Modélisation aléatoire

Labos de Maths

13 juin 2019

Résumé

Ce document regroupe plusieurs jeux de hasard basés sur des lancers de dés ou de pièces. Ils ont inspiré des activités variées en collège et lycée en lien avec de nombreuses thématiques au programme : probabilités, statistique, algorithmique, programmation en Scratch et Python, étude de suites.

Table des matières

1 Introduction

Ce document présente plusieurs jeux aléatoires autour du temps de premier succès dans une succession d'expériences aléatoires indépendantes et de même loi de Bernoulli : combien de fois doit-on lancer une pièce pour obtenir le premier "Pile" ou un dé à six faces pour obtenir le premier "6" ? Cette question très naturelle en pratique n'est pas évidente dans la mesure où cette variable aléatoire peut prendre des valeurs entières arbitrairement grandes avec une probabilité strictement positive : il est possible, bien que très peu probable, d'obtenir une suite de mille "Face" consécutifs avec une pièce équilibrée.
C'est sans doute pour cette raison que cette notion de temps de premier succès ne figure pas dans les programmes du secondaire. Il est toutefois possible de réfléchir à ce problème dès le collège ou le lycée par le biais de jeux, de manipulations et des simulations en Scratch ou Python par exemple.
Dans la section 2, nous présentons quelques jeux de dés amusants qui permettent de familiariser les élèves à l'aléatoire. En jouant plusieurs parties, ils peuvent se faire une idée assez juste des probabilités en jeu dans ces différentes situations. Pour aller plus loin, il est possible d'utiliser la simulation et bien entendu des calculs explicites pour les élèves qui y sont prêts.
Dans la section 4, le temps de premier succès dans un schéma de Bernoulli est étudié et sa loi, dite loi géométrique, est explicitée. Une première mise en situation est présentée avec la question de "construire un dé" à cinq faces équilibré à partir d'un dé classique.
La section 3 présente enfin le jeu dit de la ruine du joueur qui est construit à partir d'un schéma de Bernoulli.
Toutes ces situations se prêtent très bien à des activités différentiées dans la mesure où elles peuvent toutes être approchées de plusieurs manières complémentaires (jeu, simulation, calcul).

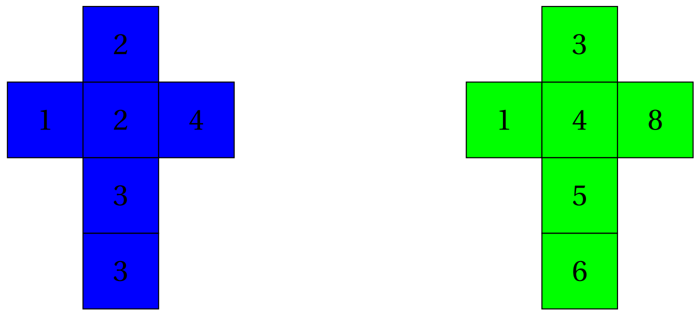
2 Quelques jeux de dés

Cette section présente quelques dés à six faces amusants. Les dés sont équilibrés au sens où toutes leurs faces apparaissent avec la même probabilité $1/6$ mais la numérotation des faces est inhabituelle… Ces jeux contre-intuitifs peuvent être abordés de multiples façons :

- par la conjecture (même si l'intuition est parfois trompeuse),
- par le jeu (répétition d'expériences, mise en commun au sein de la classe),
- par le calcul (étude de tous les cas possibles),
- par la simulation (en Scratch ou Python mais sans doute difficile au collège).

2.1 Dés de Sicherman

Les dés de Sicherman sont les dés à six faces suivants.



Jeu. Le joueur A lance les deux dés de Sicherman, le joueur B lance deux dés usuels et chacun retient la somme de ses deux résultats. Un joueur est gagnant si sa somme est strictement supérieure à celle de son adversaire. La partie est nulle en cas d'égalité.
Question. Quel joueur est-il avantagé ?
Le calcul de la distribution de la somme de deux dés normaux est classique. Le nombre d'occurrences des différents résultats possibles est rassemblé dans le tableau ci-dessous.

somme	2	3	4	5	6	7	8	9	10	11	12
occurrences	1	2	3	4	5	6	5	4	3	2	1

Pour les dés de Sicherman, la réponse est moins évidente. On peut remarquer que leur somme peut prendre toutes les valeurs entre 2 et 12 et uniquement celles-ci. Pour déterminer les probabilités d'apparition de chacun de ces entiers, le plus simple est de faire un tableau recensant tous les cas possibles.

Vert \ Bleu	1	2	2	3	3	4
1	2	3	3	4	4	5
3	4	5	5	6	6	7
4	5	6	6	7	7	8
5	6	7	7	8	8	9
6	7	8	8	9	9	10
8	9	10	10	11	11	12

On remarque alors que la distribution de la somme des dés de Sicherman est la même que celle de dés normaux. Les deux joueurs sont donc à égalité même si leurs dés ne sont pas les mêmes !
Lors d'un match "Classique vs Sicherman", la probabilité d'un match nul est donnée par la somme des probabilités que les deux sommes soient égales à 2, 3,... ou 12. Par indépendance des deux adversaires, on a donc

$$\begin{aligned}\mathbb{P}(\text{Match nul}) &= \left(\frac{1}{36}\right)^2 + \left(\frac{2}{36}\right)^2 + \cdots + \left(\frac{5}{36}\right)^2 + \left(\frac{6}{36}\right)^2 + \left(\frac{5}{36}\right)^2 + \cdots + \left(\frac{1}{36}\right)^2 \\ &= \frac{146}{1296} = \frac{73}{648}.\end{aligned}$$

Ainsi, la probabilité que les dés classiques battent ceux de Sicherman sur une manche est donnée par

$$p = \frac{1 - \mathbb{P}(\text{Match nul})}{2} = \frac{575}{1296}.$$

Simulation. Plusieurs notions peuvent être utilisées :

— simulation d'un dé équilibré à 6 faces

— boucle Si pour construire un dé de Sicherman à partir d'un dé classique et pour déterminer le gagnant ou la partie nulle à chaque manche

— compteur pour garder en mémoire le résultat de chaque manche

— représentation graphique de la proportion courante de chaque issue.

Voici l'algorithme pour une manche : on lance 4 dés classiques, les deux derniers sont retouchés pour former des dés Sicherman puis on conclut. On a volontairement évité les boucles Si imbriquées mais elles peuvent être utilisées pour raccourcir un peu l'algorithme.

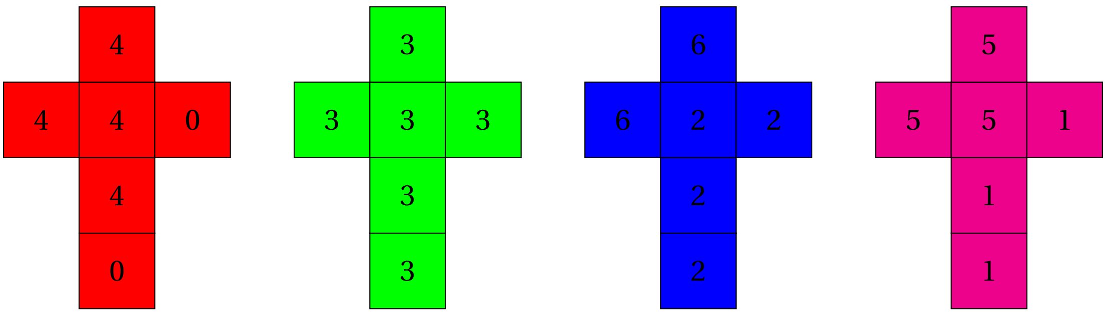
Algorithme pour une manche "Classique contre Sicherman"

```
A=Lancer d'un dé classique
B=Lancer d'un dé classique
C=Lancer d'un dé classique
D=Lancer d'un dé classique
Si C=5 alors C=2
Si C=6 alors C=3
Si D=2 alors D=8
Si A+B>C+D alors Classique gagne
Si A+B=C+D alors Match nul
Si A+B<C+D alors Sicherman gagne
```

Pour aller plus loin. On peut montrer que les dés classiques et les dés de Sicherman sont les seuls à posséder cette distribution pour la somme de leurs résultats.

2.2 Dés d'Efron

Les dés d'Efron sont les quatre dés ci-dessous.



Jeu. Le premier joueur choisit un des quatre dés. Le second joueur choisit alors un autre dé. Chacun lance son dé et le gagnant est celui qui obtient le score le plus grand. Remarquons qu'il n'y a pas d'égalité possible.
Question. Le jeu est-il équilibré entre les deux joueurs ?
Le second joueur a l'avantage de sélectionner son dé en connaissant le choix de son adversaire. On peut se demander si cette dissymétrie est décisive. Il s'avère qu'elle l'est effectivement. Les dés d'Efron sont dits non-transitifs : chaque dé est battu par un autre comme dans le jeu "Pierre, feuille, ciseaux".

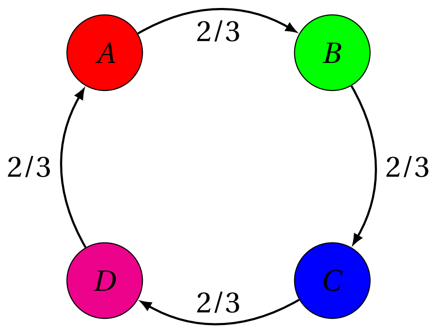
— Le dé A gagne sur le dé B si et seulement si A donne 4, donc avec probabilité $2/3$.

— Le dé B gagne sur le dé C si et seulement si C donne 2, donc avec probabilité $2/3$.

— Le dé C gagne sur le dé D si et seulement si C donne 6 ou C donne 2 et D donne 1, donc avec probabilité $1/3 + 2/3 \times 1/2 = 2/3$.

— Le dé D gagne sur le dé A si et seulement si D donne 5, ou D donne 1 et A donne 0, donc avec probabilité $1/2 + 1/2 \times 1/3 = 2/3$.

On peut résumer la dépendance cyclique par le graphe ci-dessous.



Pour aller plus loin. On peut également chercher les autres rapports de force :

— la probabilité que A batte C est $4/9$;

— la probabilité que B batte D est égale à $1/2$;

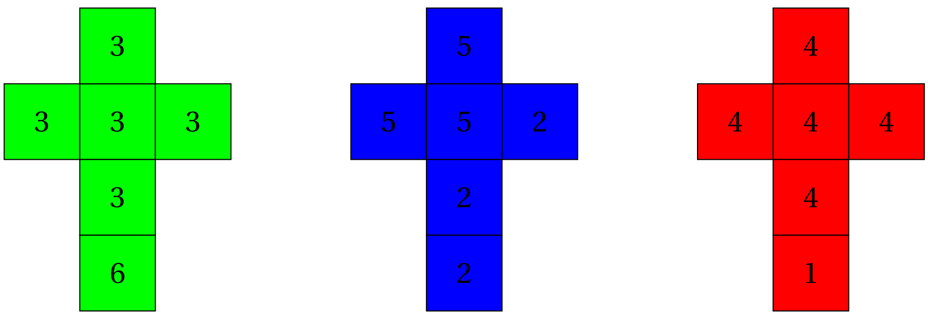
— la probabilité que C batte A vaut $5/9$;

— la probabilité que D batte B est $1/2$.

En revanche, la probabilité qu'un dé en batte un autre pris au hasard parmi les trois restants n'est pas égale suivant les dés. Elle vaut $13/27$ pour A, $1/2$ pour B, $14/27$ pour C et $1/2$ pour D.
Globalement, le meilleur dé pour gagner un jeu totalement aléatoire est donc C, qui gagne dans près de 52% des cas.

2.3 Dés du diable

Les dés du diable sont les trois dés suivants.



Bernard Delyon présente ces dés dans une vidéo disponible sur Youtube dans le cadre des "Cinq minutes Lebesgue" :

`https://www.youtube.com/watch?v=LbGTGZ_Gm44.`

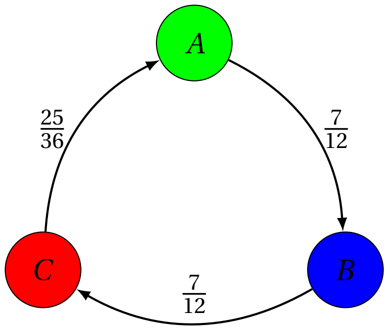
Les tableaux ci-dessous recensent, pour chaque paire de dés, le dé vainqueur en fonction des deux résultats et permettent alors de déterminer les probabilités de gain d'un dé sur un autre.

A \ B	2	2	2	5	5	5
3	A	A	A	B	B	B
3	A	A	A	B	B	B
3	A	A	A	B	B	B
3	A	A	A	B	B	B
3	A	A	A	B	B	B
6	A	A	A	A	A	A

B \ C	1	4	4	4	4	4
2	B	C	C	C	C	C
2	B	C	C	C	C	C
2	B	C	C	C	C	C
5	B	B	B	B	B	B
5	B	B	B	B	B	B
5	B	B	B	B	B	B

C \ A	3	3	3	3	3	6
1	A	A	A	A	A	A
4	C	C	C	C	C	A
4	C	C	C	C	C	A
4	C	C	C	C	C	A
4	C	C	C	C	C	A
4	C	C	C	C	C	A

On constate que comme les dés d'Efron, ces dés sont non-transitifs. Le diagramme ci-dessous indique les probabilités de victoire d'un dé sur l'autre.



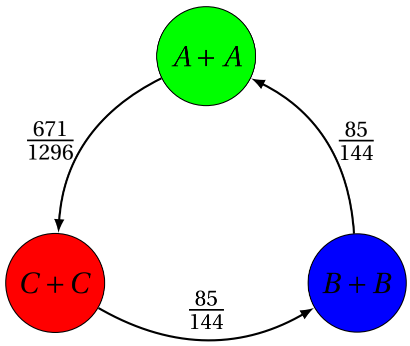
Cependant, modifions légèrement le jeu : au lieu de lancer le dé une fois, on le lance deux fois de suite et on retient la somme des deux résultats. Pour déterminer les probabilités de victoire, il faut expliciter la loi de la somme de deux dés identiques. Les distributions sont données dans les tableaux ci-dessous :

Lancer de deux dés A			
somme	6	9	12
occurrences	25	10	1

Lancer de deux dés B			
somme	4	7	10
occurrences	9	18	9

Lancer de deux dés C			
somme	2	5	8
ocurrences	1	10	25

On remarque que les matchs ne donnent pas lieu à des égalités et on obtient alors les relations suivantes :



Pour aller plus loin. Que se passe-t-il si on ajoute les résultats de trois lancers consécutifs d'un même dé ? Le calcul devient extrêmement fastidieux mais la simulation peut aider… Attention toutefois aux cas d'égalité !

3 Ruine du joueur

3.1 Description du jeu

Soit a un entier strictement positif et x un entier compris entre 0 et a. Le joueur possède initialement x jetons. Il répète le jeu suivant jusqu'à ce que sa fortune (en jetons) soit égale à 0 ou a : il lance une pièce équilibrée et gagne 1 jeton si le résultat est "Pile" et perd 1 jeton si le résultat est "Face".
On s'intéresse à deux quantités associées à ce jeu : la probabilité de finir le jeu avec a jetons et la durée moyenne du jeu. Ces deux quantités dépendent du nombre initial x de jetons. Ce sont donc des fonctions de $\{0,1,\ldots,a\}$ dans $[0,1]$ et $\mathbb{R}_+$ respectivement. Elles seront notées

$$v : \{0,1,\ldots,a\} \to [0,1] \quad \text{et} \quad t : \{0,1,\ldots,a\} \to \mathbb{R}_+$$

respectivement.

3.2 Cas particulier

Considérons dans un premier temps le cas $a = 3$ et $x = 1$. Quelle est la probabilité de finir le jeu avec 3 jetons ?

— Premier lancer : si le joueur fait "Face" la partie est perdue ; s'il fait "Pile", il a deux jetons et continue le jeu.

— Deuxième lancer : si le joueur fait "Pile", la partie est gagnée ; s'il fait "Face", il a à nouveau un jeton et revient dans la situation de départ.

Si q est la probabilité de gagner à ce jeu en démarrant avec un jeton alors

$$q = \frac{1}{2}\left(\frac{1}{2} + q\right) \quad \text{ou encore} \quad q = \frac{1}{3}.$$

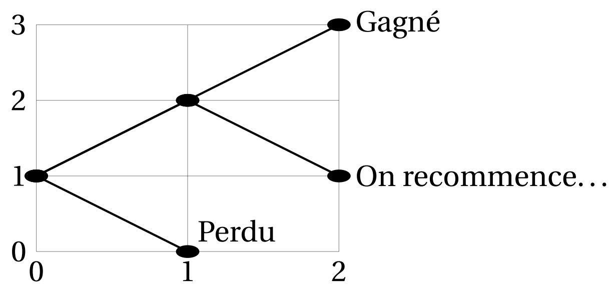


FIGURE 1 – Les deux premiers lancers de la partie.

3.3 Simulation

3.3.1 Simulation d'une partie

L'algorithme repose sur une boucle du type "Tant que". Il faut utiliser un compteur pour connaître le nombre de lancers nécessaires. Si l'on s'intéresse seulement au résultat, on

peut oublier le chemin parcouru et ne garder en mémoire à chaque étape que la position courante et le nombre de lancers déjà effectués.
Dans l'algorithme ci-dessous, c est le compteur, x est le point de départ, a la fortune recherchée et `rand` désigne une variable aléatoire de loi de Bernoulli de paramètre $1/2$ (elle vaut donc 0 ou 1 avec probabilité $1/2$) :

```
c=0
X=x
Tant que X*(a-X)≠0
  c=c+1
  X=X+2*rand-1
Fin Tant que
```

À la fin de l'algorithme, X est le point de sortie (0 ou a) et c est le nombre d'étapes nécessaire à la partie.

Remarque 1. Si $\mathbb{P}(Y=1)=\mathbb{P}(Y=0)=1/2$ alors $\mathbb{P}(2Y-1=1)=\mathbb{P}(2Y-1=-1)=1/2$.

Si l'on appelle un grand nombre de fois cette procédure, on peut estimer la probabilité de victoire et et le temps moyen d'une partie (en fonction du point de départ). Plus précisément, on simule N jeux successifs en retenant à chaque fois le résultat et le nombre de lancers du jeu. Pour le jeu i, on pose $V_i=1$ si le jeu est gagnant et $V_i=0$ dans le cas contraire et T_i le nombre de lancers du jeu i.

3.3.2 Estimation de la probabilité de gagner

Les variables aléatoires $(V_i)_{1\leq i\leq N}$ sont indépendantes et de loi de Bernoulli de paramètre p égal à la probabilité que le jeu soit gagnant. Ce paramètre dépend de la fortune initiale x et est noté $v(x)$. La somme $V_1+V_2+\cdots+V_N$ est le nombre de victoires pour N parties indépendantes de même probabilité de succès. Elle est donc de loi binomiale $\mathscr{B}(N,v(x))$. On peut estimer $v(x)$ par

$$\hat{v}_N(x)=\frac{1}{N}\sum_{i=1}^{N}V_i.$$

La loi des grands nombres assure que $\hat{v}_N(x)$ converge en probabilité vers $v(x)$ lorsque N tend vers l'infini. On peut préciser cette estimation en proposant un intervalle de confiance pour la probabilité $v(x)$...

3.3.3 Estimation du temps moyen de jeu

Les variables aléatoires $(T_i)_{1\leq i\leq N}$ sont indépendantes et de loi inconnue. L'espérance de cette loi dépend de la fortune initiale x et est noté $t(x)$. On peut estimer $t(x)$ par

$$\hat{t}(x)=\frac{1}{N}\sum_{i=1}^{N}T_i.$$

La loi des grands nombres assure que $\hat{t}(x)$ converge en probabilité vers $t(x)$ lorsque N tend vers l'infini. On peut préciser cette estimation en proposant un intervalle de confiance pour $t(x)$...

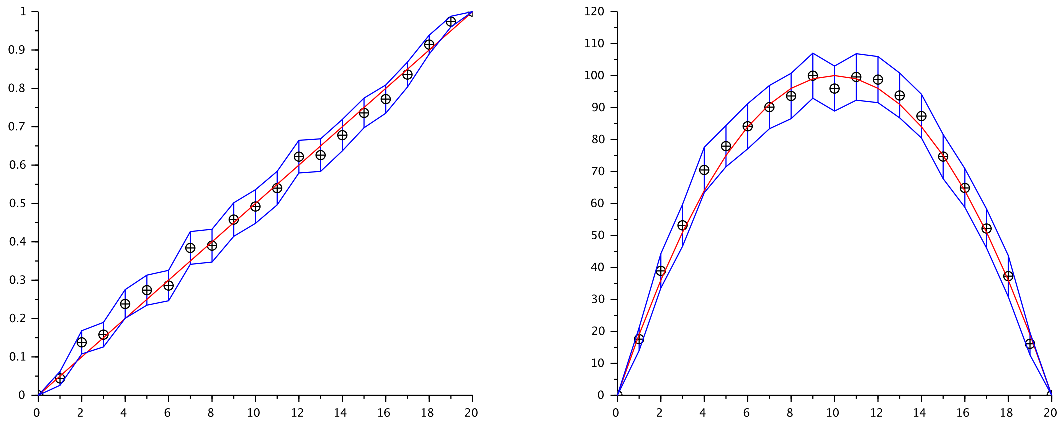


FIGURE 2 – Estimation de la probabilité de victoire (à gauche) et du temps moyen de jeu (à droite) pour $a = 20$ à partir de 500 jeux.

3.4 Expression explicite

Théorème 2. *On a, pour* $x \in \{0, 1, \dots, a\}$,

$$v(x) = \frac{x}{a} \quad \textit{et} \quad t(x) = x(a - x).$$

DÉMONSTRATION. Si $1 \leq x \leq a - 1$, alors

$$v(x) = \frac{1}{2} v(x-1) + \frac{1}{2} v(x+1). \tag{1}$$

Cette relation s'obtient en raisonnant sur la première partie jouée lorsque l'on possède initialement x jetons. Avec probabilité $1/2$, un jeton est gagné et tout se passe alors comme si un nouveau jeu commençait mais avec $x + 1$ jetons. Sinon, on perd un jeton et tout se passe cette fois-ci comme si le jeu commençait avec $x - 1$ jetons. L'équation (1) assure que $v(x+1) - v(x) = v(x) - v(x-1)$ pour tout $1 \leq x \leq a-1$. En d'autres termes, les accroissements de la fonction v sont constants. Elle est donc affine. On a bien entendu $v(0) = 0$ et $v(a) = 1$ ce qui permet de conclure.
On procède de manière similaire pour le temps moyen de jeu. On a bien entendu $t(0) = 0$ et $t(a) = 0$. De plus, si $1 \leq x \leq a - 1$,

$$t(x) = 1 + \frac{1}{2} t(x-1) + \frac{1}{2} t(x+1).$$

On peut alors vérifier que la fonction $x \mapsto x(a - x)$ est solution de ce système d'équations. □

3.5 Pour aller plus loin

Dans les casinos, lieux que tout probabiliste sait qu'il convient d'éviter, les jeux ne sont pas équilibrés. Par exemple, si l'on joue à "Pair" ou "Impair" à la roulette, les chances de gagner sont de $18/37$ car lorsque la bille tombe sur le 0, a banque empoche toutes les mises. Il

est donc intéressant d'examiner le cas où la victoire à chaque étape du jeu survient avec probabilité $p \in [0,1]$ et non plus seulement $p = 1/2$. La simulation fonctionne exactement de la même façon que dans le cas équilibré une fois que l'on a modifié la probabilité de gain.
Il est également possible d'obtenir des expressions explicites pour la probabilité de sortir avec a jetons et le temps moyen de jeu !

Théorème 3. *Si la probabilité de gain vaut $p \neq 1/2$ et $q = 1-p$, on a, pour $x \in \{0,1,\dots,a\}$,*

$$v(x) = \frac{\left(\frac{p}{q}\right)^{a-x} - \left(\frac{p}{q}\right)^{a}}{1-\left(\frac{p}{q}\right)^{a}} \quad \textit{et} \quad t(x) = \frac{x}{q-p} - \frac{a}{q-p}\frac{1-\left(\frac{q}{p}\right)^{x}}{1-\left(\frac{q}{p}\right)^{a}}.$$

DÉMONSTRATION. On a bien entendu $v(0) = 0$ et $v(a) = 1$. De plus, si $1 \le x \le a-1$,

$$v(x) = qv(x-1) + pv(x+1).$$

On a bien entendu $t(0) = 0$ et $t(a) = 0$. De plus, si $1 \le x \le a-1$,

$$t(x) = 1 + qt(x-1) + pt(x+1).$$

$\square$

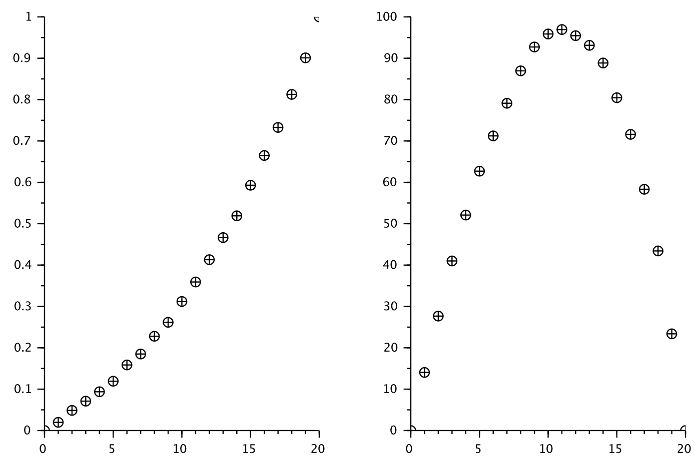


FIGURE 3 – Estimation de la probabilité de victoire (à gauche) et du temps moyen de jeu (à droite) pour $a = 20$ et $p = 0.48$ à partir de 1000 jeux.

4 Temps de premier succès

On répète une même expérience aléatoire dont les seules issues sont un succès ou un échec :

— lancer d'une pièce pour faire "Face",

— lancer d'un dé à six faces pour obtenir 6,

— deviner la première carte d'un jeu bien mélangé...

de manière indépendante et l'on s'intéresse au temps de premier succès.

4.1 La loi géométrique

La modélisation probabiliste est la suivante. On se donne une suite de variables aléatoires $(X_n)_{n\geq 1}$ indépendantes et de même distribution de Bernoulli de paramètre $p \in [0,1]$: pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = p.$$

La tentative i est un succès si X_i vaut 1 et un échec sinon. Le réel p représente la probabilité de succès en une tentative. Le temps de premier succès T est alors défini comme le plus petit indice des termes de la suite valant 1. En d'autres termes,

$$T = \inf\{n \geq 1 \ : \ X_n = 1\}.$$

Théorème 4. *La variable aléatoire T est à valeurs dans $\mathbb{N}^*$ et, pour tout $n \in \mathbb{N}^*$,*

$$\mathbb{P}(T > n) = (1-p)^n \quad \textit{et} \quad \mathbb{P}(T = n) = p(1-p)^{n-1}.$$

On dit que T suit la distribution géométrique de paramètre p.

DÉMONSTRATION. Pour tout $n \in \mathbb{N}^*$, le temps de premier succès T est strictement supérieur à n si et seulement si les n premières tentatives sont des échecs, c'est-à-dire que $X_i = 0$ pour $1 \leq i \leq n$. On obtient donc

$$\mathbb{P}(T > n) = \mathbb{P}(X_1 = 0, X_2 = 0, \ldots, X_n = 0).$$

L'indépendance des variables aléatoires X_1, X_2,..., X_n entraine que

$$\mathbb{P}(X_1 = 0, X_2 = 0, \ldots, X_n = 0) = \mathbb{P}(X_1 = 0) \times \mathbb{P}(X_2 = 0) \times \cdots \times \mathbb{P}(X_n = 0).$$

On obtient ainsi l'expression de $\mathbb{P}(T > n)$. En remarquant que

$$\mathbb{P}(T = n) = \mathbb{P}(T > n-1) - \mathbb{P}(T > n),$$

on obtient que

$$\mathbb{P}(T = n) = p(1-p)^{n-1}$$

pour tout $n \in \mathbb{N}^*$. □

Corollaire 5. *On a aussi, pour tout $n \in \mathbb{N}^*$,*

$$\mathbb{P}(T \leq n) = 1 - \mathbb{P}(T > n) = 1 - (1-p)^n.$$

Remarque 6. Comme $\mathbb{P}(T=1)+\mathbb{P}(T=2)+\cdots+\mathbb{P}(T=n)=\mathbb{P}(T\leq n)$, on a

$$p+p(1-p)+\cdots+p(1-p)^n=1-(1-p)^n.$$

On retrouve ainsi la formule pour la somme de n termes consécutifs d'une suite géométrique (d'où le nom de distribution géométrique).

Une question naturelle est de déterminer la moyenne (si elle existe) de ce temps de premier succès. La difficulté vient du fait que T peut prendre une infinité de valeurs avec une probabilité strictement positive. Cela étant dit, on a envie d'écrire

$$\mathbb{E}(T)=\sum_{k=1}^{\infty}k\mathbb{P}(T=k)=\lim_{n\to\infty}\sum_{k=1}^{n}k\mathbb{P}(T=k).$$

Notons, pour $x\in[0,1[$,

$$f(x)=\sum_{k=0}^{n}x^k=1+x+x^2+\cdots+x^n=\frac{1-x^{n+1}}{1-x}.$$

On a, en calculant les dérivées de ces expressions,

$$f'(x)=\sum_{k=1}^{n}kx^{k-1}=1+2x+3x^2+\cdots+nx^{n-1}=\frac{-(n+1)x^n}{1-x}+\frac{1-x^{n+1}}{(1-x)^2}$$

Que se passe-t-il quand n tend vers l'infini ? À gauche, on voit apparaître une somme qui ressemble à celle que l'on cherche à calculer. À droite, il faut trouver la limite de nx^n. Il s'agit d'une forme indéterminée mais on peut se débrouiller sans la croissance comparée et le logarithme.

Théorème 7. *Pour tout $x\in[0,1[$, la suite $(u_n)_n$ de terme général $u_n=nx^n$ tend vers 0 quand n tend vers $+\infty$.*

DÉMONSTRATION. Soit $x\in[0,1[$ fixé.
Étape 1. Pour tout entier $n\geq 1$, $u_n\geq 0$ donc la suite $(u_n)_n$ est minorée par 0. De plus,

$$\frac{u_{n+1}}{u_n}=\frac{(n+1)x^{n+1}}{nx^n}=\frac{n+1}{n}x.$$

Dès que[1] $n\geq x/(1-x)$,

$$\frac{n+1}{n}x\leq 1 \quad \text{donc} \quad u_{n+1}\leq u_n.$$

La suite $(u_n)_n$ est donc décroissante à partir d'un certain rang[2]. La suite $(u_n)_n$ est minorée par 0 et décroissante à partir d'un certain rang, elle est donc convergente et sa limite l est positive ou nulle.
Étape 2. Pour tout $n\geq 1$,

$$u_{n+1}=(n+1)x^{n+1}=xnx^n+x^{n+1}.$$

Ainsi, la suite $(u_n)_n$ est une suite récurrente :

$$u_{n+1}=xu_n+x^{n+1}.$$

1. Vive les inéquations !
2. Si $x\leq 1/2$, alors la suite $(u_n)_n$ est tout simplement décroissante.

On peut alors passer à la limite quand n tend vers l'infini pour obtenir

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} u_{n+1} = \lim_{n\to\infty} (xu_n + x^{n+1}).$$

Puisque $x \in [0,1[$, $\lim_{n\to\infty} x^{n+1} = 0$. Ceci entraine que $l = lx$. On en déduit que la limite l est égale à 0. □

On obtient donc la limite suivante.

Corollaire 8. *Pour tout* $x \in [0,1[$,

$$\lim_{n\to\infty} \sum_{k=1}^{n} kx^{k-1} = \frac{1}{(1-x)^2}.$$

Remplaçons alors x par $1-p$ et multiplions l'égalité par p. On obtient

$$\lim_{n\to\infty} \sum_{k=1}^{n} kp(1-p)^{k-1} = \frac{1}{p}.$$

Ce calcul donne le résultat suivant.

Théorème 9. *Si T suit la loi géométrique de paramètre p alors son espérance vaut $\frac{1}{p}$.*

Corollaire 10. *En moyenne, il faut lancer deux fois une pièce pour obtenir le premier "Face" et six fois un dé pour obtenir le premier "6".*

4.2 Le dé à cinq faces

On dispose d'un dé équilibré à six faces et on voudrait l'utiliser pour choisir de manière équiprobable un nombre entier entre 1 et 5.

Remarque 11. C'est facile pour un nombre entier entre 1 et 2 ou entre 1 et 3.

L'algorithme est le suivant.

```
A=Lancer du dé
Tant que A=6
  A=Lancer du dé
Fin Tant que
Afficher A
```

Théorème 12. *La distribution du nombre A sorti de cet algorithme est la distribution uniforme sur* $\{1,2,3,4,5\}$. *D'autre part, le nombre T de lancers nécessaire pour obtenir un résultat inférieur ou égal à* 5 *suit la loi géométrique de paramètre* $5/6$. *Enfin, les variables aléatoires A et T sont indépendantes.*

DÉMONSTRATION. Soit $i \in \{1,2,3,4,5\}$ et $n \geq 1$, alors

$$\begin{aligned}
\mathbb{P}(A=i, T=n) &= \mathbb{P}(X_1 = 6, X_2 = 6, \dots, X_{n-1} = 6, X_n = i) \\
&= \frac{1}{6^n} \\
&= \frac{1}{5} \times \frac{5}{6}\left(1 - \frac{5}{6}\right)^{n-1} \\
&= \mathbb{P}(A=i)\mathbb{P}(T=n).
\end{aligned}$$

Cette relation détermine en même temps les lois de A et T et l'indépendance de ces deux variables aléatoires. □

Exercice 1. On lance un dé à six faces une première fois puis on le relance jusqu'à obtenir un nombre différent. On fait alors la somme des deux résultats. Quels sont les résultats possibles ? Quelle est la distribution de cette variable aléatoire ?

Pour aller plus loin. Soit A un ensemble sur lequel on sait générer des variables aléatoires $(X_n)_{n\geq 1}$ uniforme et B un sous-ensemble de A tel que $\mathbb{P}(X_1 \in B)$ soit un réel strictement positif. On peut construire des variables aléatoires uniformes sur B en ne conservant que les variables tombées dans B. Cet algorithme s'appelle la méthode du rejet. Il est très utilisé en pratique pour étudier des ensembles compliqués comme l'ensemble des permutations de n éléments par exemple.

5 Modèle de Wright-Fisher

On modélise la transmission d'un allèle dans une population idéale :

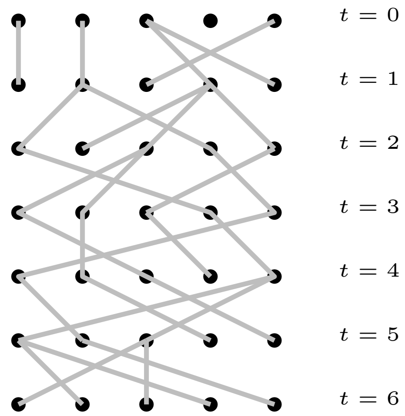
— la population est haploïde (chaque individu a un unique exemplaire du gène) ;

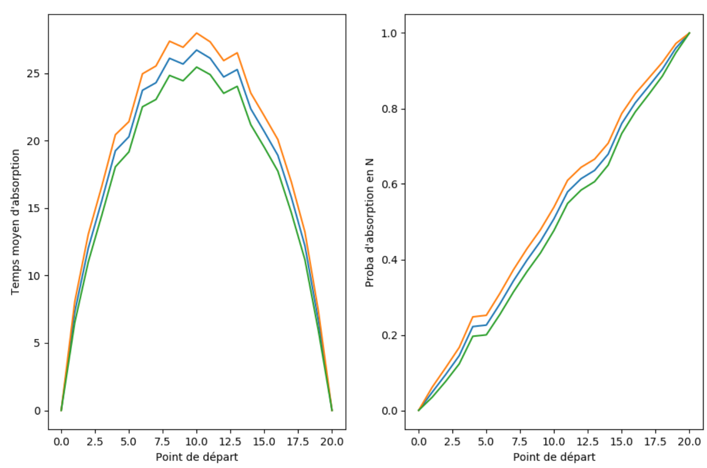
— la population est finie de taille constante égale à N ;

— le gène est présent sous la forme de deux allèles A et B ;

— chaque enfant choisi son parent uniformément parmi les individus de la génération précédent indépendamment des autres enfants ;

— il n'y a ni mutation, ni sélection.





6 Collectionneur de vignettes

On cherche à comprendre combien de vignettes il faut acheter pour finir une collection de N vignettes en supposant que les vignettes apparaissent successivement de manière indépendante et de selon la même loi uniforme.
On peut reformuler de ce problème en terme de dé. À combien de reprises doit-on lancer un dé à N faces pour les avoir toutes observées au moins une fois ?

7 Battage de cartes

Combien de fois doit-on battre un jeu de cartes pour qu'il soit bien mélangé ? Voici une question simple dont la réponse ne l'est pas. Pour apporter une réponse mathématique, il faut modéliser les battages successifs, expliquer ce que veut dire qu'un jeu est bien mélangé et être capable de dire qu'un jeu est « relativement bien mélangé ».
Un jeu de N cartes bien mélangé si les cartes sont rangées dans un ordre aléatoire. Il y a $N!$ façons de ranger les N cartes soit environ $8,06 \times 10^{67}$ pour un jeu de 52 cartes.
On considère un battage très simple : le battage par insertion. À chaque étape, la carte du dessus du paquet est positionnée au hasard à une des N places possibles. Combien de battages en moyenne pour que le jeu soit bien mélangé ?

8 Urnes et particules

Comment modéliser l'évolution de la répartition d'un gaz dans deux chambres contigües qui communiquent par une toute petite ouverture à partir d'une description microscopique (à l'échelle atomique) ?
On note a le nombre de particules total. Pour chaque instant n, on note X_n le nombre de particules dans le compartiment de gauche. On a

$$X_{n+1} = X_n + \begin{cases} 1 & \text{avec probabilité } \frac{a-X_n}{a}, \\ -1 & \text{avec probabilité } \frac{X_n}{a}. \end{cases}$$

Théorème 13. *Si $m_n = \mathbb{E}(X_n)$ alors*

$$m_{n+1} = \left(1 - \frac{2}{a}\right) m_n + 1.$$

On a donc

$$m_n = \frac{a}{2} + \left(1 - \frac{2}{a}\right)^n \left(m_0 - \frac{a}{2}\right) \xrightarrow[n\to\infty]{} \frac{a}{2}.$$

DÉMONSTRATION. On a

$$m_{n+1} = m_n + \frac{a - m_n}{a} - \frac{m_n}{a} = \left(1 - \frac{2}{a}\right) m_n + 1.$$

La suite $(m_n)_{n\geq 0}$ est arithmético-géométrique. Elle vérifie la relation

$$m_{n+1} - \frac{a}{2} = \left(1 - \frac{2}{a}\right)\left(m_n - \frac{a}{2}\right).$$

On retrouve bien l'expression annoncée. □