Probabilités et échantillonnage

I. *Loi de probabilité et modélisation*

Définition - Expérience aléatoire et univers

Une expérience aléatoire est une expérience dont les issues sont connues sans que l'on puisse déterminer laquelle sera réalisée.
L'univers d'une expérience aléatoire est l'ensemble des issues possibles de cette expérience.

Notation : L'univers se note souvent Ω et se lit « oméga ».

Exemple
On lance une pièce de monnaie et on regarde de quel côté elle tombe.
Les résultats sont Pile et Face. Pour cette expérience aléatoire, l'univers est Ω ={Pile, Face}.

Définition - Loi de probabilité

Donner une loi de probabilité associée à une expérience aléatoire, c'est en donner toutes les issues et attribuer une probabilité (un nombre compris entre O et 1) à chacune d'elles de sorte que la somme des probabilités des issues est égale à 1.
On peut présenter les résultats sous la forme d'un tableau.

Exemples
① Une étude menée sur la répartition des groupes sanguins en France a montré que 45 % de la population est du groupe A, 9 % du groupe B, 4 % du groupe AB et 42 % du groupe O.
On choisit au hasard une personne en France et on note son groupe sanguin.
Cette expérience aléatoire peut être modélisée à l'aide du tableau des fréquences suivant.

Issue	A	B	AB	O
Probabilité	0,45	0,09	0,04	0,42

② On lance un dé cubique dont les faces sont numérotées de 1 à 6. *A priori*, on peut estimer que chaque face a la même probabilité d'apparaître, égale à $\frac{1}{6}$. Cela ne signifie pas que l'on représente exactement la situation réelle : tout dé présente des imperfections physiques, aucun n'est parfaitement équilibré.

➢ **Exercice résolu 1 page 321**

Définition - Loi équirépartie

Une loi est dite équirépartie lorsque chaque issue a la même probabilité de se réaliser, qui est alors :

$$\frac{1}{nombre\ total\ d'issues} = \frac{1}{n}$$

Exemple
Lors d'un lancer d'un dé cubique équilibré à 6 faces, aucune des faces n'est favorisée, donc chacune des 6 faces a autant de chances qu'une autre d'être obtenue.
Chaque face a donc une probabilité d'apparition de $\frac{1}{6}$.

➢ **Exercice résolu 3 page 322**

Remarque : On dit aussi qu'on est dans une situation d'**équiprobabilité**.

II. Évènement

Définition

Un événement est un sous-ensemble de l'univers. Il peut s'écrire à l'aide d'issues ou être décrit à l'aide d'une phrase.

Exemple

On lance un dé cubique équilibré et on observe le résultat.
Alors l'univers associé est Ω = {1 ;6}.
L'événement A : « Obtenir un nombre pair » est A = {2 ; 4 ; 6}.

Remarque : Si une issue appartient à un événement, on dit qu'elle réalise cet événement.

Définition - Probabilité d'un événement

La probabilité d'un événement A est égale à la somme des probabilités des issues qui réalisent cet événement.

Vocabulaire

Un événement impossible est un événement qui ne se réalise jamais : sa probabilité vaut 0.
Un événement certain est un événement qui est sûr de se réaliser : sa probabilité vaut 1.

Exemple

Dans le cas de la répartition des groupes sanguins (voir l'exemple page précédente), la probabilité qu'une personne en France ait un groupe sanguin différent de A est égale à : $0{,}09 + 0{,}04 + 0{,}42 = 0{,}55$.

➢ **Exercice résolu 2 page 321**

Propriété - Cas d'équiprobabilité

Dans une situation d'équiprobabilité, où il y a n issues, la probabilité d'un événement A réalisé par k issues est :

$$p(A) = \frac{nombre\ d'issues\ qui\ réalisent\ A}{nombre\ total\ d'issues} = \frac{k}{n}$$

Exemple

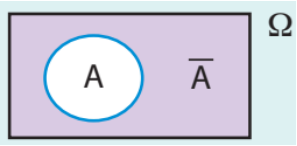
Dans le cas du lancer d'un dé cubique équilibré à 6 faces, la loi peut se modéliser par une loi équirépartie sur l'ensemble Ω ={1 ; 6}.
Le nombre total d'issues est $n = 6$.
L'événement A : « Obtenir un nombre pair » s'écrit A = {2 ; 4 ; 6}.
Le nombre d'issues qui réalisent l'événement A est $k = 3$
La probabilité de l'événement A est donc :

$$p(A) = \frac{k}{n} = \frac{3}{6} = 0{,}5$$

III. Opérations sur les événements

Définition - Événement et événement contraire

On considère une expérience aléatoire d'univers Ω.
Soit A un événement. L'événement contraire de A, noté $\bar{A}$, est l'ensemble des issues qui ne réalisent pas A, autrement dit $\bar{A}$ est réalisé par les issues de Ω qui ne sont pas dans A.



Propriété - Probabilité de l'événement contraire

Soit A un événement, on a $p(\bar{A}) = 1 - p(A)$.

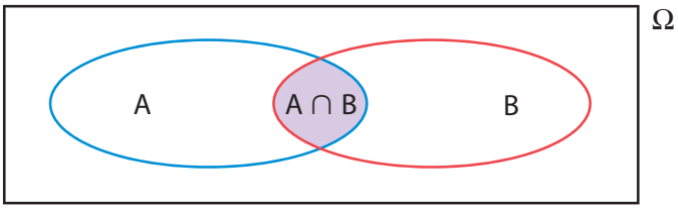
Exemple

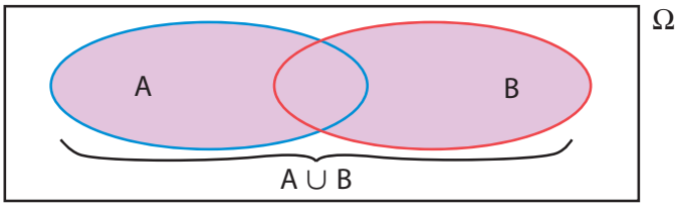
Dans le cas du lancer de dé à 6 faces, le contraire de l'événement A : « Obtenir un nombre pair » est $\bar{A}$: « Obtenir un nombre impair ».
On a donc $p(\bar{A}) = 1 - p(A) = 1 - 0{,}5 = 0{,}5$.

Définition - Union et intersection de deux événements

Soit A et B deux événements.
L'événement $A \cup B$ (se lit « A union B ») est la réunion de A et de B : c'est l'ensemble des issues qui réalisent A ou B (ou les deux à la fois).
L'événement $A \cap B$ (se lit « A inter B »)est l'intersection de A et de B : c'est l'ensemble des issues qui réalisent A et B.

Remarque : Le diagramme de Venn permet de représenter les différents événements.





Exemple

On lance un dé à 6 faces et on considère les événements A : « Obtenir un nombre pair » et B : « Obtenir un multiple de 3 » $A = \{2\,;4\,;6\}$ et $B = \{3\,;6\}$. Alors $A \cap B = \{6\}$ et $A \cup B = \{2\,;3\,;4\,;6\}$.

Propriété - Relation entre union et intersection

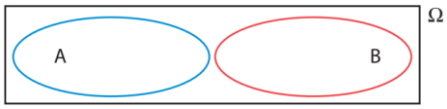
On a $p(A) + p(B) = p(A \cup B) + p(A \cap B)$. En particulier, $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

Exemple

Dans l'exemple précédent, on a :
$p(A) = \frac{3}{6} = \frac{1}{2}$ $\quad p(B) = \frac{2}{6} = \frac{1}{3}$ $\quad p(A \cap B) = \frac{1}{6}$
On a alors $p(AUB) = p(A) + p(B) - p(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}$

➢ **Exercice résolu 4 page 323**

Remarque : On dit que A et B sont des **événements disjoints** si $A \cap B = \emptyset$.
On a alors, **dans ce cas seulement**, $p(A \cap B) = 0$ et $p(A \cup B) = p(A) + p(B)$.



IV. *Échantillon et simulation*

Définition - Échantillon

Lorsque l'on réalise plusieurs fois une même expérience aléatoire de manière indépendante (c'est-à-dire que les différentes réalisations n'ont pas d'influence les unes sur les autres), l'ensemble des résultats obtenus est appelé échantillon.
Le nombre de fois où l'expérience est réalisée est appelée taille de l'échantillon.

Exemples

① Si l'on lance 10 fois un dé équilibré à 6 faces et qu'on observe le résultat obtenu, on obtient un échantillon de taille 10.

② Si l'on tire au sort 1 000 personnes dans la population française et que l'on observe si la personne est droitière ou non, on obtient un échantillon de taille 1 000.

Propriété - Simulation

On peut simuler informatiquement une expérience aléatoire à deux issues x_1 et x_2 de probabilités respectives p et $1-p$ en générant un nombre réel aléatoire entre 0 et 1 et en considérant que :

- x_1 est réalisée si ce nombre aléatoire est inférieur ou égal à p ;
- x_2 est réalisée sinon.

Démonstration

La probabilité que la simulation donne x_1 (resp. x_2) est bien p (resp. $1-p$), c'est-à-dire la même probabilité que l'expérience aléatoire donne x_1 (resp. x_2).

Remarque

Pour mettre en œuvre cette propriété, il faut disposer de commandes permettant de générer un nombre réel aléatoire entre 0 et 1 dans différents langages :

- avec la calculatrice : **NbrAléat**, (respectivement pour la calculatrice TI) génère un nombre réel aléatoire entre 0 et 1.
- avec le tableur : **ALEA()** génère un nombre réel aléatoire entre 0 et 1.
- avec PYTHON : On peut également générer des nombres aléatoires de différents types.
  Pour cela, il faut écrire import random en début de programme puis :
  - random. random ( ) renvoie un nombre réel aléatoire dans [0 ; 1[;
  - random. randint (a, b) renvoie un nombre entier aléatoire entre a et b inclus.

Exemples

① Dans la population, il y a 88 % de droitiers, ce qui signifie que la probabilité qu'une personne soit droitière est 0,88. On souhaite simuler l'expérience aléatoire consistant à tirer au sort une personne dans la population et à regarder si elle est droitière ou non à l'aide d'un tableur.
On saisit = **ALEA()** dans une cellule et on considère que :

- la personne est droitière si le nombre obtenu est inférieur ou égal à 0,88 ;
- la personne n'est pas droitière sinon.

② Avec PYTHON, on peut simuler un échantillon de 1 000 personnes selon qu'elles sont droitières ou non (voir ci-contre).

```
import random
for i in range (1,1001):
  if random.random() <= 0.88:
      print("Droitière")
  else:
      print("Non droitière")
```

➢ **Exercice résolu 5 page 323**

V. Fluctuation et estimation

Définition - Fluctuation d'échantillonnage

Deux échantillons (obtenus par l'expérience ou simulés) de même taille associés à une même expérience aléatoire ne sont, a priori, pas identiques : ce phénomène s'appelle la fluctuation d'échantillonnage.

Exemple

Si l'on lance dix fois un dé à 6 faces puis que l'on recommence, les résultats des dix premiers lancers ne seront pas identiques aux dix suivants.

Propriété - Échantillon

On considère un échantillon de taille n associé à une expérience aléatoire dont l'une des issues (ou l'un des événements) a pour probabilité p.
Pour n grand, sauf exception, la fréquence observée f de cette issue (ou événement) dans l'échantillon est proche de sa probabilité p (la plupart du temps, l'écart entre p et f est inférieur ou égal à $\frac{1}{\sqrt{n}}$).

Exemple

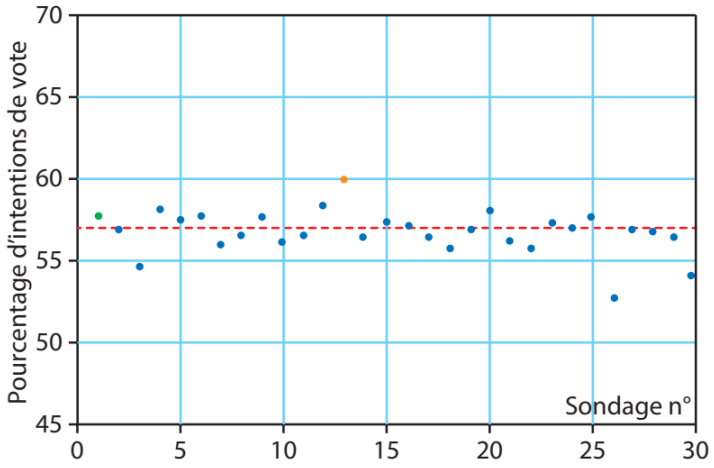
On a lancé 1 000 fois un dé à 6 faces et on a obtenu 159 fois le nombre 6. La fréquence observée de 6 est donc $f = \frac{159}{1000} = 0{,}159$, ce qui est assez proche de la probabilité d'obtenir 6 qui est $p = \frac{1}{6} \approx 0{,}167$.
L'écart entre p et f est $\frac{1}{6} - 0{,}159 \approx 0{,}008$, qui est bien inférieur à $\frac{1}{\sqrt{1000}} \approx 0{,}032$.

Propriété - Estimation

On considère un échantillon de taille n associé à une expérience aléatoire dont l'un des événements a pour probabilité p et où f est la fréquence observée de cet événement dans l'échantillon.
Pour n grand, f et p sont proches donc, si l'on ne connaît pas la valeur de p, on peut considérer que f en constitue une estimation.
On utilise généralement plusieurs échantillons de même taille pour réaliser une bonne estimation.

Exemple

On donne ci-dessous les pourcentages d'intentions de vote pour une candidate selon 30 sondages portant sur 1 000 personnes, réalisés le même jour, c'est-à-dire 30 échantillons de taille 1 000 associés à l'expérience aléatoire consistant à tirer au sort une personne dans la population et à lui demander si elle souhaite voter pour cette candidate ou non.
On constate que, dans le 1er sondage (point vert), la candidate a environ 58 % d'intention de vote, soit une fréquence observée de 0,58 : on peut penser que la probabilité p qu'une personne vote pour elle est proche de 0,58.
Dans le 13ème sondage (point orange), la fréquence observée est environ 60 % donc on peut penser que p est proche de 0,60, etc. En observant globalement ce graphique, on peut estimer cette probabilité p à environ 0,57 qui semble correspondre « au milieu » (voir droite en pointillés rouges) du nuage de points (mais cela reste une estimation !).



- **Exercice résolu 6 page 325**