Tirer un nombre au hasard

« Tirer quelque chose ou quelqu'un au hasard » est une expression redoutable. Tout le monde croit savoir ce que cela veut dire. Aussi, quand nous parlons de nombres tirés au hasard dans un ensemble fini, nous pensons, sans même y réfléchir, qu'il y a équiprobabilité. Il faut chasser cette idée car il y a une infinité d'autres cas *dont certains peuvent apparaître très simplement en classe* : un professeur ne peut pas les ignorer.

On croit aussi savoir ce que signifie « tirer un nombre au hasard dans un intervalle », ce qui présente le même inconvénient : il y a aussi une infinité de manières de procéder. On ne peut pas les confondre.

Cette note essaie de mettre en garde à ce sujet le lecteur non-probabiliste.

1 Tirer un nombre au hasard dans un ensemble fini

Voici des exemples d'équiprobabilité ou de non-équiprobabilité[1].

1.1 Exemple n˚1

Exemple 1 *Un entraineur de football veut tirer au hasard parmi les 11 joueurs de son équipe-type celui de ses joueurs qui l'accompagnera au tirage de la Coupe de France à Paris.* △

C'est facile : il met dans un chapeau 11 bouts de papier numérotés de 1 à 11, touille et en tire un, sans tricher. Clairement, tous les numéros étaient équiprobables. C'est le tirage au hasard au sens naïf.

Modèle 2 *Le modèle*[2] *associé à ce tirage est le suivant : l'ensemble Ω des résultats possibles est $\{1, 2, \ldots, 11\}$; ces événements élémentaires sont équiprobables.* □

L'affaire se corse quand cet entraineur se rappelle qu'il avait étudié autrefois le Calcul des probabilités et qu'on y utilisait beaucoup de dés. Il sait que quand on lance deux dés, la somme des chiffres qui sortent est un nombre entre 2 et 12. Si on lui retire 1, on aura donc *un nombre tiré au hasard entre 1 et 11.*

Il regrette de ne pas avoir procédé de cette manière plus savante de choisir son joueur. A-t-il raison ?

Les deux méthodes sont bien *des tirages au hasard d'un nombre entre 1 et 11.* Mais ils sont très différents car dans le premier cas, tous les joueurs avaient autant de chances d'être tirés que les autres tandis que dans le deuxième cas, les probabilités de sortie des différents chiffres sont :

1	2	3	4	5	6	7	8	9	10	11
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Tableau 3

1. Tirer un objet au hasard dans un ensemble fini, c'est toujours tirer un nombre au hasard. Il suffit en effet d'avoir recours à un *codage*, par exemple en numérotant les objets. Tirer un objet devient alors tirer un nombre.
2. On définit un modèle en indiquant les événements élémentaires et leurs probabilités.

Par exemple, dans le deuxième tirage, le joueur n° 6 a 6 fois plus de chances d'être tiré que le joueur n° 1 ! Pour fixer les idées, notons que :

Modèle 4 *Le modèle associé au deuxième tirage est le suivant : l'ensemble* Ω *des résultats possibles est* $\{1, 2, \ldots, 11\}$ *; les probabilités de ces événements élémentaires sont données par le tableau* (3). $\square$

Bien sûr, tout le monde trouvera que le modèle n°1 est le bon modèle et rejettera le n°2. Si vous avez un copain portant le numéro 5, 6 ou 7, vous préférerez peut-être le n°2. Dans des cas moins simples, on peut imaginer qu'il y ait *hésitation entre plusieurs modèles*. L'adoption d'un modèle est une affaire de *consensus*.

1.2 Exemple n°2

Dans cet exemple, nous montrons que plusieurs modèles peuvent représenter correctement une même expérience aléatoire et pourquoi il faudra choisir le modèle équiprobable quand il y en a un.

Exemple 5 *On tire* 3 *nombres au hasard dans l'ensemble* $(1, \ldots, 15)$. *Quelle est la probabilité* α *pour que ces nombres soient les longueurs d'un triangle non aplati*[3] *?* $\triangle$

Modèle 6 *Il y a à chaque tirage 15 résultats possibles et ces tirages sont indépendants. Il n'y a aucun doute qu'il y a* $15^3 = 3375$ *résultats possibles et que ces triplets, dont nous notons l'ensemble* Ω*, sont équiprobables.* $\square$

Pour trouver la probabilité demandée, il faut prendre ensuite ces résultats possibles un à un et compter le nombre n de ceux qui conduisent à un triangle non aplati. La probabilité α recherchée est $\alpha = \dfrac{n}{3375} \approx 0.502$, voir [1], [4].

On peut trouver que 3375 événements élémentaires, c'est beaucoup [5] et essayer de diminuer ce nombre, ce qui amènera à changer de modèle. Comme l'ordre dans lequel les 3 nombres sont tirés n'a pas d'importance pour le problème posé [6], on peut ne retenir de ces trois triplets que ceux dont les éléments sont rangés dans l'ordre croissant [7], [8]. Le programme « scilab » suivant donne à la fois le nombre N des triplets retenus, le nombre M de ceux qui conduisent à un triangle non aplati et le quotient $Q = \dfrac{M}{N}$:

3. On remarquera que cet énoncé est rédigé dans le style habituel, qui est notoirement insuffisant et générateur d'erreur : précisons que nous supposons que l'on tire un premier nombre, par exemple avec un chapeau comme précédemment, que l'on remet le bout de papier dans le chapeau et que l'on touille avant de tirer de nouveau un bout de papier, etc. Il faut toujours préciser complètement le protocole expérimental. Il s'agit ici d'un tirage avec remise.

4. On utilise la formule de Laplace : « dans un modèle équiprobable, la probabilité d'un événement est le quotient du nombre de cas favorables sur le nombre de cas possibles ».

5. Par exemple, si l'on travaille avec un tableur, il est fastidieux de saisir tous ces résultats possibles, ce qui occuperait 3 colonnes de 3375 lignes.

6. Par exemple les triplets $(1, 1, 3)$, $(1, 3, 1)$ et $(3, 1, 1)$ conduisent au même problème géométrique, à savoir « Existe-t-il un triangle non aplati dont les côtés ont respectivement 1, 1 et 3 pour longueur ? »

7. Ceci est une convention d'écriture qui garantit que chaque cas sera compté une fois et une seule.

8. Par exemple, des triplets $(1, 1, 3)$, $(1, 3, 1)$ et $(3, 1, 1)$, on ne retiendra que $(1, 1, 3)$.

Algorithme 7

Listing 1 – Triangles constructibles

```
clear
// Le nombre de triplets non ordonnés sera noté N
// Le nombre de triplets non ordonnés donnant un triangle non aplati
// est noté M
// Q est le quotient M/N
B=[];
for i=1:15
  for j=i:15
    for k=j:15
      A=[i,j,k];
      B=[B;A];
    end;
  end;
end;
N=size(B,"r")
M=0;
for i=1:N
  C=B(i,:);
  if ((abs(C(1)-C(2))<C(3))&(C(3)<C(1)+C(2))) then,
    M=M+1;
  end;
  Q=M/N;
end;
disp(['N=' +string(N)]);
disp(['M=' +string(M)]);
disp(['Q=' +string(Q)]);
```

On obtient $N = 680$, $M = 372$ et $Q = 0.5470588235294$.

Une erreur fréquente : Elle consiste à dire que $Q = \alpha$, ce qui est très gênant puisque l'on a trouvé précédemment que $\alpha \approx 0.502$. *L'erreur est de dire que M/N, quotient du « nombre de cas favorables » sur le « nombre de cas possibles » est la probabilité cherchée* parce que la formule de Laplace ne s'applique pas, puisque *nous ne sommes pas dans le cas équiprobable*[9], [10], [11].

En s'inspirant des notes ci-dessous, on pourra se convaincre que le modèle suivant est aussi bon que le modèle (6) ci-dessus :

Modèle 8 *L'ensemble des résultats possibles est l'ensemble des triplets (a, b, c) tels que*

$$1 \leqslant a \leqslant b \leqslant c \leqslant 15,$$

la probabilité d'un événement élémentaire étant $\dfrac{1}{3375}$ si les trois longueurs tirées sont égales (il y a 15 triplets de ce type), $\dfrac{3}{3375}$ si deux longueurs sont égales et différentes de la troisième (il

9. contrairement à la bonne vieille habitude

10. Par exemple, le triplet $(1, 1, 1)$ correspond à un seul élément de Ω, $(1, 1, 3)$ correspond à 3 éléments de Ω, enfin $(2, 3, 4)$ correspond à 6 éléments de Ω, à savoir $(2, 3, 4)$, $(2, 4, 3)$, $(3, 2, 4)$, $(3, 4, 2)$, $(4, 2, 3)$ et $(4, 3, 2)$. (1,1,1) est donc 3 fois moins probable que (1,1,3) et 6 fois moins probable que (2,3,4)

11. On pourrait illustrer ces affirmations à l'aide de simulations, en évoquant la loi des grands nombres.

y a 210 triplets de ce type), $\frac{6}{3375}$ *si les trois longueurs tirées sont différentes (il y a 455 triplets de ce type).* $\square$

Commentaires : Le modèle (8) n'est pas un modèle équiprobable mais c'est un bon modèle, à 680 événements élémentaires au lieu de 3375. On peut penser qu'il est plus simple, mais ce n'est pas vrai. En effet, on a perdu le confort des modèles équiprobables, parce que *la formule de Laplace ne s'applique plus.*

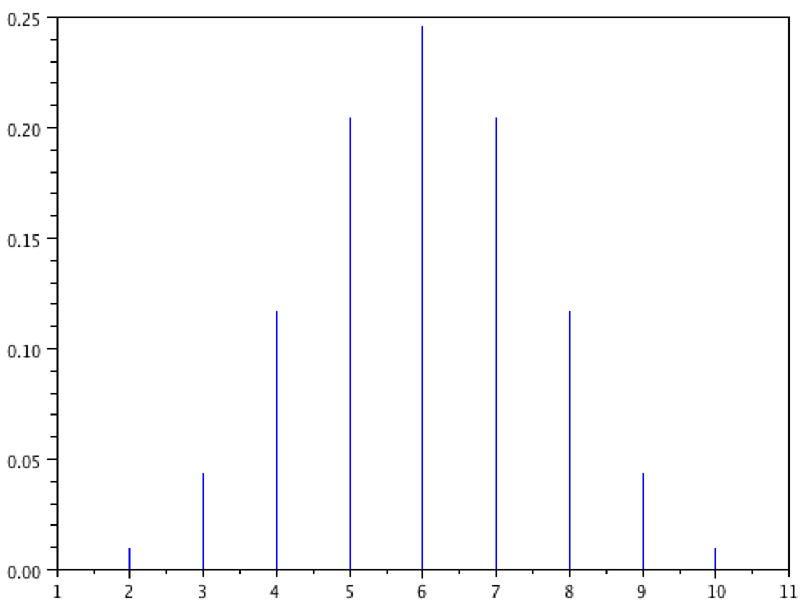
C'est en fait un inconvénient majeur : pour trouver α, on devra compter combien d'événements élémentaires de chaque catégorie conduisent à un triangle non aplati, multiplier par les probabilités correspondantes et additionner les produits. C'est beaucoup plus compliqué. $\triangledown$

1.3 Exemple n˚3

On pourrait multiplier les exemples : contentons-nous de lois binomiales, actuellement au programme de TS. On suppose que l'entraineur de l'équipe de football tire au sort un nombre entre 1 et 11 suivant le modèle [12] :

Modèle 9 *Pour k variant de 1 à 11, la probabilité que k sorte est* $\binom{10}{k-1}\cdot\frac{1}{2^{10}}\cdot$ $\square$

Ce tirage est facile à réaliser pratiquement. Il suffit, par exemple, qu'il demande à 10 supporters de lancer une pièce de monnaie quelconque en notant si c'est pile ou face qui sort et de décider que le joueur tiré au hasard est celui dont le numéro est le nombre de piles sortis plus 1. Le diagramme en bâtons qui suit représente les probabilités des différents résultats possibles :

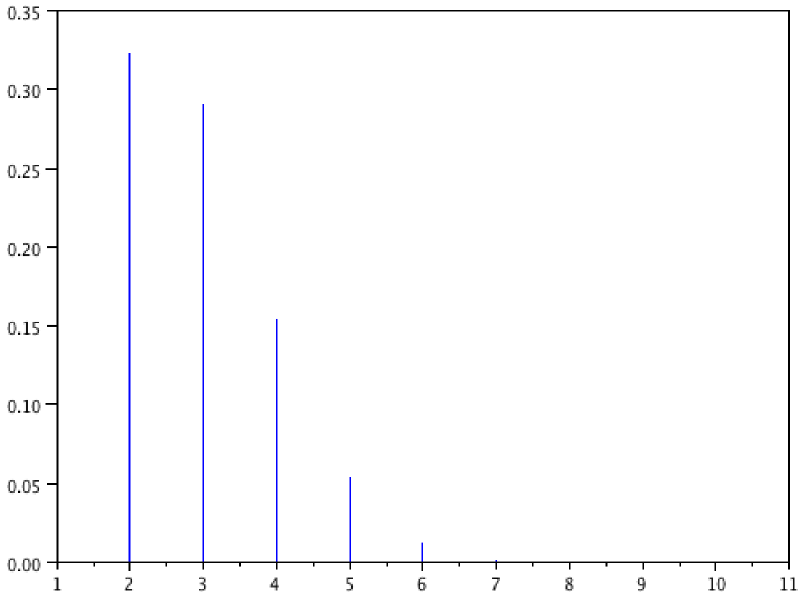


On voit que le joueur n˚5 a presque une chance sur 4 d'être tiré, les joueurs n˚1 et n˚11 n'en ont pratiquement aucune (en fait une chance sur 2^{10}).

12. On fait intervenir la loi binomiale $\mathcal{B}(10, \frac{1}{2})$.

1.4 Exemple n°4

Finalement, si l'entraineur avait demandé aux 10 supporters de lancer un dé en notant si 1 sort ou non et s'il avait décidé que le joueur tiré serait celui dont le numéro est le nombre de 1 sortis augmenté de 1 [13], on aurait obtenu le diagramme suivant [14] :



Dans ce cas, le joueur n°1 a presque une chance sur 3 d'être tiré au hasard [15]. On est bien loin du cas équiprobable !

2 Tirer un nombre au hasard dans l'intervalle $[0\,,1]$

Tirer un nombre au hasard dans l'intervalle $[0, 1]$ ne peut pas avoir de sens commun précis. On peut dire vaguement que tous les nombres de cet intervalle devraient avoir autant de chances d'être tirés que les autres, mais ceci n'a aucun sens puisqu'il y a une infinité de nombres dans cet intervalle [16]. Remarquons de plus que l'on ne peut pas imaginer une expérience qui produirait un tel nombre.

Habituellement, *un nombre tiré au hasard est interprété comme la valeur prise par une variable aléatoire.* On peut alors traduire correctement comme suit l'idée d'égalité des chances d'être tiré :

Définition 10 *On dit qu'une variable aléatoire X qui prend ses valeurs dans l'intervalle $[0\,,1]$ suit la loi uniforme sur cet intervalle si quel que soit l'intervalle $[\alpha\,,\beta]$ contenu dans $[0\,,1]$, la probabilité pour que X tombe dans $[\alpha\,,\beta]$ est la longueur de $[\alpha\,,\beta]$, soit*

$$\forall\,\alpha,\,\beta \quad \text{tels que} \quad 0 \leq \alpha \leq \beta \leq 1, \quad P(X \in [\alpha\,,\beta]) = \beta - \alpha. \quad \square \tag{1}$$

13. Ici intervient la loi binomiale $\mathcal{B}(10, \frac{1}{6})$.
14. qui est une manière agréable de remplacer la description du modèle
15. Il n'est pas très difficile de se convaincre que l'on peut favoriser qui l'on veut en organisant un tirage ad hoc. De plus, cela peut se faire très discrètement, en compliquant le modèle (penser aux jeux d'argent dans les casinos, les loteries, *etc*).
16. Si tous ces nombres avaient la même probabilité, même très petite, la probabilité totale serait infine !

On voit que cette probabilité ne dépend que de la longueur de $[\alpha\,,\beta]$. Par conséquent, si on fait glisser $[\alpha\,,\beta]$ à l'intérieur de $[0\,,1]$, elle ne change pas. C'est en ce sens que la loi de X est uniforme et traduit bien l'égalité des chances d'être tiré. Remarquons aussi que si on fait $\beta = \alpha$ dans (1), on voit que $P(X = \alpha) = 0$, ce qui montre que

> Si X est une variable aléatoire qui suit la loi uniforme sur l'intervalle $[0\,,1]$, la probabilité qu'elle prenne une valeur particulière est 0 [a].
>
> a. Ceci montre que la loi uniforme sur $[0\,,1]$ n'a vraiment rien à voir avec le tirage d'un nombre au hasard dans un *ensemble fini*.

Reprenons : si X désigne une variable aléatoire définie à partir d'une expérience aléatoire $\mathcal{E}$, qui prend ses valeurs dans l'intervalle $[0, 1]$, quand on réalise $\mathcal{E}$, X prend une valeur x qui est un nombre tiré au hasard dans $[0, 1]$.

> Cela ne veut pas du tout dire que X suit la loi uniforme sur $[0, 1]$. Si c'est le cas, il faut le préciser en disant : *soit x un nombre tiré au hasard dans l'intervalle $[0, 1]$ suivant la loi uniforme*. En effet, même au lycée dans des situations banales, « tirer un nombre au hasard dans $[0\,,1]$ » peut avoir une autre signification, comme le montrent les exemples suivants :

Exemple 11 *Soit X une variable aléatoire suivant la loi uniforme sur l'intervalle $[0\,,1]$. Alors la variable aléatoire X^2, qui prend ses valeurs dans $[0\,,1]$, ne suit pas la loi uniforme sur cet intervalle.* $\triangle$

En effet, pour tous réels α et β tels que $0 \leqslant \alpha \leqslant \beta \leqslant 1$,

$$P(X^2 \in [\alpha\,,\beta]) = P(X \in [\sqrt{\alpha}\,,\sqrt{\beta}]) = \sqrt{\beta} - \sqrt{\alpha}. \quad \square$$

Par exemple, $P(X^2 \leq \frac{1}{2}) = \frac{1}{\sqrt{2}}$ et non pas $\frac{1}{2}$, ce qui aurait été le cas si X^2 avait suivi la loi uniforme sur $[0\,,1]$. La valeur prise par X^2 est bien *un nombre tiré au hasard dans* $[0\,,1]$ puisque c'est le carré d'un nombre tiré au hasard dans $[0\,,1]$, mais ce n'est pas un nombre tiré au hasard suivant la loi uniforme sur $[0, 1]$. $\triangledown$

Exemple 12 *Soit X et Y deux variables aléatoires indépendantes*[17] *suivant la loi uniforme sur $[0\,,1]$. Alors la variable aléatoire $Z = \max(X, Y)$, qui prend ses valeurs dans $[0\,,1]$, ne suit ni la loi uniforme sur cet intervalle, ni la loi de X^2.* $\triangle$

En effet, $P(Z \leq \frac{1}{2}) = P((X \leq \frac{1}{2}) \cap (Y \leq \frac{1}{2})) = P(X \leq \frac{1}{2}) \times P(Y \leq \frac{1}{2}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ [18] et non pas $\frac{1}{2}$, ce qui aurait été le cas si Z avait suivi la loi uniforme sur $[0\,,1]$ ou $\frac{1}{\sqrt{2}}$, ce qui aurait été le cas si Z avait suivi la loi de X^2 [19]. $\triangledown$

On admettra qu'une variable aléatoire qui prend ses valeurs dans l'intervalle $[0, 1]$ peut suivre une infinité de lois différentes. La loi uniforme est un simple cas particulier.

17. La notion d'indépendance du Calcul des Probabilités ne fait pas partie du programme de cette note.
18. On a utilisé le fait que la probabilité de l'intersection de deux événements indépendants est le produit de leurs probabilités.
19. On démontre que Z suit la loi de densité $x \mapsto \frac{1}{2 \cdot \sqrt{x}} \cdot \mathbf{1}_{]0,1[}(x)$. La notion de loi à densité se trouve actuellement au programme de TS, limité aux lois uniformes et aux lois exponentielles. La densité de Z est plus problématique car elle n'est pas bornée sur $]0, 1[$.

Génération de nombres au hasard : Les nombres au hasard que l'on rencontre habituellement proviennent de nombres tirés au hasard suivant la loi uniforme sur $[0,1]$ et produits par un tableur, une calculatrice ou un logiciel de calcul [20]. Par exemple, les commandes « scilab » suivantes :

```
X=rand(1,10,[,'uniform"]);
disp(X);
```

produisent 10 nombres écrits en ligne et les affichent. On les considère comme les valeurs prises par 10 variables aléatoires *indépendantes* suivant la loi uniforme sur $[0,1]$. Plus généralement, on admettra que :

> Les nombres produits par une machine et/ou un logiciel de calcul sont les valeurs prises par des variables aléatoires indépendantes suivant la loi qui est spécifiée (par défaut, la loi uniforme sur $[0,1]$).

Cela ne peut pas se démontrer, mais est très vraisemblable d'après de nombreux arguments de statistique.

3 Conclusion

> « Tirer un nombre au hasard » est une expression extrêmement ambigüe, à bannir telle quelle. Il est absolument nécessaire de préciser la loi suivant laquelle le tirage est fait, que l'on tire au hasard dans un ensemble fini ou dans un intervalle. Quand la loi du tirage n'est pas spécifiée, on suppose par défaut qu'il s'agit de la loi uniforme. Cela contribue à aggraver les mauvaises pratiques mathématiques.

Références

[1] Raymond Moché : *Triangles constructibles et loi des grands nombres*
http://gradus-ad-mathematicam.fr/Remarquables1.htm

[2] Odile Guillon : *Triangles constructibles et fréquences*, activité proposée sur le site de l'IREM de Lille pour la classe de Cinquième, en géométrie.
http://irem.univ-lille1.fr/activites/spip.php?article162

20. L'examen des multiples problèmes liés à la génération de nombres aléatoires ne fait pas partie du programme de cette note.