, de probabilité $\dfrac{1}{4}$ par indépendance, et le produit $\mathrm{P}(\Pi_1 = a)\mathrm{P}(\Pi_2 = b)$ vaut aussi $\dfrac{1}{2} \times \dfrac{1}{2}$, puisque $u_2 = v_2 = \dfrac{1}{2}$. D'où

$$\boxed{\Pi_1 \perp\!\!\!\perp \Pi_2 \iff p = \tfrac{1}{2}}$$

4. On suppose $p = \dfrac{1}{2}$. Soit φ l'application de $\{-1, 1\}^n$ dans lui-même définie par

$$\varphi(x_1, \dots, x_n) = (x_1, x_1x_2, \dots, x_1x_2 \cdots x_n).$$

C'est une bijection : si $\pi = \varphi(x)$, on retrouve $x_1 = \pi_1$ et $x_k = \pi_k \pi_{k-1}$ pour $k \geqslant 2$, puisque $\pi_{k-1}^2 = 1$. Par indépendance, $\mathrm{P}(X_1 = x_1, \dots, X_n = x_n) = \dfrac{1}{2^n}$ pour tout x, et, pour tout $\pi \in \{-1,1\}^n$,

$$\mathrm{P}(\Pi_1 = \pi_1, \dots, \Pi_n = \pi_n) = \mathrm{P}\big((X_1, \dots, X_n) = \varphi^{-1}(\pi)\big) = \frac{1}{2^n} = \prod_{k=1}^{n} \mathrm{P}(\Pi_k = \pi_k),$$

puisque $u_k = v_k = \dfrac{1}{2}$ pour $p = \dfrac{1}{2}$. Par la proposition 15,

$$\boxed{\Pi_1, \dots, \Pi_n \text{ sont mutuellement indépendantes}}$$

Ce résultat a été contrôlé par énumération pour $n = 4$, avec $p = \dfrac{1}{2}$, et son échec pour $p = \dfrac{1}{3}$ et $p = \dfrac{3}{4}$ ✓.

Piège. Le lemme des coalitions ne s'applique pas ici : toutes les variables Π_k dépendent de X_1, et les paquets ne sont pas disjoints. L'erreur la plus répandue est de croire que des fonctions de paquets qui se chevauchent sont « un peu » indépendantes ; la question 3 montre qu'elles ne le sont pas en général, et l'indépendance obtenue pour $p = \dfrac{1}{2}$ tient à la bijection φ, qui transporte la loi uniforme de $\{-1,1\}^n$ sur elle-même.

EXERCICE 68 ⋆⋆⋆

1. Tant que le fumeur n'a pas trouvé de boîte vide, chaque choix de poche lui fournit une allumette. Les deux boîtes en contiennent $2n$ au total : si les $2n$ premiers choix ont tous fourni une allumette, les deux boîtes sont vides, et le choix suivant tombe sur une boîte vide. La constatation a donc lieu $\boxed{\text{au plus tard au } (2n+1)\text{-ième choix}}$.

On prend pour univers $\Omega = \{G, D\}^{2n+1}$, muni de la probabilité uniforme : l'issue ω, de coordonnées ω_1, …, ω_{2n+1}, donne les poches choisies, à gauche ou à droite, lors des choix successifs. Par la proposition 17, les coordonnées sont mutuellement indépendantes et de loi uniforme sur $\{G, D\}$, comme le demande l'énoncé. Les choix postérieurs à la constatation ne sont pas effectués ; on les ajoute, sans qu'ils influent sur rien, pour que toutes les issues aient la même longueur.

2. Soit $k \in [\![0, n]\!]$, et G_k l'événement « la boîte de gauche est trouvée vide alors qu'il reste k allumettes dans celle de droite ». Montrons que

$$G_k = \{\omega \in \Omega \mid \text{exactement } n \text{ des } 2n-k \text{ premiers choix sont } G, \text{ et } \omega_{2n-k+1} = G\}.$$

Si ω appartient à G_k, la boîte de gauche a été vidée par n choix G, celle de droite a fourni $n-k$ allumettes, et la constatation a lieu au choix suivant, qui est un choix G : c'est le choix numéro $n + (n-k) + 1 = 2n - k + 1$, et les $2n-k$ précédents comportent n choix G. Réciproquement, si ω est dans l'ensemble de droite, aucune constatation n'a lieu avant le choix $2n-k+1$: elle exigerait $n+1$ choix d'une même poche parmi les $2n-k$ premiers, alors qu'il y a n choix G et $n-k \leqslant n$ choix D. Au choix $2n-k+1$, la boîte de gauche a fourni ses n allumettes, et il en reste $n-(n-k) = k$ à droite.

On dénombre cet ensemble : on choisit les positions des n choix G parmi les $2n-k$ premiers, de $\binom{2n-k}{n}$ façons (théorème 2), le choix $2n-k+1$ est imposé, et les k derniers sont quelconques. Ainsi

$$\mathrm{P}(G_k) = \binom{2n-k}{n} \frac{2^k}{2^{2n+1}} = \binom{2n-k}{n} \frac{1}{2^{2n-k+1}}.$$

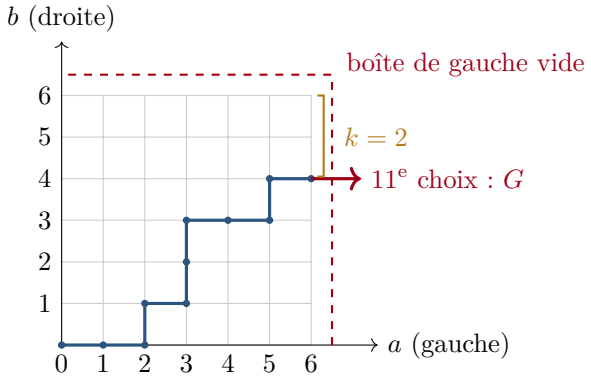
L'échange de G et de D dans toutes les coordonnées est une bijection de Ω qui transforme G_k en l'événement analogue D_k pour la poche droite, donc $\mathrm{P}(D_k) = \mathrm{P}(G_k)$. Les événements G_k et D_k sont incompatibles, puisque la constatation n'a lieu qu'une fois, et $(X = k) = G_k \cup D_k$:

$$\boxed{\mathrm{P}(X = k) = \binom{2n-k}{n} \frac{1}{2^{2n-k}}, \qquad k \in [\![0, n]\!]}$$

Comme ces probabilités ont pour somme 1 (proposition 11), on obtient au passage l'identité

$$\sum_{k=0}^{n} \binom{2n-k}{n} 2^k = 4^n,$$

qu'il ne serait pas immédiat d'établir par le calcul.



La figure suit, pour $n = 6$ et $k = 2$, l'état (a, b) des nombres d'allumettes prises à gauche et à droite : un choix G est un pas vers la droite, un choix D un pas vers le haut. Le chemin en bleu est l'un des $\binom{10}{6}$ chemins de $(0, 0)$ à $(6, 4)$; le onzième choix, en rouge, tente de franchir la barrière $a = 6$: la boîte de gauche est trouvée vide, et il reste $k = 2$ allumettes à droite.

3. Pour $n = 6$, les probabilités $\mathrm{P}(X = k) = \binom{12-k}{6} 2^{k-12}$, pour $k = 0, \ldots, 6$, valent

$$\frac{231}{1024}, \frac{231}{1024}, \frac{105}{512}, \frac{21}{128}, \frac{7}{64}, \frac{7}{128}, \frac{1}{64},$$

soit environ 0,226, 0,226, 0,205, 0,164, 0,109, 0,055 et 0,016. Le nombre moyen d'allumettes restantes est

$$\sum_{k=0}^{6} k\,\mathrm{P}(X = k) = \frac{231 + 2 \times 210 + 3 \times 168 + 4 \times 112 + 5 \times 56 + 6 \times 16}{1024},$$

soit

$$\boxed{\sum_{k=0}^{6} k\,\mathrm{P}(X = k) = \frac{1979}{1024} \approx 1{,}93}$$

Il reste en moyenne près de deux allumettes dans l'autre boîte, et les valeurs 0 et 1 sont les plus probables, à égalité. Plus généralement, on peut montrer que cette moyenne vaut $\dfrac{2n+1}{4^n}\binom{2n}{n} - 1$; ce résultat est admis ici, il a été vérifié par énumération exhaustive pour $n \leqslant 7$ ✓. Avec l'équivalent $\binom{2n}{n} \sim \dfrac{4^n}{\sqrt{\pi n}}$ que donnent les intégrales de Wallis du chapitre 18, elle est équivalente à $2\sqrt{\dfrac{n}{\pi}}$: elle croît comme la racine carrée de n.

Piège. L'erreur la plus répandue est de compter les suites de $2n - k$ choix comportant n choix G sans imposer que le choix suivant soit G : on oublie que la boîte doit être *trouvée* vide, et non simplement vidée, et l'on obtient une formule fausse, dont la somme ne vaut pas 1.

EXERCICE 69 ⋆⋆⋆

1. Le sous-groupe engendré par σ est le plus petit sous-groupe de $\mathfrak{S}_n$ qui contient σ (chapitre 11, remarque qui suit la proposition 8). C'est $H = \{\sigma^k \mid k \in \mathbb{Z}\}$: cet ensemble contient σ, il est un sous-groupe, puisque $\mathrm{id} = \sigma^0$ et $\sigma^a(\sigma^b)^{-1} = \sigma^{a-b}$, et tout sous-groupe qui contient σ contient ses puissances. Pour $x \in [\![1\,,n]\!]$, l'ensemble $\{h(x) \mid h \in H\} = \{\sigma^k(x) \mid k \in \mathbb{Z}\}$ est donc l'orbite de x sous σ au sens du chapitre 17 (démonstration du théorème 1). Le sous-groupe H est transitif si et seulement si l'orbite de 1 est $[\![1\,,n]\!]$ tout entier, c'est-à-dire si σ n'a qu'une orbite, sur laquelle elle agit comme un cycle :

$$\boxed{\langle\sigma\rangle \text{ est transitif} \iff \sigma \text{ est un cycle de longueur } n}$$

2. Un cycle de longueur n s'écrit de façon unique $(1\ a_2\ \cdots\ a_n)$ en le commençant par 1, avec $a_k = \sigma^{k-1}(1)$; la liste $(a_2, \dots, a_n)$ est une permutation de $[\![2\,,n]\!]$, et toute telle liste définit un cycle de longueur n. Il y a donc $(n-1)!$ cycles de longueur n (théorème 1), et, la loi de X_n étant uniforme,

$$\boxed{p_n = \frac{(n-1)!}{n!} = \frac{1}{n}}$$

Une permutation aléatoire n'engendre presque jamais un sous-groupe transitif quand n est grand.

3. (a) Notons H le sous-groupe engendré par X_n et Y_n, pour une issue fixée.

Supposons H non transitif : il existe x, y tels que $h(x) \neq y$ pour tout $h \in H$. Soit O l'ensemble des $h(x)$, pour $h \in H$; il contient $x = \mathrm{id}(x)$ et pas y. Pour $h \in H$, $X_n h \in H$, donc $X_n(O) \subset O$; comme X_n est injective et O fini, $X_n(O)$ a autant d'éléments que O, et $X_n(O) = O$. De même $Y_n(O) = O$. Le complémentaire $O' = [\![1\,,n]\!] \setminus O$ vérifie aussi $X_n(O') = O'$ et $Y_n(O') = O'$, puisque X_n et Y_n sont des bijections de $[\![1\,,n]\!]$. Les parties O et O' sont non vides, l'une contenant x et l'autre y, et leurs cardinaux ont pour somme n : l'une d'elles, que l'on note A, vérifie $1 \leqslant |A| \leqslant \dfrac{n}{2}$.

Réciproquement, soit A une telle partie. L'ensemble $K = \{g \in \mathfrak{S}_n \mid g(A) = A\}$ est un sous-groupe de $\mathfrak{S}_n$: il contient id, et si $g(A) = h(A) = A$, alors $h^{-1}(A) = A$ et $gh^{-1}(A) = A$. Il contient X_n et Y_n, donc il contient H. Soit $x \in A$ et $y \notin A$, qui existe puisque $|A| \leqslant \dfrac{n}{2} < n$: pour tout $h \in H$, $h(x) \in A$, donc $h(x) \neq y$, et H n'est pas transitif.

$$\boxed{H \text{ non transitif} \iff \exists A \subset [\![1\,,n]\!],\ 1 \leqslant |A| \leqslant \tfrac{n}{2},\ X_n(A) = A \text{ et } Y_n(A) = A}$$

(b) Une permutation σ vérifie $\sigma(A) = A$ si et seulement si elle induit une permutation de A et une permutation de $[\![1\,,n]\!] \setminus A$; elle est déterminée par ce couple, et tout couple convient. Il y a donc $k!\,(n-k)!$ telles permutations, et

$$\boxed{\mathrm{P}(X_n(A) = A) = \frac{k!\,(n-k)!}{n!} = \binom{n}{k}^{-1}}$$

(c) Par indépendance de X_n et Y_n (définition de la section 21.5.5, avec la partie $\{\sigma \mid \sigma(A) = A\}$ de $\mathfrak{S}_n$), $\mathrm{P}(X_n(A) = A, Y_n(A) = A) = \binom{n}{k}^{-2}$ pour $|A| = k$. Par la question (a), l'événement « H n'est pas transitif » est la réunion de ces événements, sur les parties A telles que $1 \leqslant |A| \leqslant \dfrac{n}{2}$. La sous-additivité (proposition 7 (iv)), en regroupant les $\binom{n}{k}$ parties de cardinal k, donne

$$1 - q_n \leqslant \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k}\binom{n}{k}^{-2}, \qquad \boxed{1 - q_n \leqslant \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k}^{-1}}$$

Soit $n \geqslant 4$. Le rapport $\binom{n}{k+1}/\binom{n}{k} = \dfrac{n-k}{k+1}$ est au moins égal à 1 pour $k \leqslant \dfrac{n-1}{2}$: les coefficients binomiaux croissent jusqu'à l'indice $\lfloor \dfrac{n}{2} \rfloor$, et $\binom{n}{k} \geqslant \binom{n}{2} = \dfrac{n(n-1)}{2}$ pour $2 \leqslant k \leqslant \lfloor \dfrac{n}{2} \rfloor$. Il y a au plus $\dfrac{n}{2}$ tels indices, d'où

$$0 \leqslant 1 - q_n \leqslant \frac{1}{n} + \frac{n}{2} \times \frac{2}{n(n-1)} = \frac{1}{n} + \frac{1}{n-1}, \qquad \text{donc} \qquad \boxed{q_n \xrightarrow[n\to+\infty]{} 1}$$

L'énumération exhaustive des couples donne $q_2 = \dfrac{3}{4}$, $q_3 = \dfrac{13}{18}$, $q_4 = \dfrac{71}{96}$ et $q_5 = \dfrac{461}{600}$, en accord avec la majoration ✓.

À retenir. Une permutation seule engendre un sous-groupe transitif avec une probabilité $\dfrac{1}{n}$ qui tend vers 0 ; deux permutations le font avec une probabilité qui tend vers 1. La méthode est typique : on caractérise le mauvais événement comme une réunion d'événements simples, puis la sous-additivité, grossière mais sans hypothèse, suffit à montrer qu'il est rare.

Espérance

EXERCICE 70 ⋆

1. Par la formule de transfert (théorème 10), puis la somme des carrés du chapitre 3, en posant $j = k - 1$,

$$\mathrm{E}((X-1)^2) = \frac{1}{n}\sum_{k=1}^{n}(k-1)^2 = \frac{1}{n}\sum_{j=0}^{n-1} j^2 = \frac{1}{n}\cdot\frac{(n-1)n(2n-1)}{6}, \qquad \boxed{\mathrm{E}((X-1)^2) = \frac{(n-1)(2n-1)}{6}}$$

De même, avec la somme géométrique de raison $\mathrm{e} \neq 1$,

$$\boxed{\mathrm{E}(\mathrm{e}^X) = \frac{1}{n}\sum_{k=1}^{n}\mathrm{e}^k = \frac{\mathrm{e}\,(\mathrm{e}^n - 1)}{n(\mathrm{e}-1)}}$$

Pour $n = 1$, on retrouve $\mathrm{E}((X-1)^2) = 0$ et $\mathrm{E}(\mathrm{e}^X) = \mathrm{e}$, puisque $X = 1$ ✓.

2. On note $q = 1 - p$. Par la formule de transfert, l'égalité de l'énoncé, puis le changement d'indice $j = k + 1$,

$$\mathrm{E}(Y) = \sum_{k=0}^{n}\frac{1}{k+1}\binom{n}{k}p^k q^{n-k} = \frac{1}{n+1}\sum_{k=0}^{n}\binom{n+1}{k+1}p^k q^{n-k} = \frac{1}{(n+1)p}\sum_{j=1}^{n+1}\binom{n+1}{j}p^j q^{n+1-j}.$$

La dernière somme est celle de la formule du binôme (théorème 3), privée de son terme d'indice $j = 0$, qui vaut q^{n+1} :

$$\boxed{\mathrm{E}\Big(\frac{1}{X+1}\Big) = \frac{1-(1-p)^{n+1}}{(n+1)p}}$$

L'égalité de l'énoncé est la formule du pion (chapitre 3, proposition 9) :

$$\frac{1}{k+1}\binom{n}{k} = \frac{n!}{(k+1)!\,(n-k)!} = \frac{1}{n+1}\binom{n+1}{k+1}.$$

On n'a pas $\mathrm{E}(Y) = \dfrac{1}{\mathrm{E}(X)+1}$ en général. Pour $n = 1$ et $p = \dfrac{1}{2}$, on trouve $\mathrm{E}(Y) = \dfrac{3}{4}$ et $\dfrac{1}{\mathrm{E}(X)+1} = \dfrac{2}{3}$. Plus précisément, posons $m = \mathrm{E}(X) = np$. Pour $x > -1$, une mise au même dénominateur donne

$$\frac{1}{x+1} - \frac{1}{m+1} + \frac{x-m}{(m+1)^2} = \frac{(x-m)^2}{(x+1)(m+1)^2} \geqslant 0 :$$

la fonction $x \mapsto \dfrac{1}{x+1}$, convexe, est au-dessus de sa tangente en m (chapitre 9). En appliquant cette inégalité à X et en prenant l'espérance (théorème 9 (i) et (ii)), le terme $\dfrac{\mathrm{E}(X)-m}{(m+1)^2}$ disparaît, et

$$\boxed{\mathrm{E}\Big(\frac{1}{X+1}\Big) \geqslant \frac{1}{\mathrm{E}(X)+1}}$$

avec égalité si et seulement si $\mathrm{E}\Big(\dfrac{(X-m)^2}{(X+1)(m+1)^2}\Big) = 0$, c'est-à-dire, la variable étant positive, si $\mathrm{P}(X=m)=1$. Pour $p \in]0\,,1[$, la variable X n'est pas constante, puisque $\mathrm{P}(X=0) = q^n$ et $\mathrm{P}(X=n) = p^n$ sont non nuls : l'inégalité est stricte. L'égalité n'a lieu que pour $p=1$, où $X=n$.
Piège. C'est la mise en garde qui suit le théorème 10 : $\mathrm{E}(f(X)) \neq f(\mathrm{E}(X))$ dès que f n'est pas affine. L'erreur la plus répandue est d'écrire $\mathrm{E}(\dfrac{1}{X}) = \dfrac{1}{\mathrm{E}(X)}$; l'inégalité de Jensen du chapitre 9 (proposition 7), appliquée avec les coefficients $\mathrm{P}(X=x)$, donne le sens de l'écart pour une fonction convexe.

EXERCICE 71 ⋆

1. Il existe $c>0$ tel que $\mathrm{P}(X=k) = ck$ pour $k \in [\![1\,,6]\!]$. La somme des probabilités vaut 1, donc $21c=1$:

$$\boxed{\mathrm{P}(X=k) = \frac{k}{21}, \quad k \in [\![1\,,6]\!]}, \qquad \mathrm{E}(X) = \sum_{k=1}^{6} \frac{k^2}{21} = \frac{91}{21}, \qquad \boxed{\mathrm{E}(X) = \frac{13}{3} \approx 4{,}33}$$

Pour un dé équilibré, l'espérance vaut $\dfrac{7}{2} = 3{,}5$ (proposition 19) : le dé pipé favorise les grandes faces, et sa moyenne est plus élevée.

2. La variable $Y = \dfrac{1}{X}$ prend les valeurs $\dfrac{1}{k}$, deux à deux distinctes, et $(Y = \dfrac{1}{k}) = (X=k)$ (proposition 13) :

$$\boxed{\mathrm{P}\Big(Y=\frac{1}{k}\Big) = \frac{k}{21}, \quad k \in [\![1\,,6]\!]}, \qquad \mathrm{E}(Y) = \sum_{k=1}^{6} \frac{1}{k}\cdot\frac{k}{21} = \frac{6}{21}, \qquad \boxed{\mathrm{E}(Y) = \frac{2}{7}}$$

On a $\dfrac{2}{7} \approx 0{,}286$ et $\dfrac{1}{\mathrm{E}(X)} = \dfrac{3}{13} \approx 0{,}231$: $\boxed{\mathrm{E}(\frac{1}{X}) > \frac{1}{\mathrm{E}(X)}}$. C'est le phénomène de l'exercice 70, pour la fonction convexe $x \mapsto \dfrac{1}{x}$ sur $]0\,,+\infty[$: l'identité $\dfrac{1}{x} - \dfrac{1}{m} + \dfrac{x-m}{m^2} = \dfrac{(x-m)^2}{xm^2}$, avec $m = \dfrac{13}{3}$, donne l'inégalité, stricte puisque X n'est pas constante.

EXERCICE 72 ⋆⋆

1. On note $q = 1-p$. Par la formule de transfert, puis la formule du binôme,

$$\mathrm{E}((-1)^X) = \sum_{k=0}^{2n} (-1)^k \binom{2n}{k} p^k q^{2n-k} = (q-p)^{2n}, \qquad \boxed{\mathrm{E}((-1)^X) = (1-2p)^{2n}}$$

Comme $(-1)^X = \mathbf{1}_{(X \text{ pair})} - \mathbf{1}_{(X \text{ impair})}$, on en déduit $\mathrm{P}(X \text{ pair}) = \dfrac{1+(1-2p)^{2n}}{2}$, toujours au moins égal à $\dfrac{1}{2}$, puisque l'exposant $2n$ est pair.

2. Si X est pair, $Y = \dfrac{X}{2}$; si X est impair, $Y = \dfrac{X-1}{2}$. Dans les deux cas, $Y = \dfrac{1}{2}(X - \dfrac{1-(-1)^X}{2})$, puisque $\dfrac{1-(-1)^X}{2}$ vaut 0 si X est pair et 1 sinon :

$$\boxed{Y = \frac{X}{2} - \frac{1}{4} + \frac{(-1)^X}{4}}$$

Par linéarité de l'espérance (théorème 9 (i)), avec $\mathrm{E}(X) = 2np$ (proposition 19) et la question 1,

$$\boxed{\mathrm{E}(Y) = np - \frac{1-(1-2p)^{2n}}{4}}$$

Pour $p = \dfrac{1}{2}$, on trouve $\mathrm{E}(Y) = \dfrac{n}{2} - \dfrac{1}{4}$. Ces formules ont été vérifiées sur la loi exacte pour $n \leqslant 4$ et trois valeurs de p ✓.

Piège. L'erreur la plus répandue est d'écrire $\mathrm{E}(\lfloor \frac{X}{2} \rfloor) = \lfloor \frac{\mathrm{E}(X)}{2} \rfloor = \lfloor np \rfloor$: la partie entière n'est pas linéaire. L'astuce consiste à exprimer la partie entière comme une fonction affine de X et de $(-1)^X$, dont on sait calculer les espérances.

EXERCICE 73 ⋆⋆

On prend pour univers l'ensemble des parties à r éléments de l'ensemble des $2n$ animaux, muni de la probabilité uniforme ; il a $\binom{2n}{r}$ éléments (théorème 2). On applique la méthode des indicatrices (section 21.6.3). Pour $i \in [\![1\,,n]\!]$, soit A_i l'événement « les deux animaux de l'espèce i sont choisis ». Alors $N = \sum_{i=1}^{n} \mathbf{1}_{A_i}$. Une partie favorable à A_i est formée des deux animaux de l'espèce i et de $r-2$ animaux choisis parmi les $2n-2$ autres :

$$\mathrm{P}(A_i) = \frac{\binom{2n-2}{r-2}}{\binom{2n}{r}} = \frac{(2n-2)!}{(r-2)!\,(2n-r)!} \cdot \frac{r!\,(2n-r)!}{(2n)!} = \frac{r(r-1)}{2n(2n-1)}.$$

Par linéarité (théorème 9 (i) et (iv)),

$$\boxed{\mathrm{E}(N) = n\,\frac{r(r-1)}{2n(2n-1)} = \frac{r(r-1)}{2(2n-1)}}$$

Pour $r = 2n$, tous les animaux sont choisis, et l'on retrouve $\mathrm{E}(N) = n$; pour $r = 2$, $\mathrm{E}(N) = \dfrac{1}{2n-1}$, probabilité que le second animal soit le partenaire du premier. Les valeurs ont été vérifiées par énumération pour $n \leqslant 4$ et tous les r ✓.

À retenir. Les événements A_i ne sont pas indépendants, et la loi de N est compliquée ; la linéarité de l'espérance n'en demande pas tant. C'est tout l'intérêt de la méthode des indicatrices.

EXERCICE 74 ⋆⋆

Soit ω une issue. Comme $N(\omega) \in [\![1\,,n]\!]$, l'indicatrice $\mathbf{1}_{(N \geqslant i)}(\omega)$ vaut 1 pour $i \leqslant N(\omega)$ et 0 pour $i > N(\omega)$, d'où l'égalité suggérée :

$$S = \sum_{i=1}^{n} X_i\,\mathbf{1}_{(N \geqslant i)}.$$

Pour chaque i, la variable $\mathbf{1}_{(N \geqslant i)}$ est une fonction de N, donc elle est indépendante de X_i (proposition 16). Par la proposition 20, $\mathrm{E}\big(X_i\,\mathbf{1}_{(N \geqslant i)}\big) = \mathrm{E}(X_i)\mathrm{P}(N \geqslant i) = p\,\mathrm{P}(N \geqslant i)$, et la linéarité donne

$$\mathrm{E}(S) = p \sum_{i=1}^{n} \mathrm{P}(N \geqslant i).$$

Enfin, en appliquant le même raisonnement à la constante 1 au lieu des X_i, $N = \sum_{i=1}^{n} \mathbf{1}_{(N \geqslant i)}$, donc $\mathrm{E}(N) = \sum_{i=1}^{n} \mathrm{P}(N \geqslant i)$. Ainsi

$$\boxed{\mathrm{E}(S) = p\,\mathrm{E}(N)}$$

Pour $n = 3$, avec $\mathrm{P}(N=1) = \dfrac{1}{2}$, $\mathrm{P}(N=2) = \dfrac{1}{3}$ et $\mathrm{P}(N=3) = \dfrac{1}{6}$, le calcul exact sur l'univers produit donne $\mathrm{E}(S) = \dfrac{5}{3}p$ pour $p = \dfrac{1}{3}$ et $p = \dfrac{1}{2}$ ✓.

Le résultat est intuitif : on effectue en moyenne $\mathrm{E}(N)$ épreuves, qui réussissent chacune avec la probabilité p. Il n'est pourtant pas évident, puisque le nombre de termes de la somme est lui-même aléatoire.

Piège. L'erreur la plus répandue est d'écrire $\mathrm{E}(S) = \sum_{i=1}^{\mathrm{E}(N)} \mathrm{E}(X_i)$, ce qui n'a pas de sens si $\mathrm{E}(N)$ n'est pas entier, ou d'appliquer la linéarité à une somme dont le nombre de termes dépend de ω. L'écriture avec les indicatrices ramène S à une somme d'un nombre *fixe* de termes. L'indépendance entre N et les X_i est indispensable : si l'on arrête les épreuves au premier succès, N dépend des X_i, et le résultat tombe en défaut.

EXERCICE 75 ⋆⋆

1. On note $p_j = \mathrm{P}(X = j)$ pour $j \in [\![1\,,n]\!]$, et $m_k = \mathrm{E}(X^k)$. Par la formule de transfert,

$$m_k = \sum_{j=1}^{n} j^k p_j, \qquad k \in [\![0\,,n-1]\!].$$

C'est un système linéaire de n équations aux n inconnues $p_1, \ldots, p_n$, dont la matrice $\left(j^k\right)_{0\leqslant k\leqslant n-1,\ 1\leqslant j\leqslant n}$ est la transposée de la matrice de Vandermonde des réels 1, 2, …, n. Son déterminant, égal à celui de sa transposée (chapitre 17, théorème 6 (iv)), vaut $\prod_{1\leqslant i<j\leqslant n} (j-i) \neq 0$ (chapitre 17, proposition 12), puisque ces réels sont distincts. La matrice est inversible, et le système a une unique solution : $\boxed{\text{la loi de } X \text{ est déterminée par } \mathrm{E}(X^0), \ldots, \mathrm{E}(X^{n-1})}$. Deux variables à valeurs dans $[\![1\,,n]\!]$ qui ont les mêmes moments d'ordre 0 à $n-1$ ont la même loi.

2. Soit $x_1, \ldots, x_m$ les éléments, deux à deux distincts, de l'ensemble fini $X(\Omega) \cup Y(\Omega')$, où Ω' est l'univers sur lequel Y est définie. Soit $i \in [\![1\,,m]\!]$, et L_i le polynôme interpolateur de Lagrange tel que $L_i(x_j) = 1$ si $j = i$ et 0 sinon (chapitre 13, théorème 13). Par la formule de transfert, en sommant sur les x_j et en ajoutant des termes nuls,

$$\mathrm{E}\big(L_i(X)\big) = \sum_{j=1}^{m} L_i(x_j)\mathrm{P}(X = x_j) = \mathrm{P}(X = x_i),$$

et de même $\mathrm{E}\big(L_i(Y)\big) = \mathrm{P}(Y = x_i)$. Si $L_i = \sum_k a_k X^k$, la linéarité de l'espérance et l'hypothèse donnent

$$\mathrm{E}\big(L_i(X)\big) = \sum_k a_k \mathrm{E}(X^k) = \sum_k a_k \mathrm{E}(Y^k) = \mathrm{E}\big(L_i(Y)\big).$$

Ainsi $\mathrm{P}(X = x_i) = \mathrm{P}(Y = x_i)$ pour tout i, et

$$\boxed{X \text{ et } Y \text{ ont la même loi}}$$

On n'a utilisé que les moments d'ordre au plus $m-1$, puisque $\deg L_i \leqslant m-1$.
Piège. Les premiers moments ne suffisent pas. Supposons que X suive la loi uniforme sur $\{-1, 1\}$, et que Y prenne les valeurs $-\sqrt{2}$ et $\sqrt{2}$ avec la probabilité $\frac{1}{4}$ et la valeur 0 avec la probabilité $\frac{1}{2}$. Alors $\mathrm{E}(X) = \mathrm{E}(Y) = 0$ et $\mathrm{E}(X^2) = \mathrm{E}(Y^2) = 1$, mais $\mathrm{E}(X^4) = 1 \neq 2 = \mathrm{E}(Y^4)$. L'erreur la plus répandue est de croire que l'espérance et la variance déterminent la loi ; il faut autant de moments que de valeurs possibles, moins une.

EXERCICE 76 ⋆⋆

1. Soit $i<j$. L'application $\sigma \mapsto \sigma \circ (i\ j)$ est une bijection de $\mathfrak{S}_n$ sur lui-même, qui préserve donc la loi uniforme, et elle échange les événements $(\Sigma(i)<\Sigma(j))$ et $(\Sigma(i)>\Sigma(j))$. Ces deux événements, incompatibles et de réunion Ω puisque $\Sigma(i)\neq\Sigma(j)$, ont donc la même probabilité $\dfrac{1}{2}$; c'est le calcul de l'exercice 66 (question 1 (a)), fait pour $i=1$ et $j=2$. Le nombre d'inversions est $I=\displaystyle\sum_{i<j}\mathbf{1}_{(\Sigma(i)>\Sigma(j))}$, somme de $\binom{n}{2}$ indicatrices, et

$$\boxed{\mathrm{E}(I)=\frac{1}{2}\binom{n}{2}=\frac{n(n-1)}{4}}$$

Une permutation aléatoire a en moyenne la moitié du nombre maximal $\binom{n}{2}$ d'inversions, atteint par la permutation $k\mapsto n+1-k$. Les inversions sont les « dérangements » de Cramer évoqués au chapitre 17, dont la parité donne la signature (chapitre 17, démonstration de la proposition 4).

2. Soit $k\in[\![1\,,n]\!]$. La permutation Σ est croissante sur $[\![1\,,L]\!]$, donc sur $[\![1\,,k]\!]$ si $L\geqslant k$; réciproquement, si elle est croissante sur $[\![1\,,k]\!]$, le plus grand intervalle de cette forme sur lequel elle l'est a au moins k éléments. Ainsi $(L\geqslant k)=(\Sigma(1)<\Sigma(2)<\cdots<\Sigma(k))$. Une permutation de cet événement est déterminée par l'ensemble $\{\sigma(1),\dots,\sigma(k)\}$, que l'on choisit de $\binom{n}{k}$ façons, l'ordre étant alors imposé, et par la bijection de $[\![k+1\,,n]\!]$ sur les $n-k$ valeurs restantes, choisie de $(n-k)!$ façons :

$$\mathrm{P}(L\geqslant k)=\frac{1}{n!}\binom{n}{k}(n-k)!,\qquad\boxed{\mathrm{P}(L\geqslant k)=\frac{1}{k!}}$$

Comme $(L=k)=(L\geqslant k)\setminus(L\geqslant k+1)$, avec $(L\geqslant k+1)\subset(L\geqslant k)$, la proposition 7 (ii) donne

$$\boxed{\mathrm{P}(L=k)=\frac{1}{k!}-\frac{1}{(k+1)!}=\frac{k}{(k+1)!}\ \text{ pour } k\in[\![1\,,n-1]\!],\qquad \mathrm{P}(L=n)=\frac{1}{n!}}$$

3. Comme L est à valeurs dans $[\![1\,,n]\!]$, on a, pour toute issue, $L=\displaystyle\sum_{k=1}^{n}\mathbf{1}_{(L\geqslant k)}$: la somme compte les entiers $k\in[\![1\,,n]\!]$ tels que $k\leqslant L$, qui sont au nombre de L. La linéarité donne

$$\mathrm{E}(L)=\sum_{k=1}^{n}\mathrm{P}(L\geqslant k),\qquad\boxed{\mathrm{E}(L)=\sum_{k=1}^{n}\frac{1}{k!}\xrightarrow[n\to+\infty]{}\mathrm{e}-1}$$

par la série exponentielle du chapitre 20. La convergence est très rapide : pour $n=6$, $\mathrm{E}(L)\approx 1{,}71806$, contre $\mathrm{e}-1\approx 1{,}71828$. Les résultats ont été vérifiés par énumération pour $n\leqslant 6$ ✓.

À retenir. La formule $\mathrm{E}(L)=\displaystyle\sum_{k\geqslant 1}\mathrm{P}(L\geqslant k)$, pour une variable à valeurs entières positives, est souvent plus commode que la définition : ici, les probabilités $\mathrm{P}(L\geqslant k)$ sont plus simples que les $\mathrm{P}(L=k)$. Elle resservira dans les exercices 83 et 90.

EXERCICE 77 ⋆⋆

On modélise les n lancers par l'univers $\Omega=\{P,F\}^n$, muni de la probabilité qui donne à une suite comportant j piles la probabilité p^jq^{n-j} : par la proposition 17, les résultats des lancers sont alors mutuellement indépendants, et chacun donne pile avec la probabilité p.

1. Une n-chaîne de piles occupe tous les lancers ; elle existe si et seulement si tous les lancers donnent pile. Ainsi Y_n prend les valeurs 0 et 1, et $\mathrm{P}(Y_n=1)=p^n$:

$$\boxed{Y_n\sim\mathcal{B}(p^n),\qquad\mathrm{E}(Y_n)=p^n}$$

2. Deux chaînes distinctes sont séparées par au moins un face ; deux $(n-1)$-chaînes occuperaient donc au moins $2(n-1)+1=2n-1>n$ lancers. Il y a donc au plus une $(n-1)$-chaîne, et Y_{n-1} prend

les valeurs 0 et 1. Une $(n-1)$-chaîne occupe soit les lancers 1 à $n-1$, et elle est suivie d'un face au lancer n, soit les lancers 2 à n, et elle est précédée d'un face au lancer 1. Ces deux cas sont les issues $P\cdots PF$ et $FP\cdots P$, distinctes, d'où

$$\boxed{\mathrm{P}(Y_{n-1}=1)=2qp^{n-1}} \qquad \text{et} \qquad \boxed{\mathrm{E}(Y_{n-1})=2qp^{n-1}}$$

3. (a) Soit $k\in[\![1\,,n-2]\!]$. Une k-chaîne commence au lancer 1 si et seulement si les lancers 1 à k donnent pile et le lancer $k+1$, qui existe puisque $k+1\leqslant n$, donne face. De même, une k-chaîne commence au lancer $n-k+1$ si et seulement si les k derniers lancers donnent pile et le lancer $n-k\geqslant 1$ donne face. Par indépendance,

$$\boxed{\mathrm{P}(X_{1,k}=1)=\mathrm{P}(X_{n-k+1,k}=1)=qp^k}$$

Pour $i\in[\![2\,,n-k]\!]$, une k-chaîne commence au lancer i si et seulement si le lancer $i-1$ donne face, les lancers i à $i+k-1$ donnent pile, et le lancer $i+k\leqslant n$ donne face : $\boxed{\mathrm{P}(X_{i,k}=1)=q^2p^k}$.

(b) Chaque k-chaîne a un premier lancer, d'indice $i\in[\![1\,,n-k+1]\!]$, et deux chaînes distinctes commencent à des lancers distincts : $Y_k=\sum_{i=1}^{n-k+1}X_{i,k}$. Les variables $X_{i,k}$ suivent des lois de Bernoulli, et la linéarité donne, avec les deux extrémités et les $n-k-1$ positions intérieures,

$$\boxed{\mathrm{E}(Y_k)=2qp^k+(n-k-1)q^2p^k, \qquad k\in[\![1\,,n-2]\!]}$$

Pour l'exemple de l'énoncé, l'issue $PPFFPPPFPFP$ a bien $X_{1,2}=1$, $X_{5,3}=1$, $X_{9,1}=1$ et $X_{11,1}=1$, d'où $Y_2=Y_3=1$ et $Y_1=2$. Les trois formules ont été vérifiées par énumération pour $n\leqslant 7$, avec $p=\dfrac{1}{3}$ et $p=\dfrac{1}{2}$ ✓. La formule de la question 3 reste vraie pour $k=n-1$, où le second terme est nul, mais pas pour $k=n$, puisqu'une n-chaîne n'est entourée d'aucun face.

Piège. L'erreur la plus répandue est d'oublier les effets de bord : une chaîne placée au début ou à la fin n'a besoin que d'un face pour être délimitée, d'où le facteur q au lieu de q^2.

EXERCICE 78 ⋆⋆

Pour $i\in[\![1\,,n]\!]$, soit M_i l'indicatrice de l'événement « la vache i est malade » ; les M_i sont mutuellement indépendantes, de loi $\mathcal{B}(0{,}15)$, et l'on n'a pas besoin de décrire l'univers (proposition 17).

1. Si aucune vache n'est malade, la seconde méthode demande une analyse ; sinon, elle en demande $1+n$. L'événement « aucune vache n'est malade » est $(M_1=0,\dots,M_n=0)$, de probabilité $0{,}85^n$ par indépendance. La variable Y_n prend donc deux valeurs :

$$\boxed{\mathrm{P}\Big(Y_n=\frac{1}{n}\Big)=0{,}85^n, \qquad \mathrm{P}\Big(Y_n=1+\frac{1}{n}\Big)=1-0{,}85^n}$$

et

$$\mathrm{E}(Y_n)=\frac{1}{n}\,0{,}85^n+\Big(1+\frac{1}{n}\Big)(1-0{,}85^n), \qquad \boxed{\mathrm{E}(Y_n)=1+\frac{1}{n}-0{,}85^n}$$

2. La fonction f est dérivable sur $]0\,,+\infty[$, et $f'(x)=a+\dfrac{1}{x}$, du signe de $1+ax$, puisque $x>0$. Comme $a<0$, f' s'annule en $x_0=-\dfrac{1}{a}\approx 6{,}15$, et

$$\boxed{f \text{ croît sur }]0\,,x_0] \text{ et décroît sur } [x_0\,,+\infty[\,, \qquad x_0=-\frac{1}{\ln(0{,}85)}\approx 6{,}15}$$

De plus $f(x)\to-\infty$ quand $x\to 0$, et aussi quand $x\to+\infty$, puisque $\ln x=o(x)$ (chapitre 10) ; le maximum vaut $f(x_0)=-1+\ln(x_0)\approx 0{,}817>0$. Par stricte monotonie de part et d'autre de x_0,

l'ensemble des x tels que $f(x) > 0$ est un intervalle ouvert qui contient x_0. Or $f(n) > 0$ équivaut à $n\,0{,}85^n > 1$, et

$$2 \times 0{,}85^2 = 1{,}445 > 1, \qquad 17 \times 0{,}85^{17} \approx 1{,}073 > 1, \qquad 18 \times 0{,}85^{18} \approx 0{,}966 < 1.$$

Comme $2 < x_0 < 17$, la monotonie de f sur $[2\,,x_0]$ et sur $[x_0\,,17]$ donne $f(x) \geqslant \min\big(f(2), f(17)\big) > 0$ sur $[2\,,17]$, puis $f(n) \leqslant f(18) < 0$ pour $n \geqslant 18$, par décroissance :

$$\boxed{f(n) > 0 \iff n \in [\![2\,,17]\!]}$$

3. Par la question 1, $\mathrm{E}(Y_n) < 1 \iff \dfrac{1}{n} < 0{,}85^n \iff -\ln n < n\ln(0{,}85) \iff f(n) > 0$, la fonction ln étant strictement croissante. La seconde méthode demande en moyenne moins d'une analyse par vache, et elle est donc avantageuse en moyenne, pour les troupeaux de 2 à 17 vaches.

Pour l'effectif optimal, on calcule

n	2	3	4	5	6	7
$\mathrm{E}(Y_n)$	0,7775	0,7192	0,7280	0,7563	0,7895	0,8223

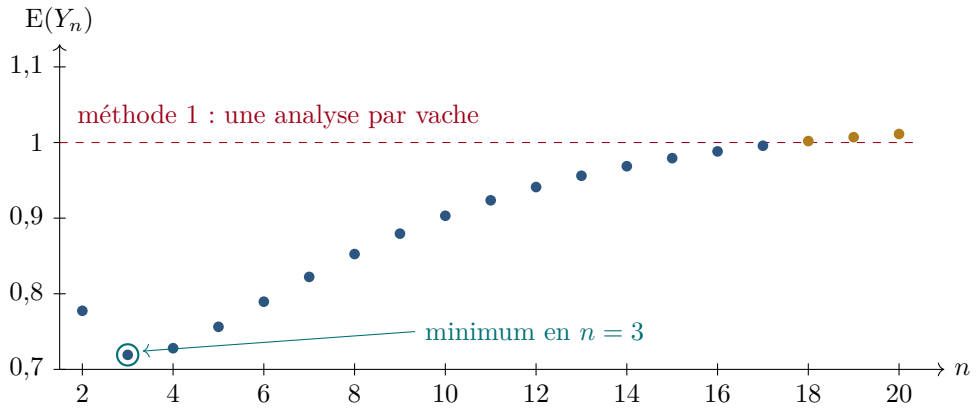
les valeurs autres que la première étant arrondies. Pour $n \geqslant 8$, on a $0{,}85^n \leqslant 0{,}85^8$, donc

$$\mathrm{E}(Y_n) > 1 - 0{,}85^n \geqslant 1 - 0{,}85^8 \approx 0{,}7275 > \mathrm{E}(Y_3).$$

Ainsi l'effectif $n = 3$ minimise le nombre moyen d'analyses par vache :

$$\boxed{n = 3, \qquad \mathrm{E}(Y_3) = 1 + \frac{1}{3} - 0{,}85^3 \approx 0{,}719}$$

En regroupant les vaches par trois, on économise en moyenne 28 % des analyses.



Les points bleus, de $n = 2$ à $n = 17$, sont sous la droite rouge d'ordonnée 1, qui représente la première méthode : le regroupement y est avantageux ; à partir de $n = 18$, en ocre, il ne l'est plus. Le minimum, entouré, est atteint pour $n = 3$, bien avant le maximum $x_0 \approx 6{,}15$ de f, qui ne mesure que le signe de $1 - \mathrm{E}(Y_n)$, et non sa valeur.

Piège. L'erreur la plus répandue est de chercher l'effectif optimal en maximisant f, ce qui donnerait $n = 6$: la fonction f ne sert qu'à comparer $\mathrm{E}(Y_n)$ à 1, et la quantité à minimiser est $\dfrac{1}{n} - 0{,}85^n$, qui n'a pas les mêmes variations.

EXERCICE 79 ⋆⋆

1. Pour chaque i, $p_i \in [0\,,1]$, donc $\ln p_i \leqslant 0$ si $p_i > 0$, et $-p_i \ln p_i \geqslant 0$, ce qui reste vrai, avec la convention, si $p_i = 0$. Ainsi $\boxed{H(X) \geqslant 0}$, comme somme de termes positifs. Elle est nulle si et seulement si chaque terme est nul, c'est-à-dire si $p_i \in \{0, 1\}$ pour tout i ; comme $\sum p_i = 1$, cela signifie qu'un des p_i vaut 1 et les autres 0 :

$$\boxed{H(X) = 0 \iff \exists i,\ p_i = 1}$$

Une variable d'entropie nulle prend une valeur avec la probabilité 1 : il n'y a aucune incertitude sur sa valeur.

2. Soit $\varphi(x) = 1 - x + x\ln x$ pour $x > 0$, et $\varphi(0) = 1$, conformément à la convention. Sur $]0\,,+\infty[$, $\varphi'(x) = -1 + \ln x + 1 = \ln x$: φ est strictement décroissante sur $]0\,,1]$ et strictement croissante sur $[1\,,+\infty[$, de minimum $\varphi(1) = 0$. Donc $\varphi(x) > 0$ pour $x > 0$, $x \neq 1$, et $\varphi(0) = 1 > 0$:

$$\boxed{-x\ln x \leqslant 1 - x \ \text{ pour tout } x \geqslant 0, \text{ avec égalité si et seulement si } x = 1}$$

Pour $x = np_k \geqslant 0$, on obtient $\boxed{-np_k \ln(np_k) \leqslant 1 - np_k}$, avec égalité si et seulement si $p_k = \dfrac{1}{n}$.

3. Si $p_k > 0$, $-np_k \ln(np_k) = -np_k \ln n - np_k \ln p_k$, et cette égalité reste vraie, avec la convention, si $p_k = 0$, les trois termes étant nuls. En sommant les inégalités de la question 2 sur k, et comme $\sum\limits_k p_k = 1$,

$$-n\ln n + nH(X) = \sum_{k=1}^{n} \big(-np_k \ln(np_k)\big) \leqslant \sum_{k=1}^{n} (1 - np_k) = n - n = 0, \qquad \text{d'où} \qquad \boxed{H(X) \leqslant \ln n}$$

Il y a égalité si et seulement si chacune des n inégalités est une égalité, c'est-à-dire si $p_k = \dfrac{1}{n}$ pour tout k :

$$\boxed{H(X) = \ln n \iff X \text{ suit la loi uniforme sur } \{x_1, \dots, x_n\}}$$

La majoration a été contrôlée sur 1 000 lois tirées au hasard, dont certaines ont des probabilités nulles ✓.

L'entropie mesure l'incertitude sur la valeur de X : elle est nulle pour une variable constante, et maximale pour la loi uniforme, où aucune valeur n'est privilégiée. Elle ne dépend que des probabilités, et non des valeurs x_i ; pour la loi uniforme sur 2^m valeurs, $H = m\ln 2$, soit m « bits » d'information.
Piège. Certaines probabilités p_k peuvent être nulles : c'est pourquoi l'inégalité de la question 2 est demandée pour $x \geqslant 0$, et non seulement pour $x > 0$. L'erreur la plus répandue est d'écrire $\ln(np_k)$ sans vérifier que $p_k > 0$.

EXERCICE 80 ⋆⋆⋆

On prend pour univers $\Omega = [\![1\,,n]\!]^r$, ensemble des r-listes $\omega = (T_1, \dots, T_r)$, où T_j est le tiroir de la boule j, muni de la probabilité uniforme ; il a n^r éléments (proposition 3). Par la proposition 17, les variables T_1, …, T_r sont mutuellement indépendantes, de loi uniforme sur $[\![1\,,n]\!]$.

1. Pour $k \in [\![1\,,n]\!]$, on a $X_k = \sum\limits_{j=1}^{r} \mathbf{1}_{(T_j = k)}$. Les indicatrices $\mathbf{1}_{(T_j=k)}$, fonctions des T_j, sont mutuellement indépendantes (proposition 16), de loi $\mathcal{B}(\dfrac{1}{n})$. Par le théorème 8,

$$\boxed{X_k \sim \mathcal{B}(r, \tfrac{1}{n}), \qquad \mathrm{E}(X_k) = \frac{r}{n}}$$

2. Les variables X_1, …, X_n ont pour somme r, et ne sont pas indépendantes. Par exemple, comme $n \geqslant 2$ et $r \geqslant 1$, l'événement $(X_1 = r, X_2 = r)$ est impossible, puisque $X_1 + X_2 \leqslant r < 2r$, alors que $\mathrm{P}(X_1 = r)\mathrm{P}(X_2 = r) = \dfrac{1}{n^{2r}} > 0$. Par la proposition 15, déjà $\boxed{X_1 \text{ et } X_2 \text{ ne sont pas indépendantes}}$, et la famille ne l'est pas non plus.

3. On a $V = \sum_{k=1}^{n} \mathbf{1}_{(X_k=0)}$, et $\mathrm{P}(X_k = 0) = (1 - \frac{1}{n})^r$ par la question 1. Par la méthode des indicatrices,

$$\boxed{\mathrm{E}(V) = n\Big(1 - \frac{1}{n}\Big)^r}$$

Pour $r = n$, $\dfrac{\mathrm{E}(V)}{n} = (1 - \frac{1}{n})^n = \exp\big(n \ln(1 - \frac{1}{n})\big)$, et $n \ln(1 - \frac{1}{n}) = n(-\frac{1}{n} + o(\frac{1}{n})) \to -1$ (chapitre 10), d'où

$$\boxed{\frac{\mathrm{E}(V)}{n} \xrightarrow[n\to+\infty]{} \frac{1}{\mathrm{e}} \approx 0{,}368}$$

Quand on répartit au hasard n boules dans n tiroirs, plus d'un tiroir sur trois reste vide en moyenne, alors qu'une répartition équitable n'en laisserait aucun. Pour $n = 100$, la proportion moyenne vaut déjà 0,366. On retrouve la même limite $\dfrac{1}{\mathrm{e}}$ que pour la probabilité qu'une permutation aléatoire n'ait aucun point fixe (remarque de la section 21.6.3).

4. Soit $j \in [\![0\,,n]\!]$ et J une partie de $[\![1\,,n]\!]$ à j éléments. Les répartitions pour lesquelles l'ensemble des tiroirs vides est exactement J sont les applications de $[\![1\,,r]\!]$ dans $[\![1\,,n]\!]$ dont l'image est $[\![1\,,n]\!] \setminus J$, c'est-à-dire les surjections de $[\![1\,,r]\!]$ sur cet ensemble à $n - j$ éléments : il y en a $S(r, n-j)$, avec les notations de l'exercice 15. En regroupant selon J, qui se choisit de $\binom{n}{j}$ façons,

$$\boxed{\mathrm{P}(V = j) = \frac{1}{n^r}\binom{n}{j} S(r, n-j) = \frac{1}{n^r}\binom{n}{j} \sum_{i=0}^{n-j} (-1)^{n-j-i} \binom{n-j}{i} i^r, \qquad j \in [\![0\,,n]\!]}$$

par la formule de l'exercice 15 (proposition 6 du cours), avec $S(r, 0) = 0$ puisque $r \geqslant 1$: on a toujours $V \leqslant n - 1$. Comme $S(r, p) = 0$ pour $p > r$, on a aussi $V \geqslant n - r$, ce qui est clair : r boules occupent au plus r tiroirs. Les lois de X_1 et de V, et l'espérance de V, ont été vérifiées par énumération pour $n \leqslant 4$ et $r \leqslant 4$ ✓.

À retenir. La question 3 donne l'espérance de V en une ligne, alors que sa loi, à la question 4, demande le nombre de surjections, qui n'a pas d'expression simple. C'est l'exemple type de la méthode des indicatrices : les événements $(X_k = 0)$ ne sont pas indépendants, et la linéarité n'en a pas besoin.

Variance et covariance

EXERCICE 81 ⋆

On note $q = 1 - p$. Le couple (X, Y) prend les quatre valeurs de $\{0, 1\}^2$, avec les probabilités q^2, pq, qp, p^2 par indépendance, et le couple $(U, V) = (X + Y, X - Y)$ prend les valeurs $(0, 0)$, $(1, 1)$, $(1, -1)$ et $(2, 0)$, deux à deux distinctes. On en déduit la loi conjointe, en lisant u en ligne et v en colonne :

	$v = -1$	$v = 0$	$v = 1$	$\mathrm{P}(U = u)$
$u = 0$	0	q^2	0	q^2
$u = 1$	pq	0	pq	$2pq$
$u = 2$	0	p^2	0	p^2
$\mathrm{P}(V = v)$	pq	$p^2 + q^2$	pq	1

Par la bilinéarité et la symétrie de la covariance (proposition 23),

$$\mathrm{Cov}(U, V) = \mathrm{Cov}(X, X) - \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, X) - \mathrm{Cov}(Y, Y) = \mathrm{V}(X) - \mathrm{V}(Y),$$

d'où $\boxed{\mathrm{Cov}(U, V) = 0}$, puisque X et Y ont la même loi. Pourtant $\mathrm{P}(U = 1, V = 0) = 0$, alors que le produit $\mathrm{P}(U = 1)\mathrm{P}(V = 0)$ vaut $2pq(p^2 + q^2) > 0$:

$$\boxed{U \text{ et } V \text{ sont décorrélées, mais pas indépendantes}}$$

On le lit sur le tableau : la case centrale est vide, alors que sa ligne et sa colonne ne le sont pas. Savoir que $U = 1$ force $V = \pm 1$.
Piège. C'est le phénomène de la remarque qui suit la proposition 20 : l'indépendance entraîne la décorrélation (proposition 23 (iii)), et la réciproque est fausse. L'erreur la plus répandue est de conclure à l'indépendance à partir d'une covariance nulle.

EXERCICE 82 ⋆

L'espérance et la variance ne dépendent que de la loi (section 21.6.1) : comme X et $aX+b$ ont la même loi, $\mathrm{E}(X) = a\mathrm{E}(X) + b$ et $\mathrm{V}(X) = \mathrm{V}(aX+b) = a^2\mathrm{V}(X)$ (théorème 11 (ii)).

1. Si $a \notin \{-1, 1\}$, alors $1 - a^2 \neq 0$, et $(1-a^2)\mathrm{V}(X) = 0$ donne $\mathrm{V}(X) = 0$. Par le théorème 11 (iii), $\mathrm{P}(X = \mathrm{E}(X)) = 1$, et comme $a \neq 1$, la relation $\mathrm{E}(X) = a\mathrm{E}(X) + b$ donne $\mathrm{E}(X) = \dfrac{b}{1-a}$:

$$\boxed{\mathrm{P}(X = c) = 1, \qquad c = \frac{b}{1-a}}$$

Réciproquement, une variable égale à c avec la probabilité 1 convient, puisque $ac + b = c$.

2. Si $a = 1$, la relation $\mathrm{E}(X) = \mathrm{E}(X) + b$ donne $b = 0$: $\boxed{\text{le cas } a = 1,\ b \neq 0 \text{ est impossible}}$. On peut aussi raisonner sans espérance : la plus grande valeur prise avec une probabilité non nulle serait la même pour X et pour $X + b$, ce qui impose $b = 0$.

3. Si $a = -1$ et $b \neq 0$, soit X de loi uniforme sur $\{0, b\}$: la variable $b - X$ prend les valeurs b et 0, chacune avec la probabilité $\dfrac{1}{2}$, donc $b - X \sim X$, et X n'est pas constante. Si $b = 0$, X de loi uniforme sur $\{-1, 1\}$ convient. Plus généralement, toute variable dont la loi est symétrique par rapport à $\dfrac{b}{2}$ convient :

$$\boxed{X \sim \mathcal{U}(\{0, b\}) \ \text{ si } b \neq 0, \qquad X \sim \mathcal{U}(\{-1, 1\}) \ \text{ si } b = 0}$$

Dans ce cas, la relation entre espérances impose seulement $\mathrm{E}(X) = \dfrac{b}{2}$, et la relation entre variances est toujours vraie.
Piège. L'énoncé dit que X et $aX + b$ ont la même *loi*, et non qu'elles sont égales : on ne peut pas écrire $X = aX + b$ et résoudre. C'est la mise en garde de la remarque qui suit la proposition 13.

EXERCICE 83 ⋆⋆

1. Soit $k \in [\![0\,, n]\!]$. On a $(M \leqslant k) = (X \leqslant k) \cap (Y \leqslant k)$, et, par indépendance,

$$\boxed{\mathrm{P}(M \leqslant k) = \mathrm{P}(X \leqslant k)\mathrm{P}(Y \leqslant k) = \frac{k^2}{n^2}}$$

Pour $k \in [\![1\,, n]\!]$, $(M = k) = (M \leqslant k) \setminus (M \leqslant k-1)$, d'où

$$\boxed{\mathrm{P}(M = k) = \frac{k^2 - (k-1)^2}{n^2} = \frac{2k-1}{n^2}}$$

Avec les sommes usuelles du chapitre 3,

$$\mathrm{E}(M) = \frac{1}{n^2}\sum_{k=1}^{n}(2k^2 - k) = \frac{1}{n^2}\Big(\frac{n(n+1)(2n+1)}{3} - \frac{n(n+1)}{2}\Big), \qquad \boxed{\mathrm{E}(M) = \frac{(n+1)(4n-1)}{6n}}$$

2. Pour toute issue, $\{M, N\} = \{X, Y\}$ en tant que valeurs, donc $M + N = X + Y$. Par linéarité, et avec $\mathrm{E}(X) = \mathrm{E}(Y) = \dfrac{n+1}{2}$ (proposition 19),

$$\mathrm{E}(N) = (n+1) - \frac{(n+1)(4n-1)}{6n}, \qquad \boxed{\mathrm{E}(N) = \frac{(n+1)(2n+1)}{6n}}$$

3. Comme N est à valeurs dans $[\![1\,,n]\!]$, $N=\sum_{k=1}^{n}\mathbf{1}_{(N\geqslant k)}$ pour toute issue, comme à l'exercice 76, d'où $\boxed{\mathrm{E}(N)=\sum_{k=1}^{n}\mathrm{P}(N\geqslant k)}$. Or $(N\geqslant k)=(X\geqslant k)\cap(Y\geqslant k)$, de probabilité $\big(\frac{n-k+1}{n}\big)^2$, et, en posant $j=n-k+1$,

$$\mathrm{E}(N)=\frac{1}{n^2}\sum_{j=1}^{n}j^2=\frac{(n+1)(2n+1)}{6n},$$

ce qui confirme la question 2 ✓.

4. Par la formule de transfert et la somme des cubes,

$$\mathrm{E}(M^2)=\frac{1}{n^2}\sum_{k=1}^{n}(2k^3-k^2)=\frac{1}{n^2}\Big(\frac{n^2(n+1)^2}{2}-\frac{n(n+1)(2n+1)}{6}\Big)=\frac{(n+1)(3n^2+n-1)}{6n}.$$

La formule de König-Huygens (théorème 11 (i)) donne, après réduction au dénominateur $36n^2$,

$$\mathrm{V}(M)=\frac{(n+1)(3n^2+n-1)}{6n}-\frac{(n+1)^2(4n-1)^2}{36n^2},\qquad\boxed{\mathrm{V}(M)=\frac{(n^2-1)(2n^2+1)}{36n^2}}$$

Pour $n=1$, on trouve $\mathrm{V}(M)=0$, puisque $M=1$. Les quatre résultats ont été vérifiés par énumération pour $n\leqslant 7$ ✓. Quand n est grand, $\mathrm{E}(M)\sim\frac{2n}{3}$ et $\mathrm{E}(N)\sim\frac{n}{3}$: les deux tirages partagent l'intervalle en trois parts égales en moyenne, ce que l'exercice 92 retrouvera.

Piège. L'erreur la plus répandue est d'écrire que la probabilité $\mathrm{P}(M=k)$ vaut $\mathrm{P}(X=k)\mathrm{P}(Y=k)$, ou bien $\mathrm{P}(X=k)+\mathrm{P}(Y=k)$. L'événement $(M=k)$ ne s'écrit pas simplement en termes de X et de Y ; c'est l'événement $(M\leqslant k)$ qui est une intersection, d'où le passage par les probabilités $\mathrm{P}(M\leqslant k)$.

EXERCICE 84 ⋆⋆

Pour une variable Z et un réel b, la linéarité de l'espérance donne

$$\mathrm{E}((Z-b)^2)=\mathrm{E}(Z^2)-2b\mathrm{E}(Z)+b^2=\mathrm{V}(Z)+\big(\mathrm{E}(Z)-b\big)^2,$$

par la formule de König-Huygens. On l'applique à $Z=Y-aX$, dont la variance se développe par bilinéarité (proposition 23) :

$$\mathrm{V}(Y-aX)=\mathrm{V}(Y)-2a\,\mathrm{Cov}(X,Y)+a^2\mathrm{V}(X).$$

On obtient, en posant $a^*=\dfrac{\mathrm{Cov}(X,Y)}{\mathrm{V}(X)}$ et en complétant le carré,

$$\mathrm{E}\Big(\big(Y-(aX+b)\big)^2\Big)=\mathrm{V}(X)\,(a-a^*)^2+\big(b-\mathrm{E}(Y)+a\mathrm{E}(X)\big)^2+\mathrm{V}(Y)-\frac{\mathrm{Cov}(X,Y)^2}{\mathrm{V}(X)}.$$

Les deux premiers termes sont positifs, puisque $\mathrm{V}(X)>0$, et ils sont nuls simultanément si et seulement si $a=a^*$ et $b=\mathrm{E}(Y)-a^*\mathrm{E}(X)$. Ainsi le minimum est atteint en un unique couple :

$$\boxed{a=\frac{\mathrm{Cov}(X,Y)}{\mathrm{V}(X)},\qquad b=\mathrm{E}(Y)-a\,\mathrm{E}(X),\qquad \min=\mathrm{V}(Y)-\frac{\mathrm{Cov}(X,Y)^2}{\mathrm{V}(X)}}$$

Ce minimum est positif : on retrouve l'inégalité de Cauchy-Schwarz $\mathrm{Cov}(X,Y)^2\leqslant\mathrm{V}(X)\mathrm{V}(Y)$ du théorème 12, et il est nul si et seulement si $\mathrm{P}(Y=aX+b)=1$ (théorème 11 (iii)). Si $\mathrm{V}(Y)>0$, il s'écrit $\mathrm{V}(Y)(1-\rho^2)$, où ρ est le coefficient de corrélation : plus $|\rho|$ est proche de 1, mieux Y est approchée par une fonction affine de X.

Si l'on munit un nuage de points (x_1,y_1), …, (x_m,y_m) de la probabilité uniforme, et que X et Y sont l'abscisse et l'ordonnée d'un point tiré au hasard, la quantité minimisée est $\dfrac{1}{m}\sum_i\big(y_i-(ax_i+b)\big)^2$,

et l'on retrouve exactement la droite des moindres carrés de la proposition 14 du chapitre 19 : pente $a = \dfrac{\mathrm{Cov}(X,Y)}{\mathrm{V}(X)}$, comme le dit la remarque qui suit le théorème 12, et droite passant par le point moyen $(\bar{x},\bar{y}) = (\mathrm{E}(X),\mathrm{E}(Y))$.
À retenir. La minimisation se fait sans dériver, en complétant deux carrés : d'abord en b, à a fixé, puis en a. C'est la méthode du chapitre 19, où la droite des moindres carrés était une projection orthogonale ; ici, $aX+b$ est la meilleure approximation de Y par une fonction affine de X, au sens de l'écart quadratique moyen.

EXERCICE 85 ⋆⋆

On note $a = \mathrm{P}(A)$, $b = \mathrm{P}(B)$ et $c = \mathrm{P}(A\cap B)$. Comme $\mathbf{1}_A\mathbf{1}_B = \mathbf{1}_{A\cap B}$, la proposition 23 (i) et le théorème 9 (iv) donnent

$$\mathrm{Cov}(\mathbf{1}_A,\mathbf{1}_B) = \mathrm{E}(\mathbf{1}_{A\cap B}) - \mathrm{E}(\mathbf{1}_A)\mathrm{E}(\mathbf{1}_B) = c - ab.$$

Les indicatrices suivent les lois $\mathcal{B}(a)$ et $\mathcal{B}(b)$, de variances $a(1-a)$ et $b(1-b)$ (proposition 21). Par l'inégalité de Cauchy-Schwarz (théorème 12), puis la majoration $x(1-x) = \dfrac{1}{4} - (x-\dfrac{1}{2})^2 \leqslant \dfrac{1}{4}$,

$$|c-ab| \leqslant \sqrt{a(1-a)}\,\sqrt{b(1-b)} \leqslant \frac{1}{2}\times\frac{1}{2}, \qquad \boxed{|\mathrm{P}(A\cap B) - \mathrm{P}(A)\mathrm{P}(B)| \leqslant \frac{1}{4}}$$

Cas d'égalité. S'il y a égalité, les deux inégalités sont des égalités, donc $a(1-a) = b(1-b) = \dfrac{1}{4}$, c'est-à-dire $a = b = \dfrac{1}{2}$, et $|c-\dfrac{1}{4}| = \dfrac{1}{4}$, soit $c = \dfrac{1}{2}$ ou $c = 0$.

— Si $c = \dfrac{1}{2}$, alors $\mathrm{P}(A\setminus B) = a - c = 0$ et $\mathrm{P}(B\setminus A) = b - c = 0$ (proposition 7 (ii)) : les événements A et B ne diffèrent que par un événement de probabilité nulle, et $\mathrm{Cov}(\mathbf{1}_A,\mathbf{1}_B) = \dfrac{1}{4}$.

— Si $c = 0$, alors $\mathrm{P}(A\cup B) = a+b-c = 1$: l'événement $\overline{B}$ ne diffère de A que par un événement de probabilité nulle, puisque $\mathrm{P}(A\cap B) = 0$ et $\mathrm{P}(\overline{A}\cap\overline{B}) = 1 - \mathrm{P}(A\cup B) = 0$, et $\mathrm{Cov}(\mathbf{1}_A,\mathbf{1}_B) = -\dfrac{1}{4}$.

Réciproquement, ces deux situations donnent l'égalité. Ainsi

$$\boxed{\text{égalité} \iff \mathrm{P}(A) = \mathrm{P}(B) = \tfrac{1}{2} \text{ et } \big(\mathrm{P}(A\cap\overline{B}) = \mathrm{P}(\overline{A}\cap B) = 0 \ \text{ ou } \ \mathrm{P}(A\cap B) = \mathrm{P}(\overline{A}\cap\overline{B}) = 0\big)}$$

Pour un lancer de pièce équilibrée, $A = B =$ « pile » donne $\dfrac{1}{2} - \dfrac{1}{4} = \dfrac{1}{4}$, et $A =$ « pile », $B =$ « face » donne $0 - \dfrac{1}{4} = -\dfrac{1}{4}$.
À retenir. La quantité $\mathrm{P}(A\cap B) - \mathrm{P}(A)\mathrm{P}(B)$ mesure l'écart à l'indépendance ; elle est bornée par $\dfrac{1}{4}$, et cette borne n'est atteinte que dans le cas le plus dépendant qui soit, celui où chaque événement détermine l'autre. La borne $\dfrac{1}{4}$ vient de la proposition 22, qui majore par $\dfrac{1}{4}$ la variance d'une variable à valeurs dans $[0\,,1]$, ici celle de chaque indicatrice.

EXERCICE 86 ⋆⋆

1. Par linéarité et la proposition 19, $\mathrm{E}(Z) = \dfrac{n}{2} - \dfrac{n}{2}$, soit $\boxed{\mathrm{E}(Z) = 0}$. Les variables X et $-Y$ sont indépendantes (proposition 16), et le théorème 13, avec $\mathrm{V}(-Y) = \mathrm{V}(Y)$ (théorème 11 (ii)) et la proposition 21, donne

$$\boxed{\mathrm{V}(Z) = \mathrm{V}(X) + \mathrm{V}(Y) = \frac{n}{4} + \frac{n}{4} = \frac{n}{2}}$$

2. Posons $W = n - Y$. Pour $k \in [\![0\,,n]\!]$, $\mathrm{P}(W = k) = \mathrm{P}(Y = n-k) = \binom{n}{n-k}\dfrac{1}{2^n} = \binom{n}{k}\dfrac{1}{2^n}$, par symétrie des coefficients binomiaux : W suit la loi $\mathcal{B}(n, \dfrac{1}{2})$, et elle est indépendante de X (proposition 16). Pour $m \in [\![0\,,2n]\!]$, la proposition 7 (v), appliquée au système complet $(X = k)_k$, puis l'indépendance, donnent

$$\mathrm{P}(X + W = m) = \sum_{k=0}^{n} \mathrm{P}(X = k, W = m - k) = \frac{1}{4^n}\sum_{k=0}^{n} \binom{n}{k}\binom{n}{m-k} = \frac{1}{4^n}\binom{2n}{m},$$

par la formule de Vandermonde, démontrée par dénombrement après la proposition 4 (on convient que $\binom{n}{j} = 0$ si $j \notin [\![0\,,n]\!]$). Ainsi $X + W \sim \mathcal{B}(2n, \dfrac{1}{2})$. Comme $Z = X + W - n$,

$$\boxed{\mathrm{P}(Z = k) = \mathrm{P}(X + W = n + k) = \frac{1}{4^n}\binom{2n}{n+k}, \qquad k \in [\![-n\,,n]\!]}$$

La loi de Z est symétrique autour de 0, et sa variance $\dfrac{2n}{4} = \dfrac{n}{2}$, celle de la loi $\mathcal{B}(2n, \dfrac{1}{2})$, confirme la question 1 ✓.

3. Par bilinéarité, et comme $\mathrm{Cov}(X, Y) = 0$ par indépendance (proposition 23 (iii)),

$$\mathrm{Cov}(X, Z) = \mathrm{Cov}(X, X) - \mathrm{Cov}(X, Y) = \mathrm{V}(X), \qquad \boxed{\mathrm{Cov}(X, Z) = \frac{n}{4} \neq 0}$$

Par contraposée de la proposition 23 (iii), $\boxed{X \text{ et } Z \text{ ne sont pas indépendantes}}$. C'est naturel : lorsque X est grand, $Z = X - Y$ a tendance à l'être aussi.

EXERCICE 87 ⋆⋆

1. Les nombres de l'énoncé sont positifs. Pour calculer leur somme, on sépare la ligne $j = 0$, formée de n termes égaux à $\dfrac{q^n}{n}$, et les couples (k, k) avec $k \geqslant 1$:

$$\sum_{j,k} \mathrm{P}(X = j, Y = k) = q^n + \sum_{k=1}^{n} \binom{n}{k} p^k q^{n-k} = q^n + (p+q)^n - q^n = 1,$$

par la formule du binôme : il s'agit bien d'une loi conjointe.

Par la proposition 14, on a $\mathrm{P}(X = 0) = n \times \dfrac{q^n}{n}$, soit q^n, et, pour $j \geqslant 1$, $\mathrm{P}(X = j) = \binom{n}{j} p^j q^{n-j}$:

$$\boxed{X \sim \mathcal{B}(n, p), \qquad \mathrm{E}(X) = np}$$

Pour $k \in [\![1\,,n]\!]$, la colonne k contient $\mathrm{P}(X = 0, Y = k) = \dfrac{q^n}{n}$ et $\mathrm{P}(X = k, Y = k)$:

$$\boxed{\mathrm{P}(Y = k) = \binom{n}{k} p^k q^{n-k} + \frac{q^n}{n}}, \qquad \mathrm{E}(Y) = \sum_{k=1}^{n} k\binom{n}{k} p^k q^{n-k} + \frac{q^n}{n}\sum_{k=1}^{n} k,$$

d'où, par la proposition 19 et la somme des entiers,

$$\boxed{\mathrm{E}(Y) = np + \frac{(n+1)q^n}{2}}$$

On peut décrire le couple ainsi : X suit une loi binomiale, $Y = X$ lorsque $X \geqslant 1$, et lorsque $X = 0$, Y est choisie uniformément dans $[\![1\,,n]\!]$.

2. Comme p et q sont dans $]0\,,1[$, $\mathrm{P}(X=j) > 0$ pour tout $j \in [\![0\,,n]\!]$. Par définition,

$$\mathrm{P}_{(X=j)}(Y=k) = \frac{\mathrm{P}(X=j, Y=k)}{\mathrm{P}(X=j)}.$$

Pour $j \geqslant 1$, ce quotient vaut 1 si $k=j$ et 0 sinon ; pour $j=0$, il vaut $\dfrac{q^n/n}{q^n} = \dfrac{1}{n}$ pour tout k :

$$\boxed{\mathrm{P}_{(X=j)}(Y=k) = \begin{cases} 1 \text{ si } k=j,\ 0 \text{ sinon} & (j \geqslant 1), \\ \dfrac{1}{n} & (j=0). \end{cases}}$$

Sachant $(X=j)$ avec $j \geqslant 1$, la variable Y vaut j, et la moyenne de la loi conditionnelle est $\boxed{j}$; sachant $(X=0)$, la loi conditionnelle de Y est uniforme sur $[\![1\,,n]\!]$, de moyenne $\boxed{\dfrac{n+1}{2}}$ (proposition 19).

3. (a) On a $\mathrm{P}(X=1, Y=1) = npq^{n-1} = \mathrm{P}(X=1)$, et $\mathrm{P}(X=1) > 0$. L'égalité de $\mathrm{P}(X=1, Y=1)$ et de $\mathrm{P}(X=1)\mathrm{P}(Y=1)$ équivaudrait donc à $\mathrm{P}(Y=1) = 1$. Or, comme $n \geqslant 2$, on a $\mathrm{P}(Y=2) \geqslant \binom{n}{2} p^2 q^{n-2}$, qui est strictement positif, donc $\mathrm{P}(Y=1) < 1$, et

$$\boxed{\mathrm{P}(X=1, Y=1) \neq \mathrm{P}(X=1)\mathrm{P}(Y=1)}$$

Les variables X et Y ne sont donc pas indépendantes, quelle que soit la valeur de q.

(b) Par la formule de transfert, les termes où $X=0$ sont nuls, et ceux où $X = j \geqslant 1$ imposent $Y = j$:

$$\mathrm{E}(XY) = \sum_{j=1}^{n} j^2 \binom{n}{j} p^j q^{n-j} = \mathrm{E}(X^2) = npq + n^2p^2,$$

par la formule de König-Huygens et la proposition 21. D'où

$$\mathrm{Cov}(X,Y) = npq + n^2p^2 - np\Big(np + \frac{(n+1)q^n}{2}\Big), \qquad \boxed{\mathrm{Cov}(X,Y) = npq\Big(1 - \frac{(n+1)q^{n-1}}{2}\Big)}$$

Comme $npq > 0$, la covariance s'annule si et seulement si $q^{n-1} = \dfrac{2}{n+1}$. Comme $n \geqslant 2$, $\dfrac{2}{n+1} \in]0\,,1[$, et la fonction $q \mapsto q^{n-1}$ est une bijection strictement croissante de $]0\,,1[$ sur lui-même : il y a une unique solution,

$$\boxed{q_0 = \Big(\frac{2}{n+1}\Big)^{1/(n-1)}}$$

Par exemple $q_0 = \dfrac{2}{3}$ pour $n=2$, et $q_0 = \dfrac{1}{\sqrt{2}} \approx 0{,}707$ pour $n=3$. La covariance est positive pour $q < q_0$ et négative pour $q > q_0$. Ces calculs ont été vérifiés avec `sympy` pour $n \leqslant 6$ ✓.

(c) Pour $q = q_0$, les variables X et Y sont décorrélées, et la question (a) montre qu'elles ne sont pas indépendantes : une covariance nulle n'entraîne pas l'indépendance, même pour un couple où Y est égale à X dès que $X \geqslant 1$. Ici, la dépendance est même très forte : sachant $X = j \geqslant 1$, la valeur de Y est connue.

EXERCICE 88 ⋆⋆

La loi de E est uniforme sur $\mathcal{P}(\llbracket 1\,,n\rrbracket)$, qui a 2^n éléments (proposition 3) : $\mathrm{P}(E=S)=\dfrac{1}{2^n}$ pour toute partie S.

1. Les parties qui contiennent i s'écrivent $\{i\}\cup S'$, avec $S'\subset\llbracket 1\,,n\rrbracket\setminus\{i\}$; il y en a 2^{n-1}, donc $\mathrm{P}(A_i)=\dfrac{1}{2}$ et $\boxed{\mathbf{1}_{A_i}\sim\mathcal{B}(\frac{1}{2})}$. Soit $(\varepsilon_1,\dots,\varepsilon_n)\in\{0,1\}^n$. L'application $S\mapsto\mathbf{1}_S$ est une bijection de $\mathcal{P}(\llbracket 1\,,n\rrbracket)$ sur $\{0,1\}^n$ (démonstration de la proposition 3) : il existe une unique partie S telle que $\mathbf{1}_S(i)=\varepsilon_i$ pour tout i, et

$$\mathrm{P}(\mathbf{1}_{A_1}=\varepsilon_1,\dots,\mathbf{1}_{A_n}=\varepsilon_n)=\mathrm{P}(E=S)=\frac{1}{2^n}=\prod_{i=1}^{n}\mathrm{P}(\mathbf{1}_{A_i}=\varepsilon_i).$$

Par la proposition 15, $\boxed{\text{les indicatrices } \mathbf{1}_{A_1},\dots,\mathbf{1}_{A_n} \text{ sont mutuellement indépendantes}}$. Choisir une partie au hasard revient à décider, pour chaque élément, à pile ou face, s'il en fait partie.

2. On a $N=\sum_{i=1}^{n}\mathbf{1}_{A_i}$. Par le théorème 8 et les propositions 19 et 21,

$$\boxed{N\sim\mathcal{B}(n,\tfrac{1}{2}),\qquad \mathrm{E}(N)=\frac{n}{2},\qquad \mathrm{V}(N)=\frac{n}{4}}$$

On retrouve que $\mathcal{P}(\llbracket 1\,,n\rrbracket)$ contient $\binom{n}{k}$ parties à k éléments.

3. Pour toute issue, $T=\sum_{i=1}^{n} i\,\mathbf{1}_{A_i}$, y compris si E est vide. Par linéarité, puis par le théorème 13, les variables $i\,\mathbf{1}_{A_i}$ étant mutuellement indépendantes (proposition 16), et par le théorème 11 (ii),

$$\boxed{\mathrm{E}(T)=\frac{1}{2}\sum_{i=1}^{n} i=\frac{n(n+1)}{4}},\qquad \boxed{\mathrm{V}(T)=\sum_{i=1}^{n} i^2\,\frac{1}{4}=\frac{n(n+1)(2n+1)}{24}}$$

Ces valeurs ont été vérifiées par énumération des 2^n parties pour $n\leqslant 6$ ✓. L'espérance est la moitié de la somme totale $\dfrac{n(n+1)}{2}$, ce que l'on voit aussi par symétrie : $S\mapsto\llbracket 1\,,n\rrbracket\setminus S$ est une bijection qui échange T et $\dfrac{n(n+1)}{2}-T$.

EXERCICE 89 ⋆⋆

On numérote les boules de 1 à n, les blanches portant les numéros 1 à p, et l'on prend pour univers l'ensemble des parties à s éléments de $\llbracket 1\,,n\rrbracket$, muni de la probabilité uniforme : il a $\binom{n}{s}$ éléments.

1. Une partie qui contient exactement k boules blanches est formée de k blanches choisies parmi p et de $s-k$ noires choisies parmi $n-p$. Par le lemme des bergers et le théorème 2,

$$\boxed{\mathrm{P}(X=k)=\frac{\binom{p}{k}\binom{n-p}{s-k}}{\binom{n}{s}},\qquad k\in\llbracket 0\,,s\rrbracket}$$

avec la convention $\binom{a}{j}=0$ si $j>a$; la probabilité est non nulle exactement lorsque $k\leqslant\min(p,s)$ et $k\geqslant\max(0,s-n+p)$. La somme de ces probabilités vaut 1 : c'est encore la formule de Vandermonde.

2. On a $X=\sum_{k=1}^{p}\mathbf{1}_{A_k}$. Une partie favorable à A_k contient la boule k et $s-1$ des $n-1$ autres, donc $\mathrm{P}(A_k)=\binom{n-1}{s-1}/\binom{n}{s}=\dfrac{s}{n}$, et

$$\boxed{\mathrm{E}(X)=\frac{ps}{n}}$$

Pour $k \neq \ell$, de même, $\mathrm{P}(A_k \cap A_\ell) = \binom{n-2}{s-2}/\binom{n}{s} = \dfrac{s(s-1)}{n(n-1)}$, ce qui reste vrai pour $s = 1$, et

$$\mathrm{Cov}(\mathbf{1}_{A_k}, \mathbf{1}_{A_\ell}) = \frac{s(s-1)}{n(n-1)} - \frac{s^2}{n^2} = -\frac{s(n-s)}{n^2(n-1)}.$$

Par le théorème 13, avec p variances $\dfrac{s}{n}\Big(1 - \dfrac{s}{n}\Big)$ et $p(p-1)$ covariances,

$$\mathrm{V}(X) = p\,\frac{s(n-s)}{n^2} - p(p-1)\frac{s(n-s)}{n^2(n-1)} = \frac{ps(n-s)}{n^2} \cdot \frac{n-p}{n-1}, \qquad \boxed{\mathrm{V}(X) = s\,\frac{p}{n}\Big(1 - \frac{p}{n}\Big)\frac{n-s}{n-1}}$$

Ces formules ont été vérifiées par énumération sur cinq urnes ✓. Les covariances sont négatives : tirer une boule blanche donnée rend les autres un peu moins probables, puisque le nombre de places est fixé.

3. Pour la loi $\mathcal{B}(s, \dfrac{p}{n})$, qui est celle du nombre de boules blanches en s tirages *avec remise* (théorème 8), l'espérance vaut $s\dfrac{p}{n}$ et la variance $s\dfrac{p}{n}\Big(1 - \dfrac{p}{n}\Big)$ (propositions 19 et 21). Les espérances sont égales, et

$$\boxed{\mathrm{V}(X) = \frac{n-s}{n-1}\,\mathrm{V}(B), \qquad B \sim \mathcal{B}(s, \tfrac{p}{n})}$$

Le facteur $\dfrac{n-s}{n-1}$ est au plus égal à 1 : un tirage sans remise est moins dispersé qu'un tirage avec remise, et il ne l'est plus du tout pour $s = n$, où l'on tire toute l'urne et $X = p$. Lorsque n et p tendent vers $+\infty$, avec $\dfrac{p}{n} = \theta$ et s fixés, ce facteur tend vers 1. La loi elle-même converge : pour $k \in [\![0\,, s]\!]$, en développant les coefficients binomiaux,

$$\mathrm{P}(X = k) = \binom{s}{k} \frac{p(p-1)\cdots(p-k+1)\,(n-p)\cdots(n-p-s+k+1)}{n(n-1)\cdots(n-s+1)} \xrightarrow[n\to+\infty]{} \binom{s}{k}\theta^k(1-\theta)^{s-k},$$

chacun des k premiers facteurs, divisé par le facteur correspondant du dénominateur, tendant vers θ, et chacun des $s-k$ suivants vers $1-\theta$. Dans une urne très grande, tirer sans remise ou avec remise revient presque au même : c'est ce qui justifie l'usage de la loi binomiale dans les sondages.

EXERCICE 90 ⋆⋆

L'univers est l'ensemble des parties à r éléments de $[\![1\,, n]\!]$, muni de la probabilité uniforme. Pour le paquet de cartes, chaque ensemble de r premières cartes correspond à $r!\,(n-r)!$ ordres du paquet, le même nombre pour tous (lemme des bergers) : l'ensemble des r premières cartes suit bien la loi uniforme sur les parties à r éléments, et les deux descriptions donnent la même loi pour X.

1. La variable X prend ses valeurs dans $[\![r\,, n]\!]$. Pour $k \in [\![r\,, n]\!]$, une partie Y vérifie $\max Y = k$ si et seulement si $k \in Y$ et $Y \setminus \{k\}$ est une partie à $r-1$ éléments de $[\![1\,, k-1]\!]$; l'application $Y \mapsto Y \setminus \{k\}$ est une bijection entre ces parties et celles à $r-1$ éléments de $[\![1\,, k-1]\!]$:

$$\boxed{\mathrm{P}(X = k) = \frac{\binom{k-1}{r-1}}{\binom{n}{r}}, \qquad k \in [\![r\,, n]\!]}$$

La somme de ces probabilités vaut 1 (proposition 11), d'où l'identité

$$\boxed{\sum_{k=r}^{n} \binom{k-1}{r-1} = \binom{n}{r}}$$

C'est la formule de la crosse de hockey de l'exercice 35 de la feuille du chapitre 3, $\displaystyle\sum_{k=p}^{N} \binom{k}{p} = \binom{N+1}{p+1}$, avec $p = r-1$ et $N = n-1$. Le chapitre 3 l'obtenait par télescopage de la formule de Pascal (proposition 10) ; elle est ici démontrée en comptant les parties à r éléments selon leur plus grand élément.

2. Avec l'égalité suggérée, qui est la formule du pion, puis la crosse de hockey avec $p = r$ et $N = n$,

$$\mathrm{E}(X) = \frac{1}{\binom{n}{r}} \sum_{k=r}^{n} k\binom{k-1}{r-1} = \frac{r}{\binom{n}{r}} \sum_{k=r}^{n} \binom{k}{r} = r\,\frac{\binom{n+1}{r+1}}{\binom{n}{r}}, \qquad \boxed{\mathrm{E}(X) = \frac{r(n+1)}{r+1}}$$

Pour $r = 1$, on retrouve l'espérance $\dfrac{n+1}{2}$ de la loi uniforme, et pour $r = n$, $\mathrm{E}(X) = n$. Tout se passe comme si les r nombres choisis découpaient $[\![0\,,n+1]\!]$ en $r+1$ intervalles de même longueur moyenne $\dfrac{n+1}{r+1}$: le maximum se trouve en moyenne à une telle longueur de $n+1$, puisque $\mathrm{E}(X) = (n+1) - \dfrac{n+1}{r+1}$.

3. On calcule $\mathrm{E}\big(X(X+1)\big)$. Par la formule du pion, appliquée deux fois,

$$k(k+1)\binom{k-1}{r-1} = r(k+1)\binom{k}{r} = r(r+1)\binom{k+1}{r+1},$$

et par la crosse de hockey,

$$\mathrm{E}\big(X(X+1)\big) = \frac{r(r+1)}{\binom{n}{r}} \sum_{k=r}^{n} \binom{k+1}{r+1} = r(r+1)\frac{\binom{n+2}{r+2}}{\binom{n}{r}} = \frac{r(n+1)(n+2)}{r+2}.$$

Par la formule de König-Huygens, $\mathrm{V}(X) = \mathrm{E}\big(X(X+1)\big) - \mathrm{E}(X) - \mathrm{E}(X)^2$, et une réduction au même dénominateur donne

$$\boxed{\mathrm{V}(X) = \frac{r(n-r)(n+1)}{(r+1)^2(r+2)}}$$

Pour $r = 1$, on retrouve la variance $\dfrac{n^2-1}{12}$ de la loi uniforme (proposition 21), et pour $r = n$, $\mathrm{V}(X) = 0$. La loi, l'espérance et la variance ont été vérifiées par énumération pour $n \leqslant 8$ et tous les r, ainsi que l'équivalence avec la version des cartes pour $n \leqslant 5$ ✓.

À retenir. Pour une variable dont la loi s'écrit avec $\binom{k-1}{r-1}$, on calcule les espérances de X, de $X(X+1)$, …, dont les coefficients se simplifient par la formule du pion, plutôt que $\mathrm{E}(X^2)$; c'est la démarche de la proposition 21 pour la loi binomiale, avec $\mathrm{E}\big(X(X-1)\big)$.

EXERCICE 91 ⋆⋆

On numérote les boules, et l'on prend pour univers l'ensemble des $(2n)!$ ordres de sortie, muni de la probabilité uniforme. L'ensemble R des rangs de sortie des boules rouges est une partie à n éléments de $[\![1\,,2n]\!]$; chacune correspond à $n!\,n!$ ordres, donc R suit la loi uniforme sur les parties à n éléments de $[\![1\,,2n]\!]$.

1. La dernière boule rouge sort au rang $\max R$, et il reste alors $X = 2n - \max R$ boules noires. Pour $k \in [\![0\,,n]\!]$, $(X = k)$ est réalisé si et seulement si la boule de rang $2n-k$ est rouge et les $n-1$ autres rouges ont des rangs dans $[\![1\,,2n-k-1]\!]$, d'où

$$\boxed{\mathrm{P}(X = k) = \frac{\binom{2n-k-1}{n-1}}{\binom{2n}{n}}, \qquad k \in [\![0\,,n]\!]}$$

C'est l'exercice 90 avec $r = n$ et $2n$ au lieu de n, appliqué à $\max R = 2n - X$.

2. Pour chaque boule noire b, soit B_b l'événement « b sort après toutes les boules rouges ». Alors $X = \sum_b \mathbf{1}_{B_b}$, somme de n indicatrices. Considérons les $n+1$ boules formées des rouges et de b. Pour une boule rouge ρ, l'échange des rangs de b et de ρ est une bijection de l'univers, qui préserve la probabilité uniforme et transforme B_b en l'événement « ρ sort après les autres boules de ce groupe ». Ces $n+1$ événements, un par boule du groupe, forment un système complet et ont la même probabilité, comme dans l'exemple des records de la section 21.6.3 :

$$\mathrm{P}(B_b) = \frac{1}{n+1}, \qquad \boxed{\mathrm{E}(X) = \frac{n}{n+1}}$$

Il reste en moyenne moins d'une boule noire dans l'urne, quel que soit n. On retrouve ce résultat avec l'exercice 90 : $\mathrm{E}(X) = 2n - \dfrac{n(2n+1)}{n+1} = \dfrac{n}{n+1}$ ✓.

3. Soit $b \neq b'$ deux boules noires. L'événement $B_b \cap B_{b'}$ est réalisé si et seulement si, parmi les $n+2$ boules formées des rouges, de b et de b', les deux dernières à sortir sont b et b'. Le même argument d'échange montre que la paire des deux dernières boules de ce groupe est uniforme parmi les $\binom{n+2}{2}$ paires possibles : $\mathrm{P}(B_b \cap B_{b'}) = \dfrac{2}{(n+1)(n+2)}$, et

$$\mathrm{Cov}(\mathbf{1}_{B_b}, \mathbf{1}_{B_{b'}}) = \frac{2}{(n+1)(n+2)} - \frac{1}{(n+1)^2} = \frac{n}{(n+1)^2(n+2)}.$$

Par le théorème 13, avec n variances $\dfrac{1}{n+1} \cdot \dfrac{n}{n+1}$ et $n(n-1)$ covariances,

$$\mathrm{V}(X) = \frac{n^2}{(n+1)^2} + \frac{n^2(n-1)}{(n+1)^2(n+2)}, \qquad \boxed{\mathrm{V}(X) = \frac{n^2(2n+1)}{(n+1)^2(n+2)}}$$

ce qui coïncide avec la variance de l'exercice 90 pour $r = n$ et $2n$ boules ✓. La loi, l'espérance et la variance ont été vérifiées par énumération pour $n \leqslant 5$, et $\mathrm{P}(B_b \cap B_{b'})$ sur les 720 ordres de sortie pour $n = 3$. Quand n tend vers $+\infty$, $\mathrm{E}(X) \to 1$ et $\mathrm{V}(X) \to 2$.

Piège. Les événements B_b ne sont pas indépendants, et leurs covariances sont positives : si une boule noire sort après toutes les rouges, c'est que la dernière rouge est sortie tôt, ce qui favorise les autres noires. L'erreur la plus répandue est d'écrire $\mathrm{V}(X) = n\dfrac{1}{n+1}(1-\dfrac{1}{n+1})$, comme pour une loi binomiale.

EXERCICE 92 ⋆⋆⋆

1. On numérote les boules. Les $N!$ ordres de sortie sont équiprobables ; l'ensemble des rangs des deux boules noires, partie à 2 éléments de $[\![1\,,N]\!]$, correspond à $2!\,(N-2)!$ ordres, le même nombre pour toutes les parties. On peut donc prendre pour univers

$$\boxed{\Omega = \{\{i,j\} \mid 1 \leqslant i < j \leqslant N\},\ \text{muni de la probabilité uniforme},\ |\Omega| = \binom{N}{2}}$$

avec $X_1(\{i,j\}) = i$ et $X_2(\{i,j\}) = j$ lorsque $i < j$.

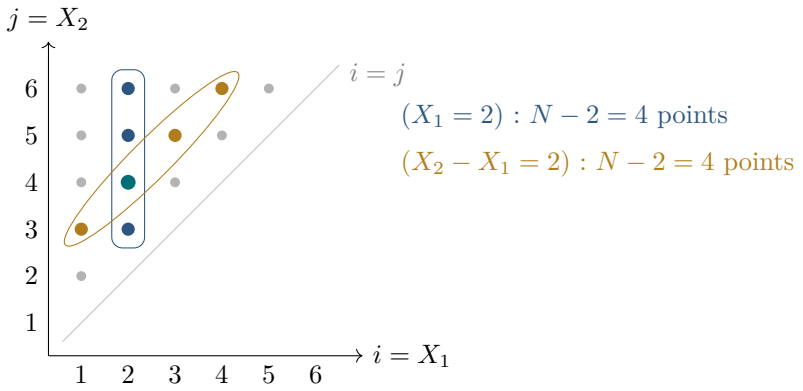
2. Pour $i < j$, l'événement $(X_1 = i, X_2 = j)$ est le singleton $\{\{i,j\}\}$, et pour $i \geqslant j$ il est vide :

$$\boxed{\mathrm{P}(X_1 = i, X_2 = j) = \frac{2}{N(N-1)} \ \text{ si } i < j, \qquad 0 \text{ sinon}}$$

Par la proposition 14, en comptant les $N - i$ valeurs de $j > i$ et les $j - 1$ valeurs de $i < j$,

$$\boxed{\mathrm{P}(X_1 = i) = \frac{2(N-i)}{N(N-1)}, \qquad \mathrm{P}(X_2 = j) = \frac{2(j-1)}{N(N-1)}, \qquad i,j \in [\![1\,,N]\!]}$$

Si $N \geqslant 3$, les variables ne sont pas indépendantes : $\mathrm{P}(X_1 = 2, X_2 = 2) = 0$, alors que le produit $\mathrm{P}(X_1 = 2)\mathrm{P}(X_2 = 2)$ vaut $\dfrac{4(N-2)}{N^2(N-1)^2} > 0$. Si $N = 2$, en revanche, $X_1 = 1$ et $X_2 = 2$ sont constantes, donc indépendantes.



Pour $N = 6$, les 15 issues sont les points (i, j) situés au-dessus de la diagonale, tous de même probabilité. La colonne $X_1 = 2$, encadrée en bleu, et la diagonale $X_2 - X_1 = 2$, en ocre, contiennent chacune $N - 2 = 4$ points : c'est la question 3 (b). Le point $(2, 4)$, commun aux deux, est en bleu pétrole.

3. (a) Pour $i \in [\![1\,, N]\!]$, $\mathrm{P}(N + 1 - X_2 = i) = \mathrm{P}(X_2 = N + 1 - i) = \dfrac{2(N-i)}{N(N-1)} = \mathrm{P}(X_1 = i)$: $\boxed{N + 1 - X_2 \sim X_1}$. C'est le tirage lu à l'envers : le rang de la dernière noire, compté depuis la fin, a la loi du rang de la première.

(b) La variable $X_2 - X_1$ prend ses valeurs dans $[\![1\,, N-1]\!]$. Pour d dans cet intervalle, les couples $(i, i + d)$ avec $1 \leqslant i$ et $i + d \leqslant N$ sont au nombre de $N - d$:

$$\boxed{\mathrm{P}(X_2 - X_1 = d) = \frac{2(N-d)}{N(N-1)} = \mathrm{P}(X_1 = d)}, \qquad X_2 - X_1 \sim X_1.$$

(c) Les trois variables X_1, $X_2 - X_1$ et $N + 1 - X_2$ ont la même loi, donc la même espérance m, et leur somme vaut $N + 1$. Par linéarité, $3m = N + 1$:

$$\boxed{\mathrm{E}(X_1) = \frac{N+1}{3}}, \qquad \mathrm{E}(X_2) = N + 1 - \mathrm{E}(N + 1 - X_2) = N + 1 - m, \qquad \boxed{\mathrm{E}(X_2) = \frac{2(N+1)}{3}}$$

Les deux boules noires partagent en moyenne la suite des tirages en trois parts égales, comme à l'exercice 83.

(d) Par la question (a) et le théorème 11 (ii), $\mathrm{V}(X_1) = \mathrm{V}(N + 1 - X_2) = \mathrm{V}(X_2)$. Par la question (b), $\mathrm{V}(X_2 - X_1) = \mathrm{V}(X_1)$, et par bilinéarité de la covariance,

$$\mathrm{V}(X_1) = \mathrm{V}(X_2 - X_1) = \mathrm{V}(X_2) + \mathrm{V}(X_1) - 2\,\mathrm{Cov}(X_1, X_2) = 2\mathrm{V}(X_1) - 2\,\mathrm{Cov}(X_1, X_2),$$

d'où $\boxed{\mathrm{V}(X_1) = \mathrm{V}(X_2)}$ et $\boxed{2\,\mathrm{Cov}(X_1, X_2) = \mathrm{V}(X_1)}$.

4. Par la formule de transfert, avec les sommes de i^2 et de i^3 pour $i \in [\![1\,, N-1]\!]$,

$$\mathrm{E}(X_1^2) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} (Ni^2 - i^3) = \frac{2}{N(N-1)} \Big(\frac{N^2(N-1)(2N-1)}{6} - \frac{N^2(N-1)^2}{4} \Big) = \frac{N(N+1)}{6}.$$

Par la formule de König-Huygens, puis la question 3 (d),

$$\mathrm{V}(X_1) = \frac{N(N+1)}{6} - \frac{(N+1)^2}{9}, \qquad \boxed{\mathrm{V}(X_1) = \frac{(N+1)(N-2)}{18}}$$

et

$$\boxed{\mathrm{Cov}(X_1, X_2) = \frac{(N+1)(N-2)}{36}}$$

Pour $N = 2$, on retrouve une variance et une covariance nulles. La covariance est positive : quand la première noire sort tard, la seconde aussi. Le coefficient de corrélation vaut $\dfrac{1}{2}$ pour tout $N \geqslant 3$. Toutes ces valeurs ont été vérifiées par énumération pour $N \leqslant 8$ ✓.

5. On a $\mathrm{P}(X=Y)=\sum_{k=1}^{N}\mathrm{P}(X=k)\mathrm{P}(Y=k)=\dfrac{N}{N^2}$, donc $\boxed{\mathrm{P}(D)=1-\dfrac{1}{N}=\dfrac{N-1}{N}}$. Pour $i<j$, l'événement $(Z_1=i,Z_2=j)\cap D$ est la réunion des événements incompatibles $(X=i,Y=j)$ et $(X=j,Y=i)$, de probabilité $\dfrac{2}{N^2}$; pour $i\geqslant j$, il est vide, puisque $Z_1<Z_2$ sur D. Ainsi

$$\mathrm{P}_D(Z_1=i,Z_2=j)=\frac{2/N^2}{(N-1)/N},\qquad \boxed{\mathrm{P}_D(Z_1=i,Z_2=j)=\frac{2}{N(N-1)}\ \text{ si } i<j,\quad 0 \text{ sinon}}$$

Sachant D, le couple (Z_1,Z_2) a la loi du couple (X_1,X_2) de la question 2. Deux tirages indépendants uniformes, conditionnés à être distincts et rangés dans l'ordre, se comportent exactement comme les positions des deux boules noires.

À retenir. La question 3 évite tout calcul de somme : les symétries du problème, lecture à l'envers et translation des écarts, donnent l'égalité des lois, puis la linéarité de l'espérance et la bilinéarité de la covariance font le reste. Le seul calcul véritable est celui de $\mathrm{E}(X_1^2)$.

EXERCICE 93 ⋆⋆⋆

Les n^2 variables sont centrées : $\mathrm{E}(X_{i,j})=0$, et $X_{i,j}^2=1$, donc $\mathrm{V}(X_{i,j})=1$.

1. On a $\mathrm{tr}(M)=\sum_{i=1}^{n}X_{i,i}$. Les variables $B_i=\dfrac{X_{i,i}+1}{2}$ sont mutuellement indépendantes (proposition 16), de loi $\mathcal{B}(\frac{1}{2})$, et $\dfrac{\mathrm{tr}(M)+n}{2}=\sum_i B_i$ suit la loi $\mathcal{B}(n,\frac{1}{2})$ (théorème 8). D'où

$$\boxed{\mathrm{P}(\,\mathrm{tr}(M)=2k-n)=\binom{n}{k}\frac{1}{2^n},\qquad k\in[\![0\,,n]\!]}$$

Par linéarité et par le théorème 13, les $X_{i,i}$ étant indépendantes,

$$\boxed{\mathrm{E}(\,\mathrm{tr}(M))=0,\qquad \mathrm{V}(\,\mathrm{tr}(M))=n}$$

C'est la marche aléatoire de la section 21.5.7 : la trace est la position d'un marcheur après n pas.

2. Comme tous les coefficients sont non nuls, $M\neq 0$ et $\mathrm{rg}(M)\geqslant 1$. La matrice est de rang 1 si et seulement si toutes ses colonnes sont proportionnelles à la première, C_1, qui est non nulle (chapitre 16). Une colonne C_j à coefficients ±1 s'écrit λC_1 si et seulement si $\lambda=\pm1$, puisque les coefficients de λC_1 sont $\pm\lambda$. Ainsi $\mathrm{rg}(M)=1$ si et seulement si $C_j\in\{C_1,-C_1\}$ pour tout $j\in[\![2\,,n]\!]$. La matrice M suit la loi uniforme sur les 2^{n^2} matrices à coefficients ±1, par indépendance, et il y a 2^n choix de C_1 et 2 choix pour chacune des $n-1$ autres colonnes :

$$\mathrm{P}(\mathrm{rg}(M)=1)=\frac{2^n\times 2^{n-1}}{2^{n^2}},\qquad \boxed{\mathrm{P}(\mathrm{rg}(M)=1)=\frac{1}{2^{(n-1)^2}}}$$

Pour $n=1$, on trouve 1, et pour $n=2$, $\dfrac{1}{2}$: une matrice 2×2 à coefficients ±1 est inversible une fois sur deux.

3. Par la définition du déterminant (chapitre 17),

$$\det(M)=\sum_{\sigma\in\mathfrak{S}_n}\varepsilon(\sigma)\,T_\sigma,\qquad T_\sigma=\prod_{j=1}^{n}X_{\sigma(j),j}.$$

Le produit T_σ porte sur n coefficients de positions distinctes $(\sigma(j),j)$, donc mutuellement indépendants : par la proposition 20, $\mathrm{E}(T_\sigma)=0$, et, par linéarité, $\boxed{\mathrm{E}(\det(M))=0}$.

Soit $\sigma\neq\tau$. Dans le produit $T_\sigma T_\tau$, le facteur de la colonne j vaut $X_{\sigma(j),j}^2=1$ si $\sigma(j)=\tau(j)$, et $X_{\sigma(j),j}X_{\tau(j),j}$ sinon. Ainsi $T_\sigma T_\tau$ est le produit des coefficients situés aux positions $(\sigma(j),j)$ et

$(\tau(j), j)$, pour les j tels que $\sigma(j) \neq \tau(j)$; ces positions sont deux à deux distinctes, et il y en a au moins deux, puisque $\sigma \neq \tau$. Par la proposition 20, $\mathrm{E}(T_\sigma T_\tau) = 0$, donc $\mathrm{Cov}(T_\sigma, T_\tau) = 0$, alors que $\mathrm{V}(T_\sigma) = \mathrm{E}(T_\sigma^2) = 1$. Les $n!$ variables $\varepsilon(\sigma)T_\sigma$ sont deux à deux décorrélées, et le théorème 13 donne

$$\boxed{\mathrm{V}\big(\det(M)\big) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)^2 \mathrm{V}(T_\sigma) = n!}$$

Pour $n = 2$, $\det(M)$ vaut 0 avec la probabilité $\dfrac{1}{2}$, et ± 2 avec la probabilité $\dfrac{1}{4}$ chacun : la variance vaut bien 2. Pour $n = 3$, l'énumération des 512 matrices donne $\mathrm{P}(\det M = 0) = \dfrac{5}{8}$ et $\mathrm{P}(\det M = \pm 4) = \dfrac{3}{16}$ chacun, de variance 6 ✓.

Piège. Les variables T_σ ne sont pas indépendantes : deux d'entre elles partagent en général des coefficients. Le théorème 13 n'a besoin que de leur décorrélation deux à deux, qui se vérifie par la proposition 20 appliquée aux coefficients, eux indépendants. L'erreur la plus répandue est d'invoquer une indépendance qui n'existe pas.

À retenir. L'écart type de $\det(M)$ vaut $\sqrt{n!}$: le déterminant d'une matrice aléatoire à coefficients ± 1 est en moyenne nul, mais il est typiquement de l'ordre de $\sqrt{n!}$, bien plus petit que $n!$, nombre de termes de la somme, en raison des compensations entre termes de signes opposés.

EXERCICE 94 ⋆⋆⋆

1. Par définition du produit matriciel, puis par la bilinéarité de la covariance (proposition 23 (ii)), appliquée à chaque argument,

$$\Lambda^T C \Lambda = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \,\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}\Big(\sum_{i=1}^n \lambda_i X_i, \sum_{j=1}^n \lambda_j X_j\Big), \qquad \boxed{\Lambda^T C \Lambda = \mathrm{V}(\lambda_1 X_1 + \cdots + \lambda_n X_n)}$$

par la proposition 23 (i). En particulier $\Lambda^T C \Lambda \geqslant 0$ pour toute colonne Λ : la matrice symétrique C est positive.

2. Pour $\Lambda \in \mathbb{R}^n$, on note $Z_\Lambda = \lambda_1 X_1 + \cdots + \lambda_n X_n$. Par bilinéarité, la i-ième coordonnée de $C\Lambda$ est

$$(C\Lambda)_i = \sum_{j=1}^n \mathrm{Cov}(X_i, X_j)\lambda_j = \mathrm{Cov}(X_i, Z_\Lambda).$$

Si C n'est pas inversible, il existe une colonne $\Lambda \neq 0$ telle que $C\Lambda = 0$ (chapitre 16). Alors $\mathrm{V}(Z_\Lambda) = \Lambda^T C \Lambda = 0$, et le théorème 11 (iii) donne $\mathrm{P}\big(Z_\Lambda = \mathrm{E}(Z_\Lambda)\big) = 1$: il existe des réels non tous nuls et un réel $c = \mathrm{E}(Z_\Lambda)$ tels que $\mathrm{P}(\lambda_1 X_1 + \cdots + \lambda_n X_n = c) = 1$.

Réciproquement, supposons $\Lambda \neq 0$ et $\mathrm{P}(Z_\Lambda = c) = 1$. Alors $Z_\Lambda(\omega) = c$ pour toute issue ω de probabilité non nulle, donc $\mathrm{E}(Z_\Lambda) = c$ (proposition 18), et, pour tout i, la proposition 18 donne

$$(C\Lambda)_i = \mathrm{Cov}(X_i, Z_\Lambda) = \sum_{\omega \in \Omega} \big(X_i(\omega) - \mathrm{E}(X_i)\big)\big(Z_\Lambda(\omega) - c\big)\mathrm{P}(\{\omega\}) = 0,$$

chaque terme étant nul. Ainsi $C\Lambda = 0$ avec $\Lambda \neq 0$, et C n'est pas inversible. Par contraposée,

$$\boxed{\begin{array}{c} C \text{ inversible} \iff \text{aucune combinaison } \lambda_1 X_1 + \cdots + \lambda_n X_n, \text{ à coefficients non tous nuls,} \\ \text{n'est constante avec probabilité 1} \end{array}}$$

Par exemple, les nombres de boules $X_1, \ldots, X_n$ des tiroirs de l'exercice 80 ont une somme constante égale à r : leur matrice de covariance n'est pas inversible, et la colonne $(1, \ldots, 1)^T$ est dans son noyau. Pour $n = 3$ et $r = 2$, on trouve

$$C = \frac{2}{9}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \qquad \det C = 0 \text{ ✓}.$$

Comparaison avec le chapitre 19. La matrice de Gram d'une famille $(x_1, \dots, x_n)$ de vecteurs d'un espace euclidien de dimension n est inversible si et seulement si la famille est libre (chapitre 19, proposition 9 (ii)). Ici, C est la matrice des $\mathrm{E}(U_iU_j)$, où $U_i = X_i - \mathrm{E}(X_i)$ sont les variables centrées : c'est une « matrice de Gram » pour la forme bilinéaire symétrique $(U, V) \mapsto \mathrm{E}(UV)$ sur l'espace $\mathbb{R}^\Omega$. Cette forme est positive, mais elle n'est pas définie dès qu'une issue a une probabilité nulle, ce qui explique qu'on ne puisse conclure qu'à une égalité avec probabilité 1. Si toutes les issues ont une probabilité non nulle, c'est un produit scalaire, et le résultat dit exactement que C est inversible si et seulement si la famille $(U_1, \dots, U_n)$ des variables centrées est libre.
À retenir. Une matrice de covariance est toujours symétrique et positive ; elle est inversible si et seulement si les variables ne sont liées par aucune relation affine certaine. C'est la même dichotomie qu'au chapitre 19, avec la variance à la place du carré de la norme.

Inégalités probabilistes et loi faible des grands nombres

EXERCICE 95 ⋆

1. Par la proposition 19, X_n a pour espérance $4n \times \dfrac{1}{2} = 2n$, et par la proposition 21, sa variance vaut

$$\boxed{\mathrm{V}(X_n) = 4n \times \frac{1}{2} \times \frac{1}{2} = n}$$

2. On applique l'inégalité de Bienaymé-Tchebychev (théorème 15) à X_n, avec $\varepsilon = n > 0$:

$$\mathrm{P}(|X_n - 2n| \geqslant n) \leqslant \frac{\mathrm{V}(X_n)}{n^2} = \frac{1}{n}, \qquad \text{d'où} \qquad \boxed{\mathrm{P}(|X_n - 2n| \geqslant n) \xrightarrow[n\to+\infty]{} 0}$$

L'écart toléré, n, est proportionnel au nombre $4n$ d'épreuves, alors que l'écart type $\sqrt{n}$ n'en est qu'une racine carrée : c'est ce décalage d'échelle que traduit la loi faible des grands nombres.

3. L'événement $(|X_n - 2n| < n)$ est le contraire de l'événement de la question 2, et il est inclus dans $(|X_n - 2n| \leqslant n)$. Par croissance et passage au contraire (proposition 7 (ii) et (iii)),

$$1 \geqslant \mathrm{P}(|X_n - 2n| \leqslant n) \geqslant 1 - \mathrm{P}(|X_n - 2n| \geqslant n) \geqslant 1 - \frac{1}{n},$$

et le théorème des gendarmes donne $\boxed{\mathrm{P}(|X_n - 2n| \leqslant n) \xrightarrow[n\to+\infty]{} 1}$. Or l'événement $(|X_n - 2n| \leqslant n)$ est $(n \leqslant X_n \leqslant 3n)$, et la loi binomiale donne

$$\mathrm{P}(|X_n - 2n| \leqslant n) = \sum_{k=n}^{3n} \binom{4n}{k} \frac{1}{2^k} \frac{1}{2^{4n-k}} = \frac{x_n}{16^n}.$$

Le quotient $\dfrac{x_n}{16^n}$ tend donc vers 1, c'est-à-dire

$$\boxed{x_n = \sum_{k=n}^{3n} \binom{4n}{k} \underset{n\to+\infty}{\sim} 16^n}$$

Pour $n = 1$, on trouve $x_1 = 4 + 6 + 4 = 14$, et $\dfrac{x_1}{16} = 0{,}875$; puis $\dfrac{x_5}{16^5} \approx 0{,}988$ et $\dfrac{x_{20}}{16^{20}} \approx 0{,}999\,997$ ✓. Comme $16^n = 2^{4n}$ est le nombre total de parties d'un ensemble à $4n$ éléments, le résultat dit que presque toutes ces parties ont entre le quart et les trois quarts des éléments.
Piège. Un équivalent se démontre par un quotient qui tend vers 1, pas par une différence qui tend vers 0 : ici $16^n - x_n$ tend vers $+\infty$, et pourtant $x_n \sim 16^n$. L'erreur la plus répandue est d'écrire $x_n - 16^n \to 0$, qui est faux.

EXERCICE 96 ⋆⋆

On prend pour univers $\Omega = \{0,1\}^n$, muni de la probabilité uniforme, et $X_n(\omega) = \omega$. Chaque issue a la probabilité $\dfrac{1}{2^n} = \displaystyle\prod_{i=1}^{n} \frac{1}{2}$: c'est la probabilité produit de la proposition 17, construite avec n fois la distribution uniforme sur $\{0,1\}$. Les coordonnées C_1, …, C_n de X_n sont donc mutuellement indépendantes, de loi $\mathcal{B}(\dfrac{1}{2})$. Comme $C_i^2 = C_i$,

$$\|X_n\|^2 = \sum_{i=1}^{n} C_i^2 = \sum_{i=1}^{n} C_i = N_n, \qquad N_n \sim \mathcal{B}(n, \tfrac{1}{2})$$

par le théorème 8 : le carré de la norme est le nombre de coordonnées égales à 1.

Comparaison des racines carrées. Pour $u \geqslant 0$, on a $\sqrt{u} + \sqrt{\dfrac{1}{2}} \geqslant \dfrac{\sqrt{2}}{2} > 0$, donc

$$\Big|\sqrt{u} - \frac{\sqrt{2}}{2}\Big| = \frac{|u - \frac{1}{2}|}{\sqrt{u} + \sqrt{\frac{1}{2}}} \leqslant \sqrt{2}\,\Big|u - \frac{1}{2}\Big|.$$

Conclusion. Avec $u = \dfrac{N_n}{n}$, on a $\dfrac{\|X_n\|}{\sqrt{n}} = \sqrt{u}$. Soit $\varepsilon > 0$. Par l'inégalité précédente, l'événement étudié est inclus dans $(|\dfrac{N_n}{n} - \dfrac{1}{2}| \geqslant \dfrac{\varepsilon}{\sqrt{2}})$, et la loi faible des grands nombres (théorème 16, cas binomial, avec $p(1-p) = \dfrac{1}{4}$) donne

$$\mathrm{P}\Big(\Big|\frac{\|X_n\|}{\sqrt{n}} - \frac{\sqrt{2}}{2}\Big| \geqslant \varepsilon\Big) \leqslant \mathrm{P}\Big(\Big|\frac{N_n}{n} - \frac{1}{2}\Big| \geqslant \frac{\varepsilon}{\sqrt{2}}\Big) \leqslant \frac{1/4}{n\,\varepsilon^2/2},$$

$$\boxed{\mathrm{P}\Big(\Big|\frac{\|X_n\|}{\sqrt{n}} - \frac{\sqrt{2}}{2}\Big| \geqslant \varepsilon\Big) \leqslant \frac{1}{2n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0}$$

Pour $\varepsilon = 0{,}05$ et $n = 1\,000$, la majoration vaut 0,2, alors que la valeur exacte, calculée avec la loi binomiale, est de l'ordre de 9×10^{-6} ✓.

Interprétation. Les sommets du cube $[0\,,1]^n$ sont les 2^n points de $\{0,1\}^n$, et X_n est un sommet choisi au hasard. Le résultat dit que, pour n grand, la plupart des sommets sont à une distance de l'origine voisine de $\sqrt{\dfrac{n}{2}}$, à un facteur $1 \pm \varepsilon\sqrt{2}$ près : *ils se concentrent près de la sphère de centre O et de rayon* $\sqrt{\dfrac{n}{2}}$, alors que les distances possibles s'étalent de 0 à $\sqrt{n}$. On peut aussi le lire avec les angles du chapitre 19 : le cosinus de l'angle entre un sommet $x \neq 0$ et la grande diagonale $(1,\dots,1)$ vaut $\dfrac{N_n}{\|x\|\sqrt{n}} = \dfrac{\|x\|}{\sqrt{n}}$, qui est proche de $\dfrac{\sqrt{2}}{2}$: vus de l'origine, presque tous les sommets font un angle voisin de $\dfrac{\pi}{4}$ avec la diagonale. Les deux extrémités de la diagonale, 0 et $(1,\dots,1)$, sont des sommets tout à fait atypiques.

EXERCICE 97 ⋆⋆

Les variables $Y_i = X_i - m$ sont mutuellement indépendantes (proposition 16) et de même loi (proposition 13), et centrées par linéarité.

1. Comme $x \mapsto x^4$ est strictement croissante sur $\mathbb{R}_+$, les événements $(|S_n| \geqslant n\varepsilon)$ et $(S_n^4 \geqslant n^4\varepsilon^4)$ sont égaux. L'inégalité de Markov (théorème 14), appliquée à la variable positive S_n^4 avec $a = n^4\varepsilon^4 > 0$, donne

$$\boxed{\mathrm{P}\Big(\Big|\frac{S_n}{n}\Big| \geqslant \varepsilon\Big) \leqslant \frac{\mathrm{E}(S_n^4)}{\varepsilon^4 n^4}}$$

C'est la démonstration de l'inégalité de Bienaymé-Tchebychev, avec la puissance 4 au lieu de la puissance 2.

2. On développe le produit de quatre sommes (chapitre 3), puis on applique la linéarité de l'espérance :

$$S_n^4 = \sum_{(i,j,k,\ell)\in[\![1\,,n]\!]^4} Y_iY_jY_kY_\ell, \qquad \mathrm{E}(S_n^4) = \sum_{(i,j,k,\ell)} \mathrm{E}(Y_iY_jY_kY_\ell).$$

Dans un quadruplet (i,j,k,ℓ), regroupons les indices égaux : si $a_1, \ldots, a_r$ sont les valeurs distinctes, de multiplicités $m_1, \ldots, m_r$, de somme 4, le produit vaut $Y_{a_1}^{m_1} \cdots Y_{a_r}^{m_r}$. Ces r facteurs sont mutuellement indépendants (propositions 15 et 16), et la proposition 20 donne

$$\mathrm{E}(Y_iY_jY_kY_\ell) = \mathrm{E}(Y_{a_1}^{m_1}) \cdots \mathrm{E}(Y_{a_r}^{m_r}).$$

Dès qu'une multiplicité vaut 1, le facteur correspondant est $\mathrm{E}(Y_a) = 0$. Il ne reste que deux types de quadruplets :

— les quatre indices égaux, ce qui donne n quadruplets, de contribution μ_4 chacun ;

— deux valeurs distinctes, chacune deux fois : on choisit la paire de valeurs, de $\binom{n}{2}$ façons, puis les deux positions qui portent la plus petite, de $\binom{4}{2} = 6$ façons ; ces $6\binom{n}{2} = 3n(n-1)$ quadruplets contribuent chacun $\mathrm{E}(Y_a^2)\,\mathrm{E}(Y_b^2) = \sigma^4$.

$$\boxed{\mathrm{E}(S_n^4) = n\mu_4 + 3n(n-1)\sigma^4}$$

La formule a été vérifiée exactement pour $n \leqslant 4$, sur une loi à trois valeurs ✓. On en déduit

$$0 \leqslant \frac{\mathrm{E}(S_n^4)}{n^4} = \frac{\mu_4}{n^3} + \frac{3(n-1)\sigma^4}{n^3} \leqslant \frac{\mu_4 + 3\sigma^4}{n^2},$$

et la série de Riemann $\sum \dfrac{1}{n^2}$ converge (chapitre 20, théorème 5) : $\boxed{\text{la série } \displaystyle\sum_{n\geqslant 1} \frac{\mathrm{E}(S_n^4)}{n^4} \text{ converge}}$ par comparaison.

3. La loi faible des grands nombres (théorème 16) majore la même probabilité par $\dfrac{\sigma^2}{n\varepsilon^2}$, qui est en $O(\dfrac{1}{n})$, alors que la question 1 donne un majorant en $O(\dfrac{1}{n^2})$. Si $\sigma > 0$, le rapport du second au premier est équivalent à $\dfrac{3\sigma^2}{n\varepsilon^2}$, qui tend vers 0 : *pour n grand, la majoration par le moment d'ordre 4 est bien meilleure.* Elle est moins bonne pour n petit. Pour des variables de loi $\mathcal{B}(\dfrac{1}{2})$, on a $\sigma^2 = \dfrac{1}{4}$ et $\mu_4 = \dfrac{1}{16}$; avec $\varepsilon = 0{,}1$, les deux majorants valent

$$\frac{25}{n} \qquad \text{et} \qquad \frac{625(3n-2)}{n^3},$$

et le second devient le meilleur à partir de $n = 75$ ✓ ; pour $n = 1\,000$, ils valent 0,025 et environ 0,0019.

À retenir. Plus on utilise un moment d'ordre élevé dans l'inégalité de Markov, plus la décroissance obtenue est rapide ; la proposition 24 et l'exercice 101 poussent l'idée jusqu'au bout avec l'exponentielle. La convergence de la série de la question 2 est le point de départ de la démonstration de la loi *forte* des grands nombres, qui porte sur une suite infinie de variables et dépasse le cadre des univers finis.

EXERCICE 98 ⋆⋆

1. On note B_i l'indicatrice de l'événement « la personne i se présente », pour $i \in [\![1\,,n]\!]$. L'énoncé dit que B_1, …, B_n sont mutuellement indépendantes, de loi $\mathcal{B}(p)$; leur somme S_n suit donc la loi binomiale (théorème 8) :

$$\boxed{S_n \sim \mathcal{B}(n, 0{,}75)}$$

2. Par les propositions 19 et 21, $\mathrm{E}(S_n) = np$ et $\mathrm{V}(S_n) = np(1-p)$. Si $S_n \geqslant 150$, alors

$$S_n - np \geqslant 150 - np > 0, \qquad \text{donc} \qquad |S_n - np| \geqslant 150 - np :$$

c'est l'inclusion de l'énoncé. Par croissance, puis par l'inégalité de Bienaymé-Tchebychev (théorème 15) avec $\varepsilon = 150 - np > 0$,

$$\boxed{\mathrm{P}(S_n \geqslant 150) \leqslant \mathrm{P}\big(|S_n - np| \geqslant 150 - np\big) \leqslant \frac{np(1-p)}{(150-np)^2}}$$

L'hypothèse $np < 150$ est indispensable : sans elle, ε ne serait pas strictement positif.

3. Sur $]0\,,150[$, on a $(150-x)^2 > 0$ et $1-p = \dfrac{1}{4}$; l'inéquation équivaut à $\dfrac{x}{4} \leqslant \dfrac{(150-x)^2}{20}$, c'est-à-dire à

$$x^2 - 305x + 22\,500 \geqslant 0.$$

Le discriminant vaut $305^2 - 90\,000 = 3\,025 = 55^2$, et les racines sont $\dfrac{305 \pm 55}{2}$, soit 125 et 180. Le trinôme est positif à l'extérieur des racines :

$$\boxed{\text{l'ensemble des solutions dans }]0\,,150[\text{ est }]0\,,125]}$$

4. « Plus de 150 personnes se présentent » est l'événement $(S_n > 150)$, inclus dans $(S_n \geqslant 150)$. Par les questions 2 et 3 avec $x = np = \dfrac{3n}{4}$, la probabilité est au plus 0,05 dès que $\dfrac{3n}{4} \leqslant 125$, c'est-à-dire $n \leqslant \dfrac{500}{3} \approx 166{,}7$. La compagnie peut donc vendre

$$\boxed{n = 166 \text{ billets}}$$

et la majoration vaut alors $\dfrac{124{,}5 \times 0{,}25}{25{,}5^2} \approx 0{,}048$. Le programme suivant calcule la probabilité exacte, puis cherche le plus grand nombre de billets pour lequel elle reste au plus 0,05.

```
from math import comb

def proba_trop(n, p=0.75, places=150):
    # P(S_n > places) pour S_n de loi B(n, p)
    return sum(comb(n, k) * p**k * (1 - p)**(n - k)
               for k in range(places + 1, n + 1))

print(proba_trop(166))          # environ 1.42e-07
n = 151
while proba_trop(n + 1) <= 0.05:
    n += 1
print(n, proba_trop(n), proba_trop(n + 1))
```

On obtient $\mathrm{P}(S_{166} > 150) \approx 1{,}4 \times 10^{-7}$, au lieu du 0,05 garanti, et le plus grand nombre de billets admissible est en réalité 187, avec $\mathrm{P}(S_{187} > 150) \approx 0{,}039$, alors que $\mathrm{P}(S_{188} > 150) \approx 0{,}052$ ✓.

À retenir. L'inégalité de Bienaymé-Tchebychev ne connaît de la loi que son espérance et sa variance : elle est universelle, donc très pessimiste pour une loi particulière. Ici, elle coûterait à la compagnie 21 billets. L'exercice 101 montre qu'une inégalité exponentielle, adaptée aux sommes de variables indépendantes, réduit beaucoup cet écart.

EXERCICE 99 ⋆⋆

On note $V = \mathrm{V}(X)$ et $Y = X - \mathrm{E}(X)$, variable centrée de variance V.

1. Soit $t \geqslant 0$. Comme $a + t > 0$, si $Y \geqslant a$, alors $Y + t \geqslant a + t > 0$, et donc $(Y+t)^2 \geqslant (a+t)^2$. Par croissance, puis par l'inégalité de Markov (théorème 14) appliquée à la variable positive $(Y+t)^2$,

$$\mathrm{P}(Y \geqslant a) \leqslant \mathrm{P}\big((Y+t)^2 \geqslant (a+t)^2\big) \leqslant \frac{\mathrm{E}\big((Y+t)^2\big)}{(a+t)^2}.$$

Par linéarité, $\mathrm{E}\big((Y+t)^2\big) = \mathrm{E}(Y^2) + 2t\,\mathrm{E}(Y) + t^2 = V + t^2$, d'où

$$\boxed{\mathrm{P}\big(X - \mathrm{E}(X) \geqslant a\big) \leqslant \frac{t^2+V}{(a+t)^2} \quad \text{pour tout } t \geqslant 0}$$

2. On minimise $g(t) = \dfrac{t^2+V}{(a+t)^2}$ sur $\mathbb{R}_+$. Le calcul de la dérivée donne

$$g'(t) = \frac{2t(a+t) - 2(t^2+V)}{(a+t)^3} = \frac{2(at - V)}{(a+t)^3},$$

du signe de $at - V$: g décroît sur $\left[0, \dfrac{V}{a}\right]$ et croît ensuite, et son minimum est atteint en $t = \dfrac{V}{a}$. En ce point,

$$g\Big(\frac{V}{a}\Big) = \frac{\frac{V^2}{a^2} + V}{\big(a + \frac{V}{a}\big)^2} = \frac{V(V+a^2)/a^2}{(a^2+V)^2/a^2}, \qquad \text{d'où} \qquad \boxed{\mathrm{P}\big(X - \mathrm{E}(X) \geqslant a\big) \leqslant \frac{V}{V+a^2}}$$

Cette inégalité *unilatérale* est optimale : la variable qui vaut a avec la probabilité $\dfrac{V}{V+a^2}$ et $-\dfrac{V}{a}$ avec la probabilité $\dfrac{a^2}{V+a^2}$ est centrée, de variance $\dfrac{a^2V + V^2}{V+a^2} = V$, et réalise l'égalité ✓.

3. La variable $-X$ a pour variance V (théorème 11 (ii)), et la question 2, appliquée à $-X$, donne

$$\mathrm{P}\big(X - \mathrm{E}(X) \leqslant -a\big) = \mathrm{P}\big(-X - \mathrm{E}(-X) \geqslant a\big) \leqslant \frac{V}{V+a^2}.$$

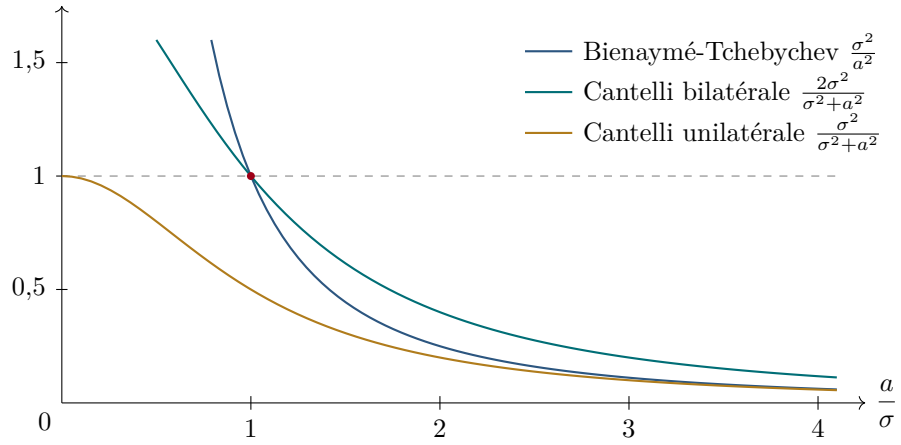
L'événement $\big(|X - \mathrm{E}(X)| \geqslant a\big)$ est la réunion de $\big(X - \mathrm{E}(X) \geqslant a\big)$ et de $\big(X - \mathrm{E}(X) \leqslant -a\big)$, et la sous-additivité (proposition 7 (iv)) donne

$$\boxed{\mathrm{P}\big(|X - \mathrm{E}(X)| \geqslant a\big) \leqslant \frac{2V}{V+a^2}}$$

Comparons avec la borne $\dfrac{V}{a^2}$ de Bienaymé-Tchebychev, lorsque $V > 0$:

$$\frac{2V}{V+a^2} < \frac{V}{a^2} \iff 2a^2 < V + a^2 \iff a^2 < V.$$

La nouvelle borne n'est donc meilleure que si $a < \sigma(X)$. Mais alors $\dfrac{V}{a^2} > 1$ et $\dfrac{2V}{V+a^2} > 1$: les deux majorations sont sans intérêt, puisqu'une probabilité est toujours au plus 1. *La version bilatérale n'améliore jamais utilement l'inégalité de Bienaymé-Tchebychev.* L'intérêt de l'inégalité de Cantelli est la version unilatérale : pour $a = \sigma(X)$, elle donne $\dfrac{1}{2}$ là où Bienaymé-Tchebychev ne donne que 1, et pour $a = 2\sigma(X)$, elle donne $\dfrac{1}{5}$ au lieu de $\dfrac{1}{4}$.



Les trois majorants sont tracés en fonction de $\dfrac{a}{\sigma}$. Les courbes de Bienaymé-Tchebychev, en bleu, et de Cantelli bilatérale, en bleu pétrole, se coupent au point rouge, d'ordonnée 1 : la seconde n'est en dessous de la première que là où toutes deux dépassent 1. La borne unilatérale, en ocre, est partout la plus petite et reste inférieure à 1.

Piège. L'erreur la plus répandue est de comparer des bornes qui ne majorent pas la même probabilité : la borne de Cantelli unilatérale porte sur un seul côté de l'écart, et doit être doublée, comme à la question 3, pour être comparée honnêtement à Bienaymé-Tchebychev.

EXERCICE 100 ⋆⋆⋆

L'univers est $\mathfrak{S}_n$, muni de la probabilité uniforme. L'entier i est un record de Σ si et seulement si $\Sigma(i) = \max \Sigma([\![1\,,i]\!])$: cette condition ne fait intervenir que les valeurs prises sur $[\![1\,,i]\!]$.

1. Soit $i < j$. On construit une permutation σ dont i et j sont des records par choix successifs, dont le nombre ne dépend pas des choix précédents (lemme des bergers) :

- on choisit l'ensemble $A = \sigma([\![1\,,j]\!])$, partie à j éléments de $[\![1\,,n]\!]$: $\binom{n}{j}$ façons ;
- on choisit la restriction de σ à $[\![j+1\,,n]\!]$, bijection sur le complémentaire de A : $(n-j)!$ façons (théorème 1) ;
- j est un record si et seulement si $\sigma(j) = \max A$: une seule façon ;
- il reste à placer les $j-1$ valeurs de $A' = A \setminus \{\max A\}$ sur $[\![1\,,j-1]\!]$. On choisit l'ensemble $B = \sigma([\![1\,,i]\!]) \subset A'$, de $\binom{j-1}{i}$ façons ; i est un record si et seulement si $\sigma(i) = \max B$; on place les autres valeurs de B sur $[\![1\,,i-1]\!]$, de $(i-1)!$ façons, et celles de $A' \setminus B$ sur $[\![i+1\,,j-1]\!]$, de $(j-1-i)!$ façons.

Chaque permutation de l'événement est obtenue exactement une fois, puisque tous les choix se lisent sur σ. Le nombre de ces permutations vaut

$$\binom{n}{j}(n-j)!\binom{j-1}{i}(i-1)!\,(j-1-i)! = \frac{n!}{j!}\cdot\frac{(j-1)!}{i!\,(j-1-i)!}\,(i-1)!\,(j-1-i)! = \frac{n!}{j}\cdot\frac{1}{i},$$

et, en divisant par $n!$,

$$\boxed{\mathrm{P}(I_i = 1, I_j = 1) = \frac{1}{ij} = \mathrm{P}(I_i = 1)\,\mathrm{P}(I_j = 1)}$$

Le résultat a été contrôlé par énumération de $\mathfrak{S}_n$ pour $n \leqslant 7$ ✓.

2. Les indicatrices I_i suivent les lois $\mathcal{B}(\frac{1}{i})$, et, pour $i < j$, la question 1 et la proposition 23 (i) donnent

$$\mathrm{Cov}(I_i, I_j) = \mathrm{E}(I_iI_j) - \mathrm{E}(I_i)\,\mathrm{E}(I_j) = \frac{1}{ij} - \frac{1}{i}\cdot\frac{1}{j} = 0.$$

Les I_i sont deux à deux décorrélées, et le théorème 13, avec la proposition 21, donne

$$\boxed{\mathrm{V}(R_n) = \sum_{i=1}^{n} \frac{1}{i}\Big(1 - \frac{1}{i}\Big) = H_n - \sum_{i=1}^{n} \frac{1}{i^2} \leqslant H_n}$$

Pour $n = 100$, on trouve $\mathrm{E}(R_{100}) \approx 5{,}19$ et $\mathrm{V}(R_{100}) \approx 3{,}55$, soit un écart type d'environ 1,9.

3. D'après le chapitre 20, $H_n - \ln n$ tend vers la constante d'Euler γ, donc $\dfrac{H_n - \ln n}{\ln n}$ tend vers 0, et $H_n \sim \ln n$. Soit $\varepsilon > 0$, et $n_0 \geqslant 2$ tel que $|H_n - \ln n| \leqslant \dfrac{\varepsilon}{2} \ln n$ pour $n \geqslant n_0$. Pour un tel n, l'inégalité triangulaire donne

$$\Big|\frac{R_n}{\ln n} - 1\Big| \leqslant \frac{|R_n - H_n|}{\ln n} + \frac{|H_n - \ln n|}{\ln n} \leqslant \frac{|R_n - H_n|}{\ln n} + \frac{\varepsilon}{2}.$$

L'événement étudié est donc inclus dans $\big(|R_n - H_n| \geqslant \dfrac{\varepsilon}{2} \ln n\big)$. Comme $\mathrm{E}(R_n) = H_n$, l'inégalité de Bienaymé-Tchebychev et la question 2 donnent

$$\mathrm{P}\Big(\Big|\frac{R_n}{\ln n} - 1\Big| > \varepsilon\Big) \leqslant \frac{4\mathrm{V}(R_n)}{\varepsilon^2 (\ln n)^2} \leqslant \frac{4H_n}{\varepsilon^2 (\ln n)^2} \underset{n\to+\infty}{\sim} \frac{4}{\varepsilon^2 \ln n},$$

$$\boxed{\mathrm{P}\Big(\Big|\frac{R_n}{\ln n} - 1\Big| > \varepsilon\Big) \xrightarrow[n\to+\infty]{} 0}$$

À retenir. La variable R_n n'est pas une somme de variables de même loi, et la loi faible des grands nombres ne s'applique pas telle quelle ; on refait sa démonstration, qui n'utilise que l'espérance, la variance et la décorrélation. La convergence est très lente : la majoration décroît comme $\dfrac{1}{\ln n}$, et les fluctuations de R_n autour de $\ln n$ sont de l'ordre de $\sqrt{\ln n}$.

EXERCICE 101 ⋆⋆⋆

1. Soit $s \geqslant 0$. Si $s = 0$, le majorant vaut 1, et l'inégalité est évidente. Si $s > 0$, la fonction $x \mapsto \mathrm{e}^{sx}$ est strictement croissante, et les événements $\big(\dfrac{S_n}{n} \geqslant a\big)$ et $\big(\mathrm{e}^{sS_n} \geqslant \mathrm{e}^{nsa}\big)$ sont égaux. L'inégalité de Markov, appliquée à la variable positive e^{sS_n}, donne

$$\mathrm{P}\Big(\frac{S_n}{n} \geqslant a\Big) \leqslant \mathrm{e}^{-nsa}\,\mathrm{E}(\mathrm{e}^{sS_n}).$$

La formule de transfert (théorème 10) et la formule du binôme donnent

$$\mathrm{E}(\mathrm{e}^{sS_n}) = \sum_{k=0}^{n} \mathrm{e}^{sk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (p\,\mathrm{e}^s)^k (1-p)^{n-k} = (1 - p + p\,\mathrm{e}^s)^n,$$

$$\boxed{\mathrm{P}\Big(\frac{S_n}{n} \geqslant a\Big) \leqslant \mathrm{e}^{-nsa}(1 - p + p\,\mathrm{e}^s)^n \quad \text{pour tout } s \geqslant 0}$$

Piège. L'hypothèse $s \geqslant 0$ est indispensable : pour $s < 0$, l'exponentielle est décroissante et l'inclusion change de sens. Pour $n = 10$, $p = \dfrac{1}{2}$ et $a = \dfrac{1}{4}$, l'énoncé étendu à $s < 0$ donnerait un majorant d'environ 0,27, alors que $\mathrm{P}(S_{10} \geqslant 2{,}5) \approx 0{,}945$ ✓.

2. Minimiser le majorant revient à minimiser son logarithme divisé par n,

$$h(s) = -sa + \ln(1 - p + p\,\mathrm{e}^s), \qquad h'(s) = -a + \frac{p\,\mathrm{e}^s}{1 - p + p\,\mathrm{e}^s} = -a + 1 - \frac{1-p}{1 - p + p\,\mathrm{e}^s}.$$

La fonction h' est strictement croissante, $h'(0) = p - a < 0$, et $h'(s)$ tend vers $1 - a > 0$ en $+\infty$. Elle s'annule en un unique point s^*, défini par $p\,\mathrm{e}^{s^*}(1-a) = a(1-p)$:

$$\mathrm{e}^{s^*} = \frac{a(1-p)}{p(1-a)} > 1,$$

puisque $a(1-p)-p(1-a)=a-p>0$. La fonction h décroît sur $[0\,,s^*]$ et croît ensuite : son minimum sur $\mathbb{R}_+$ est $h(s^*)$. Or $1-p+p\,\mathrm{e}^{s^*}=(1-p)(1+\dfrac{a}{1-a})=\dfrac{1-p}{1-a}$, d'où

$$h(s^*)=-a\ln\frac{a(1-p)}{p(1-a)}+\ln\frac{1-p}{1-a}=-a\ln\frac{a}{p}-(1-a)\ln\frac{1-a}{1-p}=-D(a,p).$$

$$\boxed{\text{la meilleure majoration est } \mathrm{P}\Big(\frac{S_n}{n}\geqslant a\Big)\leqslant \mathrm{e}^{-nD(a,p)}}$$

3. Pour $t\in[q\,,r]$, la décomposition en éléments simples (chapitre 13) donne

$$\frac{r-t}{t(1-t)}=\frac{r}{t}+\frac{r-1}{1-t},$$

ce que l'on vérifie en réduisant au même dénominateur : $r(1-t)+(r-1)t=r-t$. Comme $0<q\leqslant t\leqslant r<1$,

$$\int_q^r\frac{r-t}{t(1-t)}\,\mathrm{d}t=r\ln\frac{r}{q}+(r-1)\big[-\ln(1-t)\big]_q^r=r\ln\frac{r}{q}+(1-r)\ln\frac{1-r}{1-q},$$

$$\boxed{D(r,q)=\int_q^r\frac{r-t}{t(1-t)}\,\mathrm{d}t}$$

Pour $t\in[q\,,r]$, on a $0<t(1-t)\leqslant\dfrac14$ et $r-t\geqslant0$, donc $\dfrac{r-t}{t(1-t)}\geqslant4(r-t)$. Par croissance de l'intégrale (chapitre 18), les bornes étant dans l'ordre croissant,

$$\boxed{D(r,q)\geqslant4\int_q^r(r-t)\,\mathrm{d}t=2(r-q)^2}$$

4. Soit $\varepsilon>0$ et $a=p+\varepsilon$. Montrons d'abord que $\mathrm{P}(\dfrac{S_n}{n}-p\geqslant\varepsilon)\leqslant\mathrm{e}^{-2n\varepsilon^2}$.

— Si $a>1$, l'événement est impossible, puisque $S_n\leqslant n$.

— Si $a<1$, les questions 2 et 3, avec $r=a$ et $q=p$, donnent $\mathrm{e}^{-nD(a,p)}\leqslant\mathrm{e}^{-2n\varepsilon^2}$.

— Si $a=1$, l'événement est $(S_n=n)$, de probabilité $p^n=\mathrm{e}^{-n\ln(1/p)}$, et le calcul de la question 3 avec $r=1$ s'applique encore : $\ln\dfrac1p=\displaystyle\int_p^1\frac{\mathrm{d}t}{t}\geqslant\int_p^1 4(1-t)\,\mathrm{d}t=2(1-p)^2$, puisque $\dfrac1t=\dfrac{1-t}{t(1-t)}\geqslant4(1-t)$.

Pour l'autre côté, la variable $n-S_n$ suit la loi $\mathcal{B}(n,1-p)$, puisque

$$\mathrm{P}(n-S_n=k)=\mathrm{P}(S_n=n-k)=\binom{n}{k}(1-p)^kp^{n-k}.$$

Le résultat précédent, qui ne dépend que de la loi, appliqué à $n-S_n$ et au paramètre $1-p$, donne

$$\mathrm{P}\Big(\frac{S_n}{n}-p\leqslant-\varepsilon\Big)=\mathrm{P}\Big(\frac{n-S_n}{n}-(1-p)\geqslant\varepsilon\Big)\leqslant\mathrm{e}^{-2n\varepsilon^2}.$$

L'événement $\big(\big|\dfrac{S_n}{n}-p\big|\geqslant\varepsilon\big)$ est la réunion des deux, et la sous-additivité conclut :

$$\boxed{\mathrm{P}\Big(\Big|\frac{S_n}{n}-p\Big|\geqslant\varepsilon\Big)\leqslant2\,\mathrm{e}^{-2n\varepsilon^2}}$$

Pour $p=\dfrac12$, c'est exactement la proposition 24, obtenue ici sans la majoration du cosinus hyperbolique, et le résultat vaut maintenant pour tout p. La forme intermédiaire $\mathrm{e}^{-nD(a,p)}$ est même bien meilleure

loin de $\frac{1}{2}$, où $t(1-t)$ est petit devant $\frac{1}{4}$. Pour $n = 1\,000$, $p = 0{,}1$ et $\varepsilon = 0{,}05$, la probabilité exacte vaut environ $4{,}5 \times 10^{-7}$; la somme des deux majorants e^{-nD} vaut environ $4{,}9 \times 10^{-6}$, la borne $2\,\mathrm{e}^{-2n\varepsilon^2}$ environ 0,013, et la borne $\dfrac{p(1-p)}{n\varepsilon^2}$ du théorème 16 vaut 0,036 ✓.

À retenir. La méthode est celle de la proposition 24 : l'inégalité de Markov appliquée à e^{sS_n}, puis l'optimisation en s. Le choix de l'exponentielle transforme la somme de variables indépendantes en produit, dont l'espérance se calcule par la proposition 20 ou, comme ici, par la formule du binôme.

Exercices de synthèse

EXERCICE 102 ⋆⋆

1. Vrai. Si A et B sont incompatibles, $\mathrm{P}(A \cap B) = \mathrm{P}(\varnothing) = 0 < \mathrm{P}(A)\,\mathrm{P}(B)$. C'est la première remarque qui suit la proposition 10 : savoir que l'un est réalisé apprend que l'autre ne l'est pas.

2. Faux. On lance deux pièces équilibrées, avec $\Omega = \{P, F\}^2$ muni de la probabilité uniforme ; A est « la première donne pile », B « la seconde donne pile », C « les deux donnent le même côté ». Chacun a la probabilité $\frac{1}{2}$, chaque intersection de deux d'entre eux est réduite à $\{(P,P)\}$, de probabilité $\frac{1}{4}$: ils sont deux à deux indépendants. Mais $\mathrm{P}(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8}$. C'est le second exemple qui suit la proposition 10.

3. Vrai. Si $X \sim Y$, alors $X^2 \sim Y^2$ par la proposition 13 avec $f(x) = x^2$. La formule de transfert exprime $\mathrm{E}(X^3) = \sum_x x^3\,\mathrm{P}(X = x)$ à l'aide de la seule loi de X, donc $\mathrm{E}(X^3) = \mathrm{E}(Y^3)$.

4. Faux. Soit D le résultat d'un dé équilibré, $X = D$ et $Y = 7 - D$. Les variables X et Y suivent la loi uniforme sur $[\![1\,,6]\!]$, mais $X + Y = 7$ est constante alors que $2X$ prend six valeurs. C'est la remarque qui suit la proposition 13.

5. Faux. Soit X de loi uniforme sur $\{-1, 0, 1\}$ et $Y = X^2$. Alors

$$\mathrm{E}(XY) = \mathrm{E}(X^3) = 0 = \mathrm{E}(X)\,\mathrm{E}(Y), \qquad \text{mais} \qquad \mathrm{P}(X = 0, Y = 1) = 0 \neq \mathrm{P}(X = 0)\,\mathrm{P}(Y = 1) = \frac{2}{9}.$$

C'est la remarque qui suit la proposition 20 : la relation $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$ exprime seulement la décorrélation.

6. Faux. Soit X de loi $\mathcal{B}(\frac{1}{2})$ et $Y = X$. Alors $\mathrm{V}(X + Y) = \mathrm{V}(2X) = 4\mathrm{V}(X) = 1$, alors que $\mathrm{V}(X) + \mathrm{V}(Y) = \frac{1}{2}$. Le théorème 13 exige que les variables soient deux à deux décorrélées.

7. Vrai. Par l'indépendance et la proposition 15, la loi du couple (X, Y) est le produit de ses lois marginales ; elle ne dépend que des lois $\mathcal{B}(n,p)$ et $\mathcal{B}(m,p)$, et la loi de $X + Y$ aussi (proposition 13). Il suffit donc de la calculer sur un modèle commode. La proposition 17 fournit des variables $B_1, \ldots, B_{n+m}$ mutuellement indépendantes de loi $\mathcal{B}(p)$; les sommes $U = B_1 + \cdots + B_n$ et $W = B_{n+1} + \cdots + B_{n+m}$ suivent les lois $\mathcal{B}(n,p)$ et $\mathcal{B}(m,p)$ (théorème 8) et sont indépendantes (théorème 7). Le couple (U, W) a donc même loi que (X, Y), et $X + Y$ a même loi que $U + W = B_1 + \cdots + B_{n+m}$, qui suit la loi $\mathcal{B}(n+m,p)$ par le théorème 8. On peut aussi calculer directement, avec la formule de Vandermonde de la remarque qui suit la proposition 4 :

$$\mathrm{P}(X + Y = k) = \sum_j \binom{n}{j}\binom{m}{k-j} p^k (1-p)^{n+m-k} = \binom{n+m}{k} p^k (1-p)^{n+m-k}.$$

8. Vrai. Comme $A \cap B \subset A$ et $A \cap B \subset B$, la croissance donne $\mathrm{P}(A) > 0$ et $\mathrm{P}(B) > 0$: les deux probabilités conditionnelles sont définies. Alors

$$\mathrm{P}_B(A) = \mathrm{P}_A(B) \iff \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} \iff \mathrm{P}(A) = \mathrm{P}(B),$$

la dernière équivalence venant de $\mathrm{P}(A \cap B) \neq 0$. Sans cette hypothèse, l'équivalence tombe : deux événements incompatibles de probabilités $\frac{1}{3}$ et $\frac{1}{2}$ vérifient $\mathrm{P}_B(A) = \mathrm{P}_A(B) = 0$.

9. Vrai. Soit X une variable qui prend les valeurs -2, 0 et 2 avec les probabilités $\frac{1}{8}$, $\frac{3}{4}$ et $\frac{1}{8}$. Alors

$$\mathrm{E}(X) = 0, \qquad \mathrm{V}(X) = \mathrm{E}(X^2) = 4 \times \frac{1}{4} = 1, \qquad \mathrm{P}(|X| \geqslant 2) = \frac{1}{4} = \frac{\mathrm{V}(X)}{2^2}.$$

L'inégalité de Bienaymé-Tchebychev est une égalité pour cette variable : aucune majoration de la probabilité d'un écart à la moyenne au moins égal à ε, qui ne dépendrait que de la variance et de ε, ne peut être meilleure.

10. Vrai. Par la proposition 6 pour les surjections de $[\![1\,,5]\!]$ sur $[\![1\,,3]\!]$, et par le théorème 1 pour les injections de $[\![1\,,3]\!]$ dans $[\![1\,,5]\!]$, ces nombres valent respectivement

$$3^5 - 3 \times 2^5 + 3 \times 1^5 = 243 - 96 + 3 = 150 \qquad \text{et} \qquad A_5^3 = 5 \times 4 \times 3 = 60.$$

$$\boxed{\text{Vrai : } 1, 3, 7, 8, 9, 10; \qquad \text{Faux : } 2, 4, 5, 6}$$

Piège. Les assertions 4, 5 et 6 commettent la même faute : elles traitent des variables comme si leurs lois suffisaient à tout décrire. Or $X+Y$, $\mathrm{E}(XY)$ et $\mathrm{V}(X+Y)$ dépendent de la loi *conjointe* du couple, que les lois marginales ne déterminent pas ; l'indépendance est précisément l'hypothèse qui fait de la loi conjointe le produit des marginales (proposition 15), et c'est elle qui rend vraie l'assertion 7.

EXERCICE 103 ⋆⋆

L'univers est $[\![1\,,n]\!]$, muni de la probabilité uniforme : $\mathrm{P}(E) = \dfrac{|E|}{n}$ pour toute partie E.

1. D'après le chapitre 12, x et n ne sont pas premiers entre eux si et seulement s'ils ont un diviseur premier commun, c'est-à-dire si et seulement si l'un des p_k divise x, puisque les diviseurs premiers de n sont exactement p_1, …, p_r. Ainsi

$$\boxed{B = \bigcap_{k=1}^{r} \overline{A_{p_k}}}$$

2. Si m divise n, les multiples de m dans $[\![1\,,n]\!]$ sont m, $2m$, …, $\frac{n}{m}\,m$; il y en a $\frac{n}{m}$, et

$$\boxed{\mathrm{P}(A_m) = \frac{1}{m}}$$

3. Soit J une partie non vide de $[\![1\,,r]\!]$ et $m_J = \prod_{k\in J} p_k$. Les p_k étant des nombres premiers distincts, donc premiers entre eux deux à deux, un entier est divisible par chacun d'eux si et seulement s'il est divisible par leur produit (chapitre 12). Ainsi $\bigcap_{k\in J} A_{p_k} = A_{m_J}$, et m_J divise n ; la question 2 donne

$$\mathrm{P}\Big(\bigcap_{k\in J} A_{p_k}\Big) = \frac{1}{m_J} = \prod_{k\in J} \frac{1}{p_k} = \prod_{k\in J} \mathrm{P}(A_{p_k}).$$

$$\boxed{A_{p_1}, \ldots, A_{p_r} \text{ sont mutuellement indépendants}}$$

Pour $n = 60 = 2^2 \times 3 \times 5$, on retrouve l'exemple du cours. L'hypothèse que les p_k divisent n est essentielle : sur $[\![1\,,100]\!]$, le cours a montré que A_2 et A_3 ne sont pas indépendants, puisque $\mathrm{P}(A_2 \cap A_3) = \mathrm{P}(A_6) = \dfrac{16}{100}$ diffère de $\mathrm{P}(A_2)\,\mathrm{P}(A_3) = \dfrac{50}{100} \times \dfrac{33}{100}$.

4. Par la proposition 10 (iii), les contraires $\overline{A_{p_1}}$, …, $\overline{A_{p_r}}$ sont mutuellement indépendants, de probabilités $1 - \dfrac{1}{p_k}$. La question 1 donne

$$\boxed{\mathrm{P}(B) = \prod_{k=1}^{r} \Big(1 - \frac{1}{p_k}\Big)}$$

et, comme $\mathrm{P}(B) = \dfrac{|B|}{n} = \dfrac{\varphi(n)}{n}$,

$$\boxed{\varphi(n) = n \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right)}$$

Pour $n = 60$, on trouve $\varphi(60) = 60 \times \dfrac{1}{2} \times \dfrac{2}{3} \times \dfrac{4}{5} = 16$, et la formule a été contrôlée pour $2 \leqslant n \leqslant 300$ ✓.

5. Si $n = p$ est premier, $r = 1$ et $\boxed{\varphi(p) = p - 1}$: directement, les entiers $1, \ldots, p-1$ sont premiers avec p, et p ne l'est pas. Si $n = p^\alpha$ avec $\alpha \geqslant 1$, on a encore $r = 1$, et

$$\boxed{\varphi(p^\alpha) = p^\alpha\left(1 - \frac{1}{p}\right) = p^\alpha - p^{\alpha-1}}$$

Directement, un entier n'est pas premier avec p^α si et seulement s'il est multiple de p, et il y a $p^{\alpha-1}$ multiples de p dans $[\![1\,,p^\alpha]\!]$.
À retenir. Une formule d'arithmétique devient une conséquence de l'indépendance : être divisible par 2 et être divisible par 3 sont deux « informations indépendantes » sur un entier tiré au hasard dans $[\![1\,,n]\!]$, dès que 6 divise n. La formule montre aussi que $\varphi(ab) = \varphi(a)\varphi(b)$ lorsque a et b sont premiers entre eux, puisque les diviseurs premiers de ab sont alors ceux de a et ceux de b, sans répétition. La fonction φ intervenait déjà dans l'exercice 62 de la feuille du chapitre 20.

EXERCICE 104 ⋆⋆

Le modèle. On note W l'ensemble des g numéros gagnants, fixé, et l'on décrit le jeu par trois parties aléatoires de $[\![1\,,n]\!]$ à g éléments : T, tirée par le joueur ; R, dévoilée par le meneur ; T', tirée dans la seconde phase. Comme pour les tirages successifs de l'exemple qui suit la proposition 9, l'énoncé fixe les probabilités conditionnelles :

— T suit la loi uniforme sur les parties à g éléments de $[\![1\,,n]\!]$;

— sachant $(T = t)$, R suit la loi uniforme sur les parties à g éléments de l'ensemble L_t des numéros perdants non tirés ;

— sachant $(T = t) \cap (R = \rho)$, T' suit la loi uniforme sur les parties à g éléments de $[\![1\,,n]\!] \setminus (t \cup \rho)$.

La formule des probabilités composées (proposition 9) donne alors la probabilité de chaque triplet (t, ρ, t'). On pose $G_1 = |T \cap W|$ et $G_2 = |T' \cap W|$.

1. Le tirage de T est celui de l'exercice 89, avec n boules, dont g « blanches », les numéros gagnants, et g boules tirées ; l'hypothèse $1 \leqslant g \leqslant n-1$ de cet exercice est satisfaite, puisque $g \leqslant \dfrac{n}{3}$. Donc

$$\boxed{\mathrm{P}(G_1 = k) = \frac{\binom{g}{k}\binom{n-g}{g-k}}{\binom{n}{g}} \quad (0 \leqslant k \leqslant g), \qquad \mathrm{E}(G_1) = g \times \frac{g}{n} = \frac{g^2}{n}}$$

Toutes ces probabilités sont strictement positives, puisque $n - g \geqslant g \geqslant g - k$.

2.(a) Soit t une partie à g éléments telle que $|t \cap W| = k$. Parmi les $n - g$ numéros perdants, $g - k$ ont été tirés, donc

$$|L_t| = n - g - (g - k) = n - 2g + k \geqslant n - 2g \geqslant g,$$

puisque $n \geqslant 3g$: le meneur peut toujours dévoiler g numéros perdants. Soit $\rho \subset L_t$ à g éléments. L'ensemble $[\![1\,,n]\!] \setminus (t \cup \rho)$ a $n - 2g$ éléments ; il contient les $g - k$ numéros gagnants non tirés, puisque ρ ne contient que des perdants, et $n - 3g + k$ numéros perdants. Le nombre de ses parties à g éléments

contenant exactement j gagnants est $\binom{g-k}{j}\binom{n-3g+k}{g-j}$ (théorème 2 et lemme des bergers), d'où

$$\mathrm{P}\big(G_2 = j \mid (T=t)\cap(R=\rho)\big) = h_k(j) = \frac{\binom{g-k}{j}\binom{n-3g+k}{g-j}}{\binom{n-2g}{g}}.$$

Ce nombre ne dépend que de k. L'événement $(G_1 = k)$ est la réunion disjointe des événements $(T=t)\cap(R=\rho)$, pour les couples tels que $|t\cap W| = k$; en sommant,

$$\mathrm{P}(G_1=k, G_2=j) = \sum_{(t,\rho)} \mathrm{P}\big((T=t)\cap(R=\rho)\big)\, h_k(j) = \mathrm{P}(G_1=k)\, h_k(j),$$

$$\boxed{\mathrm{P}_{(G_1=k)}(G_2=j) = \frac{\binom{g-k}{j}\binom{n-3g+k}{g-j}}{\binom{n-2g}{g}}, \qquad 0\leqslant j\leqslant g-k}$$

C'est la loi de l'exercice 89, avec $n-2g$ boules, dont $g-k$ blanches, et g boules tirées.

2.(b) Pour une urne de N boules dont w blanches, dont on tire s boules, la formule du pion $j\binom{w}{j} = w\binom{w-1}{j-1}$ (chapitre 3, proposition 9) et la formule de Vandermonde donnent

$$\sum_j j\binom{w}{j}\binom{N-w}{s-j} = w\sum_j \binom{w-1}{j-1}\binom{N-w}{s-j} = w\binom{N-1}{s-1} = \frac{sw}{N}\binom{N}{s},$$

ce qui redonne la moyenne $\dfrac{sw}{N}$ de l'exercice 89, y compris dans les cas extrêmes $w=0$ et $w=N$. Avec $N = n-2g$, $w = g-k$ et $s = g$,

$$\boxed{E_k = \frac{g(g-k)}{n-2g}}$$

Les événements $(G_1 = k)$, pour $k\in[\![0\,,g]\!]$, forment un système complet d'événements de probabilités non nulles. Par la formule des probabilités totales (théorème 5), puis en échangeant les deux sommes finies,

$$\begin{aligned}\mathrm{E}(G_2) = \sum_j j\,\mathrm{P}(G_2=j) &= \sum_j j\sum_{k=0}^{g}\mathrm{P}(G_1=k)\,\mathrm{P}_{(G_1=k)}(G_2=j)\\ &= \sum_{k=0}^{g}\mathrm{P}(G_1=k)\sum_j j\,\mathrm{P}_{(G_1=k)}(G_2=j),\end{aligned}$$

$$\boxed{\mathrm{E}(G_2) = \sum_{k=0}^{g} E_k\,\mathrm{P}(G_1=k)}$$

2.(c) Comme E_k est une fonction affine de k, la linéarité de l'espérance, sous la forme de la formule de transfert, donne

$$\mathrm{E}(G_2) = \frac{g}{n-2g}\sum_{k=0}^{g}(g-k)\,\mathrm{P}(G_1=k) = \frac{g}{n-2g}\big(g-\mathrm{E}(G_1)\big) = \frac{g}{n-2g}\Big(g-\frac{g^2}{n}\Big),$$

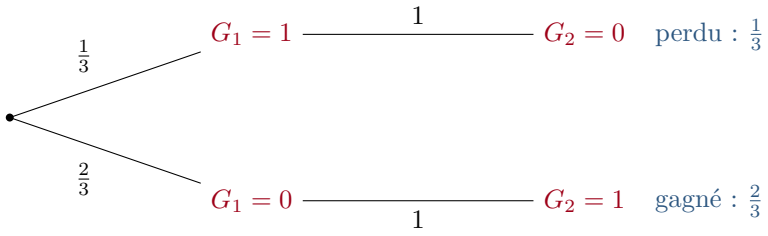
$$\boxed{\mathrm{E}(G_2) = \frac{g^2(n-g)}{n(n-2g)}}$$

Les formules de $\mathrm{E}(G_1)$ et $\mathrm{E}(G_2)$ ont été vérifiées exactement, par énumération de tous les déroulements du jeu, pour plusieurs couples (n, g) ✓.

3. Le rapport des deux espérances vaut

$$\frac{\mathrm{E}(G_2)}{\mathrm{E}(G_1)} = \frac{n-g}{n-2g} > 1 :$$

il faut changer. Pour $n = 3$ et $g = 1$, on a $\mathrm{E}(G_1) = \dfrac{1}{3}$ et $\boxed{\mathrm{E}(G_2) = \dfrac{2}{3}}$. Comme G_1 et G_2 ne prennent que les valeurs 0 et 1, ce sont les probabilités de gagner : changer de porte double les chances. L'arbre suivant décrit ce cas.



Si le joueur a d'abord choisi la bonne porte, ce qui arrive une fois sur trois, changer le fait perdre ; sinon, le meneur, qui doit ouvrir une porte perdante, ouvre la seule qui reste, et changer fait gagner à coup sûr.

Piège. L'erreur la plus répandue est de dire qu'« il reste deux portes, donc une chance sur deux pour chacune ». Le choix du meneur n'est pas indépendant de celui du joueur : il ne dévoile jamais un gagnant ni un numéro tiré. C'est cette information qui rend la réserve des $n - 2g$ numéros restants plus riche en gagnants que l'urne de départ ; le gain relatif, $\dfrac{n-g}{n-2g}$, est d'autant plus grand que g est proche de $\dfrac{n}{3}$.

EXERCICE 105 ⋆⋆⋆

On note $\mathcal{E}_X$ l'ensemble des arêtes dont les deux extrémités appartiennent à X, de sorte que $Y = |\mathcal{E}_X|$, et, pour une arête e, $\mu(e)$ son extrémité de plus grand numéro. La partie retirée est $D = \mu(\mathcal{E}_X) \subset X$, et $U = X \setminus D$.

1. Si une arête e était contenue dans U, elle le serait dans X, donc $e \in \mathcal{E}_X$, et l'extrémité $\mu(e)$ appartiendrait à D, donc pas à U : c'est absurde. Ainsi $\boxed{U \text{ est indépendante}}$. L'ensemble D est l'image de $\mathcal{E}_X$ par μ, donc $|D| \leqslant |\mathcal{E}_X| = Y$ (chapitre 2), et, comme $D \subset X$,

$$\boxed{|U| = |X| - |D| \geqslant |X| - Y}$$

L'inégalité peut être stricte : un même sommet peut être retiré à cause de plusieurs arêtes.

2. On a $|X| = \sum_{s \in S} B_s$ et $Y = \sum_{\{s,s'\} \in \mathcal{E}} B_s B_{s'}$, et, pour $s \neq s'$, $\mathrm{E}(B_s B_{s'}) = p^2$ par l'indépendance et la proposition 20. Par croissance et linéarité de l'espérance (théorème 9),

$$\mathrm{E}(|U|) \geqslant \mathrm{E}(|X|) - \mathrm{E}(Y) = np - |\mathcal{E}|\, p^2, \qquad \text{d'où} \qquad \boxed{\mathrm{E}(|U|) \geqslant np - \frac{p^2 nd}{2}}$$

C'est la méthode des indicatrices (section 21.6.3), qui ne demande pas la loi de $|U|$.

3. La fonction $f : p \mapsto np - \dfrac{p^2 nd}{2}$ est un trinôme du second degré, de coefficient dominant négatif, maximal en $p = \dfrac{1}{d}$, qui appartient à $]0\,,1]$ puisque $d \geqslant 1$:

$$f\Big(\frac{1}{d}\Big) = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d}, \qquad \mathrm{E}(|U|) \geqslant \frac{n}{2d} \quad \text{pour } p = \frac{1}{d}.$$

Par le principe de la moyenne (théorème 9 (v)), $\mathrm{P}\big(|U| \geqslant \mathrm{E}(|U|)\big) > 0$: l'événement n'est pas vide, et il existe une issue ω telle que $|U(\omega)| \geqslant \mathrm{E}(|U|) \geqslant \dfrac{n}{2d}$. La partie $U(\omega)$ est indépendante, d'où

$$\boxed{\alpha(\Gamma) \geqslant \frac{n}{2d}}$$

Pour l'hexagone, $n = 6$ et $d = 2$: la minoration donne $\alpha \geqslant \dfrac{3}{2}$, donc $\alpha \geqslant 2$, et l'on a en fait $\alpha = 3$; le calcul exact donne $\mathrm{E}(|U|) = \dfrac{13}{8}$ pour $p = \dfrac{1}{2}$ ✓. Le réel $d = \dfrac{2|\mathcal{E}|}{n}$ est le degré moyen d'un sommet. La minoration a le bon ordre de grandeur : pour la réunion disjointe de $\dfrac{n}{d+1}$ graphes complets à $d+1$ sommets, on a $\dfrac{nd}{2}$ arêtes et $\alpha = \dfrac{n}{d+1}$, puisqu'une partie indépendante contient au plus un sommet de chaque graphe complet.
À retenir. On a démontré l'existence d'une grande partie indépendante sans en construire aucune, en calculant une espérance : c'est la méthode probabiliste, déjà rencontrée avec les tournois de l'exercice 54 et avec l'exercice 50 de la feuille du chapitre 19. Le choix du paramètre p est un réglage : trop grand, il crée trop d'arêtes internes ; trop petit, la partie X est trop petite.

EXERCICE 106 ⋆⋆⋆

1.(a) Un chemin est déterminé par son point de départ et la suite de ses pas $\varepsilon_i = s_{i+1} - s_i \in \{-1, 1\}$, et il arrive en s_ℓ si et seulement si $\varepsilon_0 + \cdots + \varepsilon_{\ell-1} = s_\ell - s_0$. Si u pas valent 1 et $\ell - u$ valent -1, cette condition s'écrit $u - (\ell - u) = s_\ell - s_0$, soit $u = \dfrac{\ell + s_\ell - s_0}{2}$. Un chemin correspond donc à la partie de $[\![0\,, \ell - 1]\!]$ formée des indices des pas montants, de cardinal u, et le théorème 2 donne

$$\boxed{\text{nombre de chemins de } (0, s_0) \text{ à } (\ell, s_\ell) = \binom{\ell}{u}, \qquad u = \frac{\ell + s_\ell - s_0}{2}}$$

si $\ell + s_\ell - s_0$ est pair et $|s_\ell - s_0| \leqslant \ell$, et 0 sinon : il faut que u soit un entier de $[\![0\,, \ell]\!]$.

1.(b) *Un lemme des valeurs intermédiaires.* Un chemin qui va d'un entier strictement positif à un entier négatif s'annule : si i_0 est le plus petit indice tel que $s_{i_0} \leqslant 0$, alors $i_0 \geqslant 1$, $s_{i_0 - 1} \geqslant 1$, et $s_{i_0} \geqslant s_{i_0 - 1} - 1 \geqslant 0$, donc $s_{i_0} = 0$.

Soit $\mathcal{T}$ l'ensemble des chemins de $(0, s_0)$ à (ℓ, s_ℓ) qui touchent l'axe, et $\mathcal{T}'$ l'ensemble des chemins de $(0, s_0)$ à $(\ell, -s_\ell)$. Pour un chemin s qui touche l'axe, on note $\tau(s)$ le premier indice où il s'annule, et l'on définit le chemin réfléchi

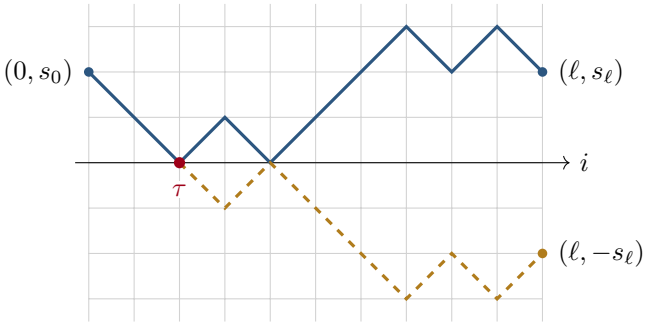
$$\Phi(s) = (s_0, \dots, s_\tau, -s_{\tau+1}, \dots, -s_\ell), \qquad \tau = \tau(s).$$

C'est bien un chemin : les pas avant l'indice τ sont conservés, ceux qui suivent l'indice $\tau + 1$ sont changés de signe, et, puisque $s_\tau = 0$,

$$|-s_{\tau+1} - s_\tau| = |s_{\tau+1}| = |s_{\tau+1} - s_\tau| = 1.$$

Si $s \in \mathcal{T}$, alors $\Phi(s) \in \mathcal{T}'$. Réciproquement, un chemin de $\mathcal{T}'$ va de $s_0 > 0$ à $-s_\ell < 0$: il touche l'axe par le lemme, et la même construction l'envoie dans $\mathcal{T}$. Les deux chemins s et $\Phi(s)$ coïncident jusqu'à l'indice $\tau(s)$, donc ont le même premier zéro, et réfléchir deux fois après ce zéro redonne le chemin initial : Φ réalise une bijection de $\mathcal{T}$ sur $\mathcal{T}'$, d'où

$$\boxed{|\mathcal{T}| = |\mathcal{T}'| = \text{nombre de chemins de } (0, s_0) \text{ à } (\ell, -s_\ell)}$$



Le chemin en bleu part de $(0, 2)$, touche l'axe pour la première fois au point rouge, d'abscisse $\tau = 2$, et arrive en $(10, 2)$; sa partie située après τ, réfléchie par rapport à l'axe, donne le chemin en pointillés ocre, qui arrive en $(10, -2)$.

2. Les $2n$ boules sont discernables, et l'univers des ordres de sortie, muni de la probabilité uniforme, a $(2n)!$ éléments. On ne retient que la suite des couleurs : chaque suite de n « blanc » et n « noir » provient de $n!\,n!$ ordres (lemme des bergers), donc la suite des couleurs suit la loi uniforme sur les $\binom{2n}{n}$ suites possibles. On lui associe le chemin $s_i =$ (blanches sorties) $-$ (noires sorties) après i tirages, de $(0, 0)$ à $(2n, 0)$; c'est une bijection entre suites de couleurs et chemins.

L'événement étudié est « $s_i \geqslant 0$ pour tout i ». En décalant d'une unité vers le haut, il correspond aux chemins de $(0, 1)$ à $(2n, 1)$ qui ne touchent pas l'axe. Par la question 1, ils sont au nombre de

$$\binom{2n}{n} - \binom{2n}{n-1},$$

le second terme comptant les chemins de $(0, 1)$ à $(2n, -1)$, pour lesquels $u = n - 1$. Comme

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

la probabilité vaut

$$\boxed{1 - \frac{n}{n+1} = \frac{1}{n+1}}$$

Le nombre de suites favorables, $\dfrac{1}{n+1}\binom{2n}{n}$, est le nombre de Catalan de l'exercice 21, obtenu ici par la réflexion plutôt que par le lemme cyclique. Le résultat a été contrôlé par énumération pour $n \leqslant 7$ ✓.

3. Comme à la question 2, l'ordre de dépouillement fournit une suite uniforme parmi les $\binom{1\,000}{400}$ suites de 600 voix A et 400 voix B, et le chemin $s_i =$ (voix de A) $-$ (voix de B) va de $(0, 0)$ à $(1\,000, 200)$. Décalé d'une unité, il doit aller de $(0, 1)$ à $(1\,000, 201)$ sans toucher l'axe. Les chemins qui le touchent correspondent à ceux de $(0, 1)$ à $(1\,000, -201)$, pour lesquels $u = \dfrac{1\,000 - 201 - 1}{2} = 399$. La probabilité vaut

$$1 - \frac{\binom{1\,000}{399}}{\binom{1\,000}{400}} = 1 - \frac{400}{601}, \qquad \boxed{\text{soit } \frac{201}{601} \approx 0{,}334}$$

Plus généralement, pour $a \geqslant b$ voix, on trouve $\dfrac{a - b + 1}{a + 1}$. Bien que A l'emporte avec 60 % des voix, il n'est resté en tête, au sens large, tout au long du dépouillement qu'avec une probabilité d'environ un

tiers. La version stricte, « A toujours strictement en tête », se traite de même : la première voix doit être pour A, puis le chemin va de $(1,1)$ à $(1\,000, 200)$ sans toucher l'axe, et l'on trouve

$$\frac{\binom{999}{599} - \binom{999}{399}}{\binom{1\,000}{400}} = \frac{600}{1\,000} - \frac{400}{1\,000} = \frac{a-b}{a+b} = \frac{1}{5},$$

ce que confirme un calcul par programmation dynamique ✓.

4. On code un client qui paie avec un billet de 5 euros par un pas $+1$ et un client qui paie avec un billet de 10 euros par un pas -1 : après le passage des i premiers clients, la caisse contient $r + s_i$ billets de 5 euros, où s est le chemin de $(0,0)$ à $(100,10)$ associé à la file. Un client qui paie avec 10 euros peut être servi si et seulement s'il reste au moins un billet de 5 euros avant son passage, c'est-à-dire si $r + s_i \geqslant 0$ après. On peut donc servir toute la file si et seulement si $r + s_i \geqslant 0$ pour tout i, c'est-à-dire si le chemin $(r + 1 + s_i)_i$, de $(0, r+1)$ à $(100, r+11)$, ne touche pas l'axe. Comme à la question 2, la file fournit une suite uniforme parmi les $\binom{100}{45}$ suites possibles.

Si $r \geqslant 45$, on peut toujours servir tout le monde, puisque $s_i \geqslant -45$. Si $0 \leqslant r \leqslant 44$, les chemins qui touchent l'axe correspondent à ceux de $(0, r+1)$ à $(100, -r-11)$, pour lesquels

$$u = \frac{100 - (r+11) - (r+1)}{2} = 44 - r.$$

La probabilité de servir toute la file vaut

$$\boxed{P(r) = 1 - \frac{\binom{100}{44-r}}{\binom{100}{45}} \quad (0 \leqslant r \leqslant 44), \qquad P(r) = 1 \quad (r \geqslant 45)}$$

Pour $r = 0$, on retrouve $1 - \dfrac{45}{56} = \dfrac{11}{56}$, valeur du scrutin avec $a = 55$ et $b = 45$. Le programme suivant cherche le plus petit r convenable.

```python
from math import comb

def P(r):
    # probabilite de servir toute la file avec r billets de 5 euros
    if r >= 45:
        return 1.0
    return 1 - comb(100, 44 - r) / comb(100, 45)

r = 0
while P(r) < 0.95:
    r += 1
print(r, P(r), P(r - 1))
```

Il affiche $r = 8$, avec $P(8) \approx 0{,}968$ et $P(7) \approx 0{,}944$:

$$\boxed{r = 8}$$

La formule a été confirmée, pour tout $r \in [\![0\,, 46]\!]$, par un calcul direct par programmation dynamique ✓. Huit billets de 5 euros suffisent dans 95 % des cas, alors que le pire cas, où les 45 billets de 10 euros arrivent en tête, en demanderait 45.

À retenir. La réflexion remplace une condition difficile à compter, « ne jamais passer sous un niveau », par un simple changement de point d'arrivée. Toute la difficulté est de bien coder le problème : choisir le pas, placer le niveau interdit, et décaler le chemin pour que ce niveau devienne l'axe.

Les quatre derniers exercices sont des problèmes en autonomie ; chacun commence par l'idée directrice.

EXERCICE 107 ⋆⋆⋆⋆

Idée. Si le premier passager s'assied à la place k, avec $2 \leqslant k \leqslant n-1$, les passagers 2 à $k-1$ s'assoient normalement, et le passager k se retrouve exactement dans la situation du premier passager d'un avion plus petit, où la place 1 joue le rôle de sa propre place. On obtient une relation de récurrence forte.

Le modèle. Le déroulement est déterminé par la suite des places choisies au hasard. Comme dans l'exemple qui suit la proposition 9, l'énoncé fixe les probabilités conditionnelles : chaque passager qui choisit au hasard prend chaque place libre avec la même probabilité, et la probabilité d'un déroulement est le produit des inverses des nombres de places libres lors des choix successifs. On note q_n la probabilité que le dernier passager s'asseye à sa place, et C_1 la place prise par le premier passager, de loi uniforme sur $[\![1\,,n]\!]$.

Conditionnement par la place du premier passager.

— Si $C_1 = 1$, chaque passager trouve sa place libre : le dernier s'assied à sa place.

— Si $C_1 = n$, la place n est prise : le dernier ne s'assied pas à sa place.

— Si $C_1 = k$ avec $2 \leqslant k \leqslant n-1$, les passagers 2, …, $k-1$ trouvent leur place libre et s'y assoient. Le passager k trouve sa place prise ; les places libres sont 1, $k+1$, …, n, au nombre de $n' = n-k+1$. Renumérotons : le passager k et la place 1 deviennent le passager et la place $1'$, et, pour $\ell \in [\![1\,,n-k]\!]$, le passager et la place $k+\ell$ deviennent $(\ell+1)'$. Le passager $1'$ choisit une place au hasard parmi les n' places ; chacun des suivants s'assied à sa place si elle est libre, et sinon au hasard parmi les places libres. C'est exactement le problème initial avec n' passagers, et les déroulements se correspondent avec les mêmes probabilités, produits des mêmes facteurs. Le dernier passager, n, devient n' et sa place devient n', donc

$$\mathrm{P}_{(C_1=k)}(\text{le dernier est à sa place}) = q_{n-k+1}.$$

Récurrence. La formule des probabilités totales (théorème 5), appliquée au système complet $\big((C_1=k)\big)_{1\leqslant k\leqslant n}$, donne, en posant $m = n-k+1$,

$$q_n = \frac{1}{n}\Big(1 + 0 + \sum_{k=2}^{n-1} q_{n-k+1}\Big) = \frac{1}{n}\Big(1 + \sum_{m=2}^{n-1} q_m\Big).$$

Pour $n = 2$, la somme est vide et $q_2 = \dfrac{1}{2}$: le premier passager prend sa place ou celle du second. Supposons $q_m = \dfrac{1}{2}$ pour $2 \leqslant m \leqslant n-1$. Alors

$$q_n = \frac{1}{n}\Big(1 + \frac{n-2}{2}\Big) = \frac{1}{n}\cdot\frac{n}{2} = \frac{1}{2},$$

et, par récurrence forte,

$$\boxed{\text{le dernier passager s'assied à sa place avec la probabilité } \frac{1}{2}, \text{ quel que soit } n \geqslant 2}$$

Le calcul exact, en énumérant tous les déroulements avec leurs probabilités, le confirme pour $n \leqslant 8$ ✓.

Remarque. Le résultat s'explique sans calcul. Pour $2 \leqslant j \leqslant n-1$, la place j est occupée au plus tard quand le passager j s'assied ; à l'arrivée du dernier passager, la place libre est donc la place 1 ou la place n. Or, à chaque choix au hasard, tant qu'aucune des deux n'est prise, elles sont toutes deux libres et ont la même probabilité d'être choisies : celle qui est prise la première l'est avec probabilité $\dfrac{1}{2}$ pour chacune, et le problème se résout au moment de ce choix. La récurrence ci-dessus est la forme rigoureuse de cet argument de symétrie.

EXERCICE 108 ⋆⋆⋆⋆

Idée. On découpe $\mathbb{N}^*$ en blocs $[\![m^2\,,(m+1)^2-1]\!]$, sur lesquels la partie entière de $\sqrt{k}$ vaut m. Dans un bloc, la première décimale de $\sqrt{k}$ vaut 1 lorsque $(m+0{,}1)^2 \leqslant k < (m+0{,}2)^2$: environ $0{,}2m$ entiers sur les $2m+1$ du bloc, soit un dixième. Le bloc en cours n'a qu'environ $2\sqrt{n}$ éléments, négligeable devant n.

Le cadre. On a $p_n = \dfrac{N(n)}{n}$, où $N(n)$ est le nombre d'entiers $k \in [\![1\,,n]\!]$ dont la première décimale de $\sqrt{k}$ vaut 1, c'est-à-dire tels que $\sqrt{k} - \lfloor\sqrt{k}\rfloor \in [0{,}1\,,0{,}2[$.

Comptage dans un bloc. Soit $m \geqslant 1$ et c_m le nombre d'entiers k tels que $\lfloor\sqrt{k}\rfloor = m$ et dont la première décimale de $\sqrt{k}$ vaut 1. Par croissance stricte de la racine carrée, ce sont les entiers de l'intervalle

$$[(m+0{,}1)^2\,,(m+0{,}2)^2[\;=\; [m^2+0{,}2m+0{,}01\,,m^2+0{,}4m+0{,}04[,$$

de longueur $0{,}2m+0{,}03$. Un intervalle $[\alpha\,,\beta[$ de longueur λ contient $\lceil\beta\rceil - \lceil\alpha\rceil$ entiers, nombre compris strictement entre $\lambda-1$ et $\lambda+1$. Ainsi

$$\frac{m}{5} - 1 < c_m < \frac{m}{5} + 2,$$

ce que confirme le calcul pour $m < 3\,000$ ✓.

Encadrement. Soit $n \geqslant 1$ et $M = \lfloor\sqrt{n}\rfloor \geqslant 1$, de sorte que $M^2 \leqslant n < (M+1)^2$. Les blocs $m = 1$, …, $M-1$ sont contenus dans $[\![1\,,n]\!]$, et $[\![1\,,n]\!]$ est contenu dans la réunion des blocs $m = 1$, …, M. Avec la somme des entiers du chapitre 3,

$$N(n) \geqslant \sum_{m=1}^{M-1}\Big(\frac{m}{5}-1\Big) = \frac{M(M-1)}{10} - (M-1), \qquad N(n) \leqslant \sum_{m=1}^{M}\Big(\frac{m}{5}+2\Big) = \frac{M(M+1)}{10} + 2M.$$

En divisant par n, avec $M^2 \leqslant n < (M+1)^2$, et pour $M \geqslant 10$, de sorte que le minorant de $N(n)$ soit positif,

$$\frac{M(M-1)-10(M-1)}{10(M+1)^2} \leqslant p_n \leqslant \frac{M(M+1)+20M}{10M^2}.$$

Conclusion. Quand n tend vers $+\infty$, $M > \sqrt{n}-1$ tend vers $+\infty$, et les deux membres extrêmes tendent vers $\dfrac{1}{10}$. Par le théorème des gendarmes,

$$\boxed{\text{la suite } (p_n) \text{ converge, et } \lim_{n\to+\infty} p_n = \frac{1}{10}}$$

Le calcul exact donne $p_{10^3} = 0{,}099$, $p_{10^5} = 0{,}100\,17$ et $p_{10^6} = 0{,}099\,9$ ✓. La suite n'est pas monotone : dans chaque bloc, elle augmente sur la portion favorable puis diminue, mais l'amplitude de ces oscillations est en $O(\dfrac{1}{M})$.

Remarque. Le même calcul montre que chacun des dix chiffres est, à la limite, la première décimale de $\sqrt{k}$ avec la probabilité $\dfrac{1}{10}$, puisque l'intervalle correspondant au chiffre c a pour longueur $0{,}2m + \dfrac{2c+1}{100}$. Ce n'est pas un phénomène général : la proportion des entiers de $[\![1\,,n]\!]$ dont le *premier chiffre* vaut 1 n'a pas de limite. Elle vaut environ $\dfrac{1}{9}$ pour $n = 10^6-1$, et environ $\dfrac{5}{9}$ pour $n = 2\times10^6-1$ ✓, car les blocs de nombres de même longueur ont des tailles qui croissent géométriquement : le dernier bloc n'est plus négligeable devant les précédents. Ici, c'est la croissance seulement linéaire de la longueur $2m+1$ des blocs qui rend le bloc en cours négligeable.

EXERCICE 109 ⋆⋆⋆⋆

Idée. Pour $n=1$, un calcul direct donne $\mathrm{Cov}\,(f(T_1),g(T_1)) = p(1-p)(f(1)-f(0))(g(1)-g(0)) \geqslant 0$. Pour passer de $n-1$ à n, on fixe la dernière coordonnée : à T_n fixé, les deux fonctions restent croissantes en les $n-1$ premières variables, et leurs moyennes en ces variables sont des fonctions croissantes de la dernière.

On note p_i le paramètre de T_i, et l'on doit montrer que $\mathrm{E}(fg) \geqslant \mathrm{E}(f)\,\mathrm{E}(g)$, où $\mathrm{E}(f)$ abrège $\mathrm{E}(f(T_1,\dots,T_n))$ (proposition 23 (i)). Pour $t' \in \{0,1\}^{n-1}$ et $x \in \{0,1\}$, on note (t',x) l'élément de $\{0,1\}^n$ obtenu en ajoutant x à la fin de t'.

Le cas $n=1$. Soit f et g croissantes sur $\{0,1\}$, et $p=p_1$, $q=1-p$. La formule de transfert donne

$$\mathrm{E}(fg) - \mathrm{E}(f)\,\mathrm{E}(g) = pf(1)g(1) + qf(0)g(0) - \big(pf(1)+qf(0)\big)\big(pg(1)+qg(0)\big).$$

En utilisant $p-p^2 = q-q^2 = pq$, on obtient

$$\mathrm{E}(fg) - \mathrm{E}(f)\,\mathrm{E}(g) = pq\big(f(1)g(1)+f(0)g(0)-f(1)g(0)-f(0)g(1)\big) = pq\big(f(1)-f(0)\big)\big(g(1)-g(0)\big),$$

qui est positif, puisque $f(1) \geqslant f(0)$ et $g(1) \geqslant g(0)$.

Un lemme de « conditionnement ». Soit $n \geqslant 2$, $T' = (T_1,\dots,T_{n-1})$ et $h : \{0,1\}^n \to \mathbb{R}$. Les variables T' et T_n sont indépendantes (théorème 7), donc la loi du couple (T',T_n) est le produit de ses lois marginales (proposition 15), et la formule de transfert donne

$$\mathrm{E}\big(h(T',T_n)\big) = \sum_{t',x} h(t',x)\,\mathrm{P}(T'=t')\,\mathrm{P}(T_n=x) = \sum_{x\in\{0,1\}} \mathrm{P}(T_n=x)\,\mathrm{E}\big(h(T',x)\big). \qquad \text{(L)}$$

C'est une moyenne pondérée, sur les valeurs de T_n, d'espérances calculées à T_n fixé.

Hérédité. Soit $n \geqslant 2$; supposons le résultat vrai pour $n-1$ variables de Bernoulli indépendantes quelconques. Soit f et g croissantes sur $\{0,1\}^n$. Pour $x \in \{0,1\}$, les fonctions $f_x : t' \mapsto f(t',x)$ et $g_x : t' \mapsto g(t',x)$ sont croissantes sur $\{0,1\}^{n-1}$, et les variables T_1, …, T_{n-1} sont mutuellement indépendantes (proposition 15). L'hypothèse de récurrence donne

$$\mathrm{E}\big(f_x(T')\,g_x(T')\big) \geqslant F(x)\,G(x), \qquad \text{où} \quad F(x) = \mathrm{E}\big(f(T',x)\big), \quad G(x) = \mathrm{E}\big(g(T',x)\big).$$

Les fonctions F et G sont croissantes sur $\{0,1\}$: pour tout t', $(t',0) \leqslant (t',1)$, donc $f(t',0) \leqslant f(t',1)$, et la croissance de l'espérance donne $F(0) \leqslant F(1)$; de même pour G. Le lemme (L), appliqué à $h=fg$, puis l'hypothèse de récurrence, multipliée par les poids positifs $\mathrm{P}(T_n=x)$, donnent

$$\mathrm{E}(fg) = \sum_x \mathrm{P}(T_n=x)\,\mathrm{E}\big(f_x(T')\,g_x(T')\big) \geqslant \sum_x \mathrm{P}(T_n=x)\,F(x)G(x) = \mathrm{E}\big(F(T_n)\,G(T_n)\big).$$

Le cas $n=1$, appliqué à F, G et à la variable T_n, donne ensuite

$$\mathrm{E}\big(F(T_n)\,G(T_n)\big) \geqslant \mathrm{E}\big(F(T_n)\big)\,\mathrm{E}\big(G(T_n)\big) = \mathrm{E}(f)\,\mathrm{E}(g),$$

la dernière égalité étant le lemme (L) appliqué à $h=f$ et à $h=g$. Ainsi $\mathrm{E}(fg) \geqslant \mathrm{E}(f)\,\mathrm{E}(g)$, et, par récurrence sur n,

$$\boxed{\mathrm{Cov}\,(f(T_1,\dots,T_n),g(T_1,\dots,T_n)) \geqslant 0 \ \text{ pour toutes } f,g \text{ croissantes}}$$

L'inégalité a été testée sur 300 couples de fonctions croissantes tirées au hasard, avec $n \leqslant 4$ et des paramètres quelconques ✓.

Remarques.

— L'hypothèse d'indépendance est indispensable. Si T_1 suit la loi $\mathcal{B}(\frac{1}{2})$ et $T_2 = 1-T_1$, les variables T_1 et T_2 sont croissantes, puisque ce sont les images de (T_1,T_2) par les deux projections, qui sont croissantes pour l'ordre produit. Mais $T_1T_2 = 0$, et $\mathrm{Cov}(T_1,T_2) = 0 - \frac{1}{4} = -\frac{1}{4}$.

— Avec des indicatrices, le résultat s'écrit $\mathrm{P}(A\cap B)\geqslant \mathrm{P}(A)\,\mathrm{P}(B)$ pour deux événements croissants, c'est-à-dire stables quand on fait passer des coordonnées de 0 à 1. Pour le graphe aléatoire de la section 21.6.3, dont les arêtes sont des variables de Bernoulli indépendantes, « le graphe contient un triangle » et « le graphe est connexe » sont croissants : ces deux événements sont positivement corrélés, sans aucun calcul.

— Le cas $n=1$ est celui de l'exemple du cours qui suit la proposition 23 : deux fonctions croissantes d'une même variable ont une covariance positive. L'inégalité de Harris l'étend à plusieurs variables indépendantes ; elle a été généralisée par Fortuin, Kasteleyn et Ginibre à des lois non produits (inégalité FKG).

EXERCICE 110 ⋆⋆⋆⋆

Idée. Le produit $i_1\cdots i_p$ compte les façons de marquer un objet dans chacun de p blocs consécutifs de tailles i_1, …, i_p. On code un découpage de n objets alignés, avec un objet marqué par bloc, en insérant $p-1$ cases séparatrices : on obtient une rangée de $n+p-1$ cases, dans laquelle les p cases marquées et les $p-1$ séparateurs forment $2p-1$ cases qui alternent.

Les configurations. Si l'un des i_k est nul, le produit est nul : la somme porte en fait sur les p-uplets d'entiers $i_k\geqslant 1$ de somme n. Pour un tel p-uplet, le produit $i_1\cdots i_p$ est le nombre de p-uplets $(j_1,\dots,j_p)$ tels que $1\leqslant j_k\leqslant i_k$ (proposition 2 (iv)). Le membre de gauche est donc le cardinal de l'ensemble $\mathcal{C}$ des *configurations* $((i_k),(j_k))$, avec $i_k\geqslant 1$, $\sum i_k=n$ et $1\leqslant j_k\leqslant i_k$: on range n objets en p blocs consécutifs non vides de tailles i_1, …, i_p, et l'on marque l'objet de rang j_k dans le bloc k.

Le codage. On note $I_k=i_1+\cdots+i_k$, avec $I_0=0$, et l'on associe à une configuration les entiers

$$c_{2k-1}=I_{k-1}+(k-1)+j_k\quad(1\leqslant k\leqslant p),\qquad c_{2k}=I_k+k\quad(1\leqslant k\leqslant p-1).$$

Dans la rangée de $n+p-1$ cases, où l'on a inséré une case séparatrice après chacun des $p-1$ premiers blocs, c_{2k} est la position du k-ième séparateur et c_{2k-1} celle de l'objet marqué du bloc k. On a $1\leqslant c_1=j_1$, $c_{2p-1}\leqslant I_p+p-1=n+p-1$, et la suite est strictement croissante :

$$c_{2k-1}\leqslant I_k+k-1<I_k+k=c_{2k},\qquad c_{2k}=I_k+k<I_k+k+j_{k+1}=c_{2k+1}.$$

On définit ainsi une application Φ de $\mathcal{C}$ dans l'ensemble des parties à $2p-1$ éléments de $[\![1\,,n+p-1]\!]$.

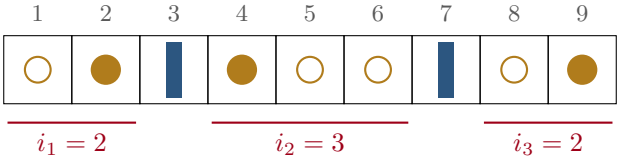
La réciproque. Soit $c_1<\cdots<c_{2p-1}$ des éléments de $[\![1\,,n+p-1]\!]$, et $c_0=0$, $c_{2p}=n+p$. On pose

$$i_k=c_{2k}-c_{2k-2}-1,\qquad j_k=c_{2k-1}-c_{2k-2}\qquad(1\leqslant k\leqslant p).$$

Comme $c_{2k}-c_{2k-2}\geqslant 2$, on a $i_k\geqslant 1$; on a aussi $1\leqslant j_k\leqslant c_{2k}-1-c_{2k-2}=i_k$, et la somme télescopique donne $\sum i_k=c_{2p}-c_0-p=n$. On obtient une configuration, et l'on vérifie que les deux constructions sont réciproques l'une de l'autre : partant d'une configuration, $c_{2k}-c_{2k-2}-1=i_k$ et $c_{2k-1}-c_{2k-2}=j_k$; partant des c, on retrouve $I_k+k=c_{2k}$ en sommant, puis c_{2k-1}. L'application Φ est une bijection (chapitre 2), et le théorème 2 donne

$$\boxed{\sum_{\substack{(i_1,\dots,i_p)\in\mathbb{N}^p\\ i_1+\cdots+i_p=n}} i_1i_2\cdots i_p=|\mathcal{C}|=\binom{n+p-1}{2p-1}}$$

avec la convention $\binom{n+p-1}{2p-1}=0$ lorsque $p>n$: les deux membres sont alors nuls, puisque p blocs non vides demandent au moins p objets.



Pour $n = 7$ et $p = 3$, la configuration formée des blocs de tailles $(2, 3, 2)$ et des rangs marqués $(2, 1, 2)$ est codée par les cases $\{2, 3, 4, 7, 9\}$ parmi 9 : les séparateurs, en bleu, occupent les cases 3 et 7, et les objets marqués, en ocre plein, les cases 2, 4 et 9. Pour $n = 4$ et $p = 2$, on trouve $0 \times 4 + 1 \times 3 + 2 \times 2 + 3 \times 1 + 4 \times 0 = 10 = \binom{5}{3}$, et l'identité a été vérifiée pour $n \leqslant 10$ et $p \leqslant 5$, ainsi que la bijection pour quelques couples (n, p) ✓.

À retenir. Pour $p = 2$, on obtient $\sum_{i=0}^{n} i(n-i) = \binom{n+1}{3}$, que l'on retrouve par le calcul avec les sommes du chapitre 3. Ranger des objets et des séparateurs dans une même rangée est aussi la façon de compter les p-uplets d'entiers naturels de somme n, au nombre de $\binom{n+p-1}{n}$: c'est la case hors programme du tableau de la section 21.2.6, où l'on choisit n éléments parmi p, avec répétitions et sans ordre.