# 13 : Variables aléatoires et lois des grands nombres

## I. Somme de variables aléatoires

### 1. Exemple

On considère deux jeux dont les gains sont donnés :
* pour le premier jeu, par la variable aléatoire $X$ qui prend les valeurs 1 et 2.
* pour le second jeu, par la variable aléatoire $Y$ qui prend les valeurs –2, 3 et 4.

Par exemple, l'évènement $(X = 1) \cap (Y = -2)$ signifie qu'on a gagné 1 € au premier jeu et perdu 2 € au deuxième jeu.

Considérons la variable aléatoire somme $X + Y$ donnant le gain total cumulé aux deux jeux.
Alors la variable aléatoire $X + Y$ peut prendre les valeurs : $-1$, 0, 4, 5 et 6.
En effet, on a par exemple $X + Y = 0$ avec $(X = 2) \cap (Y = -2)$.

Par ailleurs, pour calculer par exemple, la probabilité de l'évènement $X + Y = 5$, on cherche toutes les sommes $X + Y$ égales à 5.
On a ainsi : $P(X + Y = 5) = P\big((X = 1) \cap (Y = 4)\big) + P\big((X = 2) \cap (Y = 3)\big)$

Si de plus, les évènements $X$ et $Y$ sont indépendants, alors on a :
$P(X + Y = 5) = P(X = 1) \times (Y = 4) + P(X = 2) \times (Y = 3)$

### 2. Définition

Soit $X$ et $Y$ deux variables aléatoires.
La loi de probabilité de la variable aléatoire somme $X + Y$ est donnée par :

$$\mathrm{P}(X + Y = k) = \sum_{i+j=k} P(X = i) \cap P(Y = j)$$

Si de plus les évènements $(X = i)$ et $(Y = j)$ sont indépendants, alors on a :

$$P(X + Y = k) = \sum_{i+j=k} P(X = i) \times P(Y = j)$$

On dit dans ce cas que les variables aléatoires $X$ et $Y$ sont indépendantes.

### Exemple

Si par exemple, $k = 2$ alors :
$P(X + Y = 2) = P(X = 2) \cap P(Y = 0) + P(X = 1) \cap P(Y = 1) + P(X = 0) \cap P(Y = 2)$

## Exercice 1 : Déterminer la loi d'une somme de variables aléatoires

On considère le jeu suivant qui se déroule en deux parties :
- La 1ère partie consiste à lancer une pièce de monnaie. Si on tombe sur « pile », on gagne 1 €, si on tombe sur « face », on gagne 2 €.

- La 2e partie consiste à lancer un dé à 6 faces. Si on tombe sur un chiffre pair, on gagne 1 €, si on tombe sur le « 3 » ou le « 5 », on gagne 2 €.
Si on tombe sur le « 1 », on perd 5 €.
La variable aléatoire $X$ désigne les gains à la 1ère partie, la variable aléatoire $Y$ désigne les gains à la 2e partie.
On considère que les variables aléatoires $X$ et $Y$ sont indépendantes.
Établir la loi de probabilité de la variable aléatoire somme $S = X + Y$ donnant le gain total cumulé à la fin des deux parties.

## Correction

Dans le tableau ci-dessous, on présente toutes les sommes possibles :

| Y \ X | 1 | 2 |
|---|---|---|
| -5 | -4 | -3 |
| 1 | 2 | 3 |
| 2 | 3 | 4 |

Ainsi, on a :

$$P(S=-4) = P((X=1)\cap(Y=-5))$$
$$= P(X=1)\times P(Y=-5) \text{ en effet, les variables } X \text{ et } Y \text{ sont indépendantes.}$$
$$= \frac{1}{2}\times\frac{1}{6} = \frac{1}{12}$$

$$P(S=-3) = P(X=2)\times P(Y=-5)$$
$$= \frac{1}{2}\times\frac{1}{6} = \frac{1}{12}$$

$$P(S=2) = P(X=1)\times P(Y=1)$$
$$= \frac{1}{2}\times\frac{1}{2} = \frac{1}{4}$$

$$P(S=3) = P(X=1)\times P(Y=2) + P(X=2)\times P(Y=1)$$
$$= \frac{1}{2}\times\frac{1}{3} + \frac{1}{2}\times\frac{1}{2} = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}$$

$$P(S=4) = P(X=2)\times P(Y=2)$$
$$= \frac{1}{2}\times\frac{1}{3} = \frac{1}{6}$$

On peut présenter la loi de probabilité de $S$ dans un tableau :

| $k$ | −4 | −3 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(S=k)$ | $\frac{1}{12}$ | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{5}{12}$ | $\frac{1}{6}$ |

## II. Espérance et variance de combinaisons linéaires de variables aléatoires

### 1. Propriétés

- $E(aX+b) = aE(X) + b$
- $E(X+Y) = E(X) + E(Y)$
- $V(aX+b) = a^2V(X)$
- Si $X$ et $Y$ sont deux variables aléatoires indépendantes : $V(X+Y) = V(X) + V(Y)$

### 2. Application

Soit $X$ une variable aléatoire telle que $E(X) = 10$ et $V(X) = 4$.
Soit $Y$ une autre variable aléatoire telle que $\mathrm{E}(Y) = 0$ et $\mathrm{V}(Y) = 1$.
Calculer l'espérance et la variance des variables aléatoires suivantes.
$Z = 2X - 15$ $\quad T = X + Y$ $\quad U = -5Y$

**Correction**

Par linéarité de l'espérance, on a $E(Z) = 2E(X) - 15 = 2 \times 10 - 15 = 5$.
$V(Z) = 2^2 V(X) = 16$.

Par linéarité de l'espérance, on a $E(T) = E(X) + E(Y) = 10 + 0 = 10$.
En l'absence d'information sur l'indépendance des expériences aléatoires associées aux variables $X$ et $Y$, on ne peut appliquer la formule sur la variance de la somme. On ne peut donc pas déterminer $V(T)$.

Par linéarité de l'espérance, on a $E(U) = -5E(Y) = 0$ et $V(U) = (-5)^2 V(Y) = 25 \times 1 = 25$.

**Exercice 2 : Simplifier les calculs d'espérance et de variance**

Une entreprise qui fabrique des roulements à bille fait une étude sur une gamme de billes produites. Le diamètre théorique doit être égal à 1,3 cm mais cette mesure peut être légèrement erronée.
L'expérience consiste à tirer au hasard une bille d'un lot de la production et à mesurer son diamètre.
On considère la variable aléatoire $X$ qui, à une bille choisie au hasard, associe son diamètre.
La loi de probabilité de $X$ est résumée dans le tableau suivant :

| $x_i$ | 1,298 | 1,299 | 1,3 | 1,301 | 1,302 |
|---|---|---|---|---|---|
| $P(X = x_i)$ | 0,2 | 0,1 | 0,2 | 0,4 | 0,1 |

Calculer l'espérance et l'écart-type de la loi de probabilité de $X$.

Aide : On pourra poser $Y = 1000X - 1300$

**Correction**

Pour simplifier les calculs, on définit la variable aléatoire $Y = 1000X - 1300$.
La loi de probabilité de $Y$ est alors :

| $y_i$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $P(Y = y_i)$ | 0,2 | 0,1 | 0,2 | 0,4 | 0,1 |

Calculons l'espérance et la variance de la loi de probabilité de $Y$ :

$E(Y) = -2 \times 0{,}2 + (-1) \times 0{,}1 + 1 \times 0{,}4 + 2 \times 0{,}1 = 0{,}1$

$V(Y) = 0{,}2 \times (-2 - 0{,}1)^2 + 0{,}1 \times (-1 - 0{,}1)^2 + 0{,}2 \times (0 - 0{,}1)^2 + 0{,}4 \times (1 - 0{,}1)^2 + 0{,}1 \times (2 - 0{,}1)^2$

$= 1{,}69$

On en déduit l'espérance et la variance de la loi de probabilité de $X$ :

$E(Y) = E(1000X - 1300) = 1000E(X) - 1300$

Donc : $E(X) = \dfrac{E(Y)+1300}{1000} = \dfrac{0{,}1+1300}{1000} = 1{,}3001$

$V(Y) = V(1000X - 1300) = 10002\,V(X)$

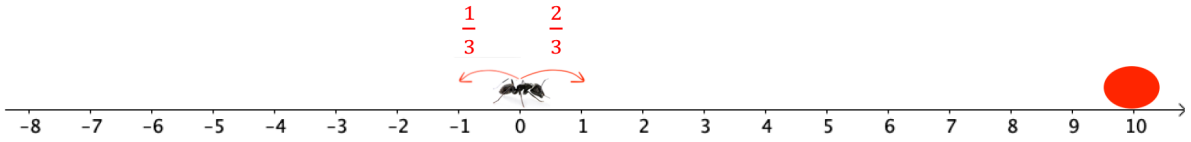
Donc : $V(X) = \dfrac{V(Y)}{1000^2} = \dfrac{1{,}69}{1000^2}$

Et donc : $\sigma(X) = \sqrt{\dfrac{1{,}69}{1000^2}} = \dfrac{1{,}3}{1000} = 0{,}0013$

Conclusion : $E(X) = 1{,}3001\ cm$ et $\sigma(X) = 0{,}0013\ cm$.

**Exercice 2 : Calculer une espérance et une variance à l'aide d'une somme de variables aléatoires**

Sur un axe gradué, on dépose une petite goutte de confiture à la fraise au point d'abscisse 10. Moshé pose une fourmi sur l'origine de l'axe gradué. Attirée par la confiture, la fourmi se déplace de façon aléatoire d'une unité vers la droite (sens positif) avec la probabilité de $\frac{2}{3}$ et d'une unité vers la gauche (sens négatif) avec la probabilité de $\frac{1}{3}$.
On suppose que les déplacements de la fourmi sont indépendants les uns des autres.
Pour tout entier naturel $n$, on note $X_n$ la variable aléatoire valant 1 si la fourmi se déplace vers la droite au $n$-ième déplacement et valant $-1$ si elle se déplace vers la gauche.
On note $S_n = X_1 + X_2 + X_3 + \cdots + X_n$ la variable aléatoire somme des $X_i$.




**1.** Calculer $\mathrm{E}(X_i)$ et $\mathrm{V}(X_i)$
**2.** En déduire $E(S_n)$ et $V(S_n)$.
**3.** Au bout de combien de déplacements, la fourmi peut-elle espérer théoriquement atteindre la goute de confiture ? Calculer $\sigma(S_n)$ dans ce cas.

**Correction**

1) On établit la loi de probabilité de $X_k$ :

| $x_i$ | $-1$ | $1$ |
|---|---|---|
| $P(X_k = x_i)$ | $\frac{1}{3}$ | $\frac{2}{3}$ |

$$E(X_k) = -1 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{1}{3}$$
$$V(X_k) = \frac{1}{3}\left(-1 - \frac{1}{3}\right)^2 + \frac{2}{3}\left(1 - \frac{1}{3}\right)^2 = \frac{8}{9}$$

2) On a : $S_n = X_1 + X_2 + \cdots + X_n$
Donc la variable aléatoire $S_n$ donne l'abscisse de la fourmi après $n$ déplacements.
Et on a : $E(S_n) = E(X_1 + X_2 + \cdots + X_n)$
$= E(X_1) + E(X_2) + \cdots + E(X_n)$
$= nE(X_k)$
$= \frac{n}{3}$

Et : $V(S_n) = V(X_1 + X_2 + \cdots + X_n)$
$= V(X_1) + V(X_2) + \cdots + V(X_n)$, car les variables sont indépendantes.
$= nV(X_k)$
$= \frac{8n}{9}$

3) $E(S_n) = 10$
$\frac{n}{3} = 10$
$n = 30$
Après 30 déplacements, Sophie peut espérer théoriquement atteindre la goûte de confiture.

$\sigma(S_{30}) = \sqrt{V(S_{30})} = \sqrt{V(S_{30})} = \sqrt{\frac{8\times30}{9}} \approx 5{,}2$ unités.

## 3. Espérance, variance et écart type pour une loi binomiale

Soit $X$ une variable aléatoire qui suit la loi binomiale de paramètres $n$ et $p$.

$E(X) = np$ $\quad V(X) = np(1-p)$ $\quad \sigma(X) = \sqrt{V(X)}$

**Exercice 3 : Calculer l'espérance, la variance et l'écart-type pour une loi binomiale**

On lance 5 fois de suite un dé à six faces. On considère comme succès le fait d'obtenir 5 ou 6.
On considère la variable aléatoire $X$ donnant le nombre de succès.
Calculer E(X), $V(X)$ et $\sigma(X)$.

**Correction**

La variable aléatoire $S$ suit la loi binomiale de paramètres $p = \frac{1}{3}$ et $n = 5$.

$$E(S) = 5 \times \frac{1}{3} = \frac{5}{3} \approx 1{,}7$$

$$V(S) = 5 \times \frac{1}{3} \times \frac{2}{3} = \frac{10}{9}$$

$$\sigma(S) = \sqrt{\frac{10}{9}} = \frac{\sqrt{10}}{3}$$

En moyenne, on peut espérer obtenir environ 1,7 fois un 5 ou un 6, en 5 lancers.

## III. Inégalité de concentration

### 1. Inégalité de Markov

**Propriété**

Soit $X$ une variable aléatoire à valeurs positives et soit $a$ un nombre réel strictement positif.

On a : $P(X \geq a) \leq \frac{E(X)}{a}$

**Exemple**

Soit $X$ une variable aléatoire d'espérance 1.

$$P(X \geq 100) \leq 0{,}01$$

### 2. Inégalité de Bienaymé Tchébychev

**Propriété**

Soit $X$ une variable aléatoire et soit $t$ un nombre réel strictement positif.

On a : $P(|X - E(X)| \geq t) \leq \frac{V(X)}{t^2}$

**Application**

Le nombre de pièces fabriquées dans une usine en une journée suit une variable aléatoire d'espérance 50 et de variance 25.
Donner deux majorations de la probabilité que la production dépasse sur une journée 75 pièces.

**Correction**

On note $X$ la variable aléatoire égale au nombre de pièces fabriquées un jour donné.
On cherche à majorer $P(X \geqslant 75)$.
D'après l'inégalité de Markov, avec $a = 75$, on a $P(X \geqslant 75) \leqslant \frac{E(X)}{75}$.
Or $\frac{E(X)}{75} = \frac{50}{75} = \frac{2}{3}$ , ce qui donne une première majoration : $P(X \geqslant 75) \leqslant \frac{2}{3}$.
Une seconde majoration peut être obtenue à l'aide de l'inégalité de Bienaymé-Tchebychev. Pour cela, on fait apparaître la bonne « forme » : on dit que l'on centre la variable $X$ en lui retranchant son espérance.
$P(X \geqslant 75) = P(X - 50 \geqslant 25) \leqslant P(|X - 50| \geqslant 25)$
Or $P(|X - 50| \geqslant 25) \leqslant \frac{V(X)}{25^2}$ d'après l'inégalité de Bienaymé-Tchebychev.
Donc $P(X \geqslant 75) \leqslant P(|X - E(X)| \geqslant 25) \leqslant \frac{V(X)}{25^2} = \frac{25}{25^2} = \frac{1}{25}$.
On a donc $P(X \geqslant 75) \leqslant \frac{1}{25}$ ce qui est une seconde majoration.
**Remarque :** Il apparaît clairement ici que l'inégalité de Bienaymé-Tchebychev fournit une majoration bien plus précise que l'inégalité de Markov. On privilégiera en général l'inégalité de Bienaymé-Tchebychev.

## IV. Loi des grands nombres

### 1. Définition

On considère $n$ expériences aléatoires identiques et indépendantes.
On note $X_1, X_2, \ldots, X_n$ les variables aléatoires associées à ces expériences, toutes de même loi.
On note $S_n = X_1 + X_2 + \ldots + X_n$ et $M_n = \frac{S_n}{n}$
$M_n$ s'appelle la moyenne empirique des variables $X_1, X_2, \ldots, X_n$.

### 2. Théorème : Loi des grands nombres

On considère une expérience aléatoire et $X$ la variable aléatoire associée à cette expérience, d'espérance $E(X)$ et de variance $V(X)$.
On répète $n$ fois cette expérience de manière indépendante. On obtient un échantillon de taille $n$ composé de $n$ variables aléatoires $X_1, X_2, \ldots, X_n$.
Les variables $X_1, X_2, \ldots, X_n$ ont la même loi (elles ont donc la même espérance $E(X)$ et la même variance $E(X)$).
Pour tout nombre réel $t > 0$, $P(|M_n - E(X)| \geq t) \leq \frac{V(X)}{nt^2}$

Autrement dit, on a $\lim\limits_{n \to +\infty} P(|M_n - E(X)| \geq t) = 0$

On dit que $M_n$ converge en probabilité vers $E(X)$ lorsque $n$ tend vers $+\infty$.

**Application**
Soit une variable aléatoire $X$ d'espérance 3 et de variance 0,2.
Soit $M_n$ la variable aléatoire moyenne d'un échantillon de taille $n$ de la variable aléatoire $X$.
**1.** Déterminer un majorant de $P(|M_n - 3| \geq 0{,}1)$ pour $n = 100$, pour $n = 1000$, puis pour $n = 10\,000$. Que constate-t-on ?
**2.** Démontrer et interpréter le résultat précédent.

**Correction**

**1.** On sait que : $P(|M_n - E(X)| \geq t) \leq \frac{V(X)}{nt^2}$

Ici on a : $P(|M_n - 3| \geq 0{,}1) \leq \frac{0{,}2}{n0{,}1^2}$

Plus simplement on a : $P(|M_n - 3| \geq 0{,}1) \leq \frac{20}{n}$

Pour $n = 100$, $P(|M_n - 3| \geq 0{,}1) \leq \frac{20}{100} = 0{,}2$

Pour $n = 1000$, $P(|M_n - 3| \geq 0{,}1) \leq \frac{20}{100} = 0{,}02$

Pour $n = 10\,000$, $P(|M_n - 3| \geq 0{,}1) \leq \frac{20}{100} = 0{,}002$

On constate que plus $n$ est grand plus la probabilité $P(|M_n - 3| \geq 0{,}1)$ se rapproche de 0.

**2.** D'après la loi des grands nombres, on a : $\lim\limits_{n\to+\infty} P(|M_n - E(X)| \geq t) = 0$

Soit, dans le contexte de l'exercice : $\lim\limits_{n\to+\infty} P(|M_n - 3| \geq t) = 0$

Plus la taille de l'échantillon est grande, plus l'écart entre la moyenne $M_n$ et l'espérance de la variable aléatoire $X$ est faible.