Sommes de variables aléatoires

Introduction :

Dans la première partie de ce cours, nous allons voir quelques définitions et propriétés importantes sur la transformation de variables aléatoires, après avoir fait quelques rappels de première.
Rappelons-nous que nous avons étudié, dans le cours précédent, la loi de probabilité associé à un schéma de Bernoulli, à savoir la loi binomiale. Nous allons voir comment calculer l'espérance, la variance et l'écart-type de la loi binomiale, grâce à des outils très puissants donnés par la linéarité de l'espérance et l'additivité de la variance (dans le cas des variables aléatoires indépendantes). Ceci sera l'objet de la deuxième partie.
Enfin, toujours grâce à ces mêmes outils, nous établirons plus généralement l'expression de l'espérance, de la variance et de l'écart-type de la somme et de la moyenne d'un échantillon de taille n d'une loi de probabilité.

→ Dans tout ce cours, Ω désigne l'univers associé à une expérience aléatoire et p une probabilité sur Ω.

1 Transformation de variables aléatoires

Nous allons tout d'abord faire quelques rappels de ce que nous avons appris sur les variables aléatoires.

a. Rappels

Commençons par redonner la définition d'une variable aléatoire.

Définition

Variable aléatoire :

Une variable aléatoire réelle définie sur l'univers Ω est une fonction, notée X, qui associe un réel à chaque éventualité de l'univers Ω. On a donc :

$$\begin{aligned} X\ :\ & \Omega \to \mathbb{R} \\ & \omega \mapsto X(\omega) \end{aligned}$$

À une variable aléatoire, nous pouvons associer sa loi de probabilité.

Définition

Loi de probabilité d'une variable aléatoire :

Soit X une variable aléatoire discrète sur Ω qui prend les valeurs x_1, x_2, ..., x_n.
Définir la loi de probabilité de X, c'est donner les valeurs de probabilités $p(X = x_i)$ pour tout entier i, avec $1 \leq i \leq n$.

Nous pouvons maintenant redonner la définition de l'espérance d'une variable aléatoire.

Définition

Espérance d'une variable aléatoire :

Soit X une variable aléatoire discrète sur Ω qui prend les valeurs x_1, x_2, ..., x_n. On note p_i la valeur de la probabilité $p(X = x_i)$ pour tout $1 \leq i \leq n$.
On appelle espérance de X le nombre réel, noté $E(X)$, défini par :

$$\begin{aligned} E(X) &= x_1 \times p_1 + x_2 \times p_2 + \ldots + x_n \times p_n \\ &= \sum_{i=1}^{n} x_i p_i \end{aligned}$$

Nous pouvons aussi rappeler les définitions de la variance et de l'écart-type d'une variable aléatoire.

Définition

Variance et écart-type d'une variable aléatoire :

La variance de la variable aléatoire X est le réel positif, noté $V(X)$, défini par :

$$\begin{aligned} V(X) &= p_1 \times (x_1 - E(X))^2 + p_2 \times (x_2 - E(X))^2 + \ldots \\ &\quad + p_n \times (x_n - E(X))^2 \\ &= \sum_{i=1}^{n} p_i (x_i - E(X))^2 \end{aligned}$$

→ L'écart-type de la variable aléatoire X est le réel positif, noté $\sigma(X)$, défini par :

$$\sigma(X) = \sqrt{V(X)}$$

Prenons un exemple pour bien revoir ces notions, en considérant le petit jeu suivant.

Exemple

On lance une pièce de monnaie parfaitement équilibrée. On gagne 5 euros si on obtient pile et on perd 3 euros si on obtient face.

1. Si X est la variable aléatoire correspondant au gain algébrique obtenu, alors l'ensemble des valeurs prises par X est $\{-3,\ 5\}$.
2. Nous pouvons donner la loi de probabilité de X, sous la forme d'un tableau :

x_i	-3	5
$p(X = x_i)$	$\frac{1}{2}$	$\frac{1}{2}$

3. Nous calculons maintenant l'espérance de X :

$$\begin{aligned} E(X) &= -3 \times p(X = -3) + 5 \times p(X = 5) \\ &= -3 \times \frac{1}{2} + 5 \times \frac{1}{2} \\ &= 1 \end{aligned}$$

→ On interprète ce résultat ($E(X) = 1$) comme le gain moyen que peut espérer un joueur sur un très grand nombre de parties.

4. Enfin, calculons la variance de X :

$$\begin{aligned} V(X) &= \frac{1}{2} \times (-3-1)^2 + \frac{1}{2} \times (5-1)^2 \\ &= \frac{1}{2} \times 16 + \frac{1}{2} \times 16 \\ &= 16 \end{aligned}$$

→ Nous pouvons donner aussi l'écart-type :

$$\begin{aligned} \sigma(X) &= \sqrt{V(X)} \\ &= 4 \end{aligned}$$

b. Variables aléatoires indépendantes

Comme nous l'avons fait pour les événements ou les épreuves, nous allons ici définir l'indépendance de 2 **variables aléatoires indépendantes**.

Définition

Indépendance de 2 variables aléatoires :

Prenons deux variables aléatoires X et Y respectivement définies sur les univers Ω_1 et Ω_2.
X et Y sont dites indépendantes si tout événement lié à la variable aléatoire X est indépendant de tout événement lié à la variable aléatoire Y.

→ C'est-à-dire si, pour tout $x \in X(\Omega_1)$ et pour tout $y \in Y(\Omega_2)$:

$$p((X = x) \cap (Y = y)) = p(X = x) \times p(Y = y)$$

Nous pouvons généraliser cette notion d'indépendance à n variables aléatoires.

À retenir

Prenons n variables aléatoires X_1, ..., X_n respectivement définies sur les univers Ω_1, ..., Ω_n.
Elles sont dites mutuellement indépendantes (ou indépendantes) si, pour tous $x_1 \in X_1(\Omega_1)$, ..., $x_n \in X_n(\Omega_n)$:

$$p((X_1 = x_1) \cap \ldots \cap (X_n = x_n)) = p(X_1 = x_1) \times \ldots \times p(X_n = x_n)$$

Attention

Les variables aléatoires deux à deux indépendantes ne sont pas nécessairement mutuellement indépendantes.

C. Transformation affine

Nous avons rappelé plus haut les formules pour calculer l'espérance et la variance d'une variable aléatoire. Ces formules peuvent entraîner des calculs assez lourds. Nous allons voir dans ce paragraphe des propriétés qui vont nous permettre d'alléger ces calculs.

Définition

Transformation affine d'une variable aléatoire :

Soit X une variable aléatoire définie sur un univers Ω, qui prend les valeurs x_1, x_2, ..., x_n. Et soit a et b deux réels.
La variable aléatoire Y définie par $Y = aX + b$ est aussi une variable aléatoire, qui prend, pour tout i allant de 1 à n, les valeurs $y_i = ax_i + b$.

Et nous avons alors les propriétés suivantes.

Propriété

Soit X une variable aléatoire.
Soit Y une variable aléatoire définie par $Y = aX + b$, avec a et b deux réels.
Nous avons alors :

- $E(Y) = aE(X) + b$,

  → c'est la linéarité de l'espérance ;

- $V(Y) = a^2V(X)$.

Nous allons maintenant voir un exemple qui va donner une méthodologie pour simplifier les calculs de l'espérance et de la variance d'une variable aléatoire.

Exemple

Une entreprise produit des vis. Un de ses modèles doit avoir pour pas de vis une distance théorique de $0,45\text{ mm}$ (c'est-à-dire que, pour chaque tour, la vis avancera, ou reculera, de $0,45\text{ mm}$).
Une vis est régulièrement tirée, au hasard, de la production, afin de mesurer le pas de vis. À cause des erreurs de mesure, ces distances peuvent légèrement varier.

Nous considérons la variable aléatoire X qui associe, à toute vis extraite, la mesure de son pas de vis, dont la loi de probabilité est donnée dans ce tableau :

x_i	$0,448$	$0,449$	$0,450$	$0,451$	$0,452$
$p(X = x_i)$	$0,1$	$0,2$	$0,2$	$0,3$	$0,2$

→ Nous allons calculer l'espérance et la variance de X.

1 **Définition de la variable Y**

Au vu des valeurs données par la loi de probabilité, nous nous rendons vite compte que les calculs de l'espérance et de la variance vont être assez lourds. Nous allons donc introduire une nouvelle variable aléatoire. Nous cherchons donc :

- à travailler avec des valeurs entières,
  - ➜ nous multiplions X par $1\,000$;
- à avoir une valeur centrale nulle,
  - ➜ nous soustrayons 450 à $1\,000X$.

Nous définissons ainsi la variable aléatoire $Y = 1\,000X - 450$.

2 Loi de probabilité de Y

Nous pouvons maintenant donner la loi de probabilité de Y, avec, pour tout i compris entre 1 et 5, $y_i = 1\,000x_i - 450$ (les probabilités ne changent évidemment pas) :

y_i	-2	-1	0	1	2
$p(Y = y_i)$	$0,1$	$0,2$	$0,2$	$0,3$	$0,2$

3 Espérance de Y

$$\begin{aligned} E(Y) &= -2 \times 0,1 + (-1) \times 0,2 + 0 \times 0,2 \\ &\quad + 1 \times 0,3 + 2 \times 0,2 \\ &= -0,2 - 0,2 + 0,3 + 0,4 \\ &= 0,3 \end{aligned}$$

4 Variance de Y

$$\begin{aligned} V(Y) &= 0,1 \times (-2 - 0,3)^2 + 0,2 \times (-1 - 0,3)^2 + 0,2 \times (0 - 0,3)^2 \\ &\quad + 0,3 \times (1 - 0,3)^2 + 0,2 \times (2 - 0,3)^2 \\ &= 0,1 \times 5,29 + 0,2 \times 1,69 + 0,2 \times 0,09 \\ &\quad + 0,3 \times 0,49 + 0,2 \times 2,89 \\ &= 1,61 \end{aligned}$$

C Indicateurs de X

Par linéarité de l'espérance, nous obtenons :

$$E(Y) = 1\,000E(X) - 450 \Leftrightarrow 0,3 = 1\,000E(X) - 450$$

→ Nous en déduisons l'espérance de X :

$$\begin{aligned} E(X) &= \frac{0,3+450}{1\,000} \\ &= 0,4503 \end{aligned}$$

Nous avons donc $Y = aX + b$, avec $a = 1\,000 = 10^3$. Avec la propriété sur la variance donnée ci-dessus, nous obtenons :

$$V(Y) = a^2V(X) \Leftrightarrow 1,61 = (10^3)^2 \times V(X)$$

→ Nous en déduisons la variance de X :

$$V(X) = 1,61 \times 10^{-6}$$

Pour vous convaincre de la simplification dans les calculs que l'introduction de Y permet, vous pouvez déterminer $E(X)$ et $V(X)$ sans passer par Y...

d. Somme de variables aléatoires

Pour bien comprendre cette notion de somme de variables aléatoires, nous allons suivre en fil rouge, tout au long de ce paragraphe, l'exemple détaillé d'un jeu qui se déroule en deux tirages indépendants (et obligatoires).

1. D'abord, dans une première urne contenant 5 boules rouges et 3 boules blanches (toutes indiscernables au toucher), on tire une boule.

→ Si la boule tirée est rouge (événement que nous notons R), on perd 5 points.

→ Si elle est blanche (événement B), on gagne 3 points.

2. Ensuite, dans une seconde urne contenant 2 boules noires, 1 boule jaune et 1 boule verte (elles aussi indiscernables au toucher), on tire encore une boule.

→ Si la boule tirée est noire (événement N), on perd 2 points.

→ Si elle est jaune (événement J), on gagne 4 points.

→ Si elle est verte (événement V), on gagne 6 points.

Considérons les variables aléatoires suivantes :

- X donne le nombre de points obtenus après le premier tirage ;
- Y donne le nombre de points obtenus après le second.

X est une variable aléatoire définie sur $\Omega_1 = \{R,\ B\}$.

→ $X(\Omega_1) = \{-5,\ 3\}$.

Y est une variable aléatoire définie sur $\Omega_2 = \{N,\ J,\ V\}$.

→ $Y(\Omega_2) = \{-2,\ 4,\ 6\}$.

Ce qui nous intéresse, c'est le nombre de points obtenus après les deux tirages. Pour cela, nous considérons maintenant la variable aléatoire Z, qui donne ce nombre total de points.

→ Logiquement, Z correspond alors à la somme des points obtenus après les deux tirages.

Définition

Somme de 2 variables aléatoires :

Soit X et Y deux variables aléatoires.
La variable aléatoire définie par $X + Y$ prend pour valeurs les sommes possibles des valeurs prises par X et Y.

Reprenons notre petit jeu.

→ Nous avons donc : $Z = X + Y$.

La variable aléatoire Z est définie sur l'univers :

$$\Omega = \{(R,\ N),\ (R,\ J),\ (R,\ V), (B,\ N),\ (B,\ J),\ (B,\ V)\}$$

Pour déterminer les valeurs que peut prendre Z en fonction des valeurs prises par X et Y, nous pouvons utiliser le tableau à deux entrées suivant :

Valeurs de Y → Valeurs de X ↓	-2	4	6
-5	-7	-1	1
3	1	7	9

→ Nous obtenons ainsi :

$$Z(\Omega) = \{-7,\ -1,\ 1,\ 7,\ 9\}$$

Astuce

Dans le paragraphe suivant, nous allons découvrir des propriétés qui nous permettront de calculer l'espérance et la variance de Z sans avoir à déterminer sa loi de probabilité.

➜ Nous allons néanmoins le faire, pour être complet et montrer comment faire.

En utilisant le tableau que nous venons de donner, nous voyons par exemple que, pour obtenir un total de -7 points, il faut d'abord tirer une boule rouge – on perd 5 points –, puis une boule noire – on perd 2 points :

$$\begin{aligned} p(Z = -7) &= p(R \cap N) \\ &= p((X = -5) \cap (Y = -2)) \end{aligned}$$

De la même façon, pour obtenir un total de 1 point, il faut soit tirer une boule rouge – on perd 5 points – et une boule verte – on gagne 6 points –, soit tirer une boule blanche – on gagne 3 points – et une boule noire – on perd 2 points.

$$\begin{aligned} p(Z = 1) &= p(R \cap V) + p(B \cap N) \\ &= p((X = -5) \cap (Y = 6)) + p((X = 3) \cap (Y = -2)) \end{aligned}$$

En outre, les tirages étant indépendants, X et Y sont indépendantes.

➜ Nous obtenons :

$$\begin{aligned} p(Z = -7) &= p(X = -5) \times p(Y = -2) \\ &= \frac{5}{8} \times \frac{1}{2} \\ &= \frac{5}{16} \end{aligned}$$

$$\begin{aligned} p(Z = 1) &= p(X = -5) \times p(Y = 6) + p(X = 3) \times p(Y = -2) \\ &= \frac{5}{8} \times \frac{1}{4} + \frac{3}{8} \times \frac{1}{2} \\ &= \frac{5}{32} + \frac{3}{16} \\ &= \frac{11}{32} \end{aligned}$$

Nous calculons de la même façon les autres probabilités, et nous obtenons la loi donnée par :

z_i	-7	-1	1	7	9
$p(Z = z_i)$	$\frac{5}{16}$	$\frac{5}{32}$	$\frac{11}{32}$	$\frac{3}{32}$	$\frac{3}{32}$

e. Indicateurs d'une somme de deux variables aléatoires

Comme nous l'avons fait pour la transformation affine, nous pouvons donner, pour la somme, les propriétés suivantes.

Propriété

Soit X et Y deux variables aléatoires.

Soit Z une variable aléatoire définie par $Z = X + Y$.

Alors l'espérance de Z est donnée par :

$$E(Z) = E(X) + E(Y)$$

➜ Nous parlons ici aussi de la linéarité de l'espérance.

Si de plus X et Y sont indépendantes, nous avons la relation d'additivité de la variance :

$$V(Z) = V(X) + V(Y)$$

Exemple

Reprenons une nouvelle fois le jeu précédent, et calculons l'espérance et la variance de Z.

1. Calculons d'abord celles de X.

$$\begin{aligned} E(X) &= -5 \times \frac{5}{8} + 3 \times \frac{3}{8} \\ &= -2 \end{aligned}$$

$$\begin{aligned} V(X) &= \frac{5}{8} \times (-5-(-2))^2 + \frac{3}{8} \times (3-(-2))^2 \\ &= \frac{5}{8} \times 9 + \frac{3}{8} \times 25 \\ &= 15 \end{aligned}$$

2. Calculons maintenant celles de Y.

$$\begin{aligned} E(Y) &= -2 \times \frac{1}{2} + 4 \times \frac{1}{4} + 6 \times \frac{1}{4} \\ &= \frac{6}{4} \\ &= 1,5 \end{aligned}$$

$$\begin{aligned} V(Y) &= \frac{1}{2} \times (-2 - 1,5)^2 + \frac{1}{4} \times (4 - 1,5)^2 + \frac{1}{4} \times (6 - 1,5)^2 \\ &= \frac{1}{2} \times 12,25 + \frac{1}{4} \times 6,25 + \frac{1}{4} \times 20,25 \\ &= 12,75 \end{aligned}$$

C Nous obtenons finalement, X et Y étant indépendantes :

$$\begin{aligned} E(Z) &= E(X) + E(Y) \\ &= -2 + 1,5 \\ &= -0,5 \end{aligned}$$

$$\begin{aligned} V(Z) &= V(X) + V(Y) \\ &= 15 + 12,75 \\ &= 27,75 \end{aligned}$$

$$\begin{aligned} \sigma(Z) &= \sqrt{V(Z)} \\ &\approx 5,268 \end{aligned}$$

On pourrait vérifier ces calculs en calculant $E(Z)$ et $V(Z)$ à partir de la loi de probabilité de Z que nous avons donnée plus haut.

➜ Mais les calculs, avec les probabilités que nous avons trouvées, seraient lourds.

Astuce

Pour calculer les indicateurs d'une variable aléatoire, il est parfois plus simple de la décomposer en une somme de deux autres variables aléatoires, dont les espérance et variance sont plus simples à calculer.

➜ Pour la variance, il ne faudra surtout pas oublier de montrer l'indépendance des variables aléatoires.

Imaginons ainsi que, en guise d'exercice, nous ayons eu le même jeu détaillé en énoncé, et que l'on nous ait demandé si ce jeu est équitable.

Il est alors inutile de calculer la loi de probabilités de Z, puis, au terme de calculs assez fastidieux, de donner l'espérance.

→ Il suffit de bien identifier X et Y, l'espérance de chacune étant plus simple à calculer, puis d'utiliser la propriété de linéarité de l'espérance.

Pour précision finale, l'espérance de Z étant égale à $-0,5$, le jeu n'est pas équitable et il est en défaveur du joueur.

2 Espérance, variance et écart-type de la loi binomiale

Dans le cours précédent, nous avons découvert la loi de Bernoulli et la loi binomiale. Nous allons ici donner les propriétés pour déterminer les indicateurs d'une variable aléatoire suivant cette dernière. Puis nous donnerons un petit exemple pour appliquer les propriétés données ci-dessous.

a. Propriétés

Propriété

Soit X_1, X_2, ..., X_n des variables aléatoires indépendantes définies sur le même univers et suivant toutes la même loi de Bernoulli de paramètre p, notée $\mathcal{B}(p)$.
Alors, leur somme $X_1 + X_2 + \ldots + X_n$ suit une loi binomiale de paramètres n et p, notée $\mathcal{B}(n,\ p)$.

Démonstration

On rappelle que la loi binomiale compte le nombre de succès dans un schéma de Bernoulli.

Chaque X_i $(1 \leq i \leq n)$ prend deux valeurs, 0 ou 1, avec une probabilité p de prendre la valeur 1.
Donc, lorsque l'on réalise successivement n épreuves indépendantes de Bernoulli, la somme (des résultats) $X_1 + X_2 + \ldots + X_n$ est égale au nombre de succès obtenus lors de ces épreuves.

→ $X_1 + X_2 + \ldots + X_n$ suit une loi binomiale de paramètres n et p.

Nous admettons la propriété suivante.

Propriété

Soit X une variable aléatoire suivant une loi binomiale $\mathcal{B}(n, p)$.
Alors X peut s'écrire comme une somme de variables aléatoires indépendantes suivant la même loi

de Bernoulli.

➜ C'est-à-dire que l'on a :

$$X = X_1 + X_2 + \ldots + X_n$$

[avec $X_1, X_2, \ldots, X_n$ des variables aléatoires suivant la loi $\mathcal{B}(p)$]

Nous pouvons maintenant donner les propriétés qui nous permettront de calculer l'espérance et la variance d'une variable aléatoire suivant une loi binomiale.

Propriété

Soit X une variable aléatoire suivant une loi binomiale de paramètres n et p.
Alors, on a :

- $E(X) = np$;
- $V(X) = np(1-p)$;
- $\sigma(X) = \sqrt{V(X)} = \sqrt{np(1-p)}$ ($\sigma(X)$ est l'écart-type de X).

En utilisant les propriétés que nous avons vues dans ce cours et celles pour l'espérance et la variance d'une variable aléatoire suivant une loi de Bernoulli, nous pouvons démontrer facilement ces dernières propriétés.

Démonstration

Par hypothèse, X suit une loi binomiale de paramètres n et p.
Donc, d'après la propriété vue plus haut, il existe n variables aléatoires de Bernoulli indépendantes X_1, X_2, ..., X_n telles que :

$$X = X_1 + X_2 + \ldots + X_n$$

Par ailleurs, pour tout entier i tel que $1 \leq i \leq n$, on a : $E(X_i) = p$ et $V(X_i) = p(1-p)$, car chaque X_i suit une loi de Bernoulli de paramètre p.

➜ Dans ces conditions, nous avons :

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \ldots + X_n) \\ &= E(X_1) + E(X_2) + \ldots + E(X_n) \\ &\text{[par linéarité de l'espérance]} \\ &= \underbrace{p + p + \ldots + p}_{n \text{ fois}} \\ &= np \end{aligned}$$

En outre, X_1, X_2, ..., X_n sont indépendantes.

➜ Nous avons alors :

$$\begin{aligned} V(X) &= V(X_1 + X_2 + \cdots + X_n) \\ &= V(X_1) + V(X_2) + \cdots + V(X_n) \\ &\text{[par la relation d'additivité de la variance]} \\ &= \underbrace{p(1-p) + p(1-p) + \ldots + p(1-p)}_{n \text{ fois}} \\ &= np(1-p) \end{aligned}$$

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

À retenir

Il est plus pratique d'utiliser ces dernières propriétés pour calculer l'espérance, la variance et l'écart-type d'une variable aléatoire suivant une loi binomiale.
Pour cela, on commence toujours par justifier que X suit une loi binomiale de paramètres n et p, à déterminer.

Nous allons traiter un exemple, pour montrer une application de cette méthode.

b. Exemple

Un institut politique souhaite constituer des groupes de réflexion, chacun de 150 Français, et s'intéresse particulièrement à leur participation au premier tour des dernières élections.
Ces 150 personnes sont choisies aléatoirement à partir des listes électorales de l'ensemble du territoire français.

➜ On admet que le nombre d'inscrits à ces listes est suffisamment grand pour considérer que le « tirage » des 150 noms est effectué avec remise.

Connaissant le taux de participation à ce premier tour, que nous considérons comme égal à $45\,\%$, l'institut souhaite étudier la variable aléatoire X qui donne, parmi les 150 personnes tirées au sort du groupe, le nombre de celles et ceux qui sont allés voter.

1 **Montrons que X suit une loi binomiale de paramètres n et p, et déterminons ces derniers.**

- Il s'agit donc de tirer au sort une personne et de regarder si elle a voté. Ce tirage est répété 150 fois et, pour chaque tirage, nous avons 2 issues :
  - « La personne choisie a voté », considérée comme succès (S) ;
  - « La personne choisie n'a pas voté », considérée comme échec (E).
- Nous considérons que le tirage se fait avec remise. Il y a donc 150 expériences identiques et indépendantes.
- La variable aléatoire compte le nombre de personnes qui ont voté parmi les 150 personnes choisies, soit le nombre de succès.
- Enfin, la proportion des votants parmi la population des inscrits est égale à $0,45$.

➜ X suit une loi binomiale de paramètres $n = 150$ et $p = 0,45$.

2 **Écrivons X sous la forme de variables aléatoires.**

La propriété donnée plus haut nous dit que nous pouvons alors écrire X comme somme de variables aléatoires indépendantes suivant la même loi de Bernoulli.

➜ Nous obtenons ainsi :

$$X = X_1 + X_2 + \ldots + X_{150}$$
$$\text{[avec } X_1,\ X_2,\ \ldots,\ X_{150} \text{ des variables aléatoires suivant la loi de Bernoulli de paramètre } p = 0,45\text{]}$$

3 **Calculons l'espérance de X.**

Nous connaissons donc les paramètres de la loi binomiale suivie par X.

➜ Grâce à la propriété que nous avons vue, nous avons :

$$\begin{aligned} E(X) &= np \\ &= 150 \times 0,45 \\ &= 67,5 \end{aligned}$$

4 **Calculons enfin la variance et l'écart-type de X.**

De la même façon, nous avons :

$$\begin{aligned} V(X) &= np(1-p) \\ &= 150 \times 0,45 \times 0,55 \\ &= 37,125 \end{aligned}$$

$$\begin{aligned} \sigma(X) &= \sqrt{V(X)} \\ &\approx 6,093 \end{aligned}$$

La variable aléatoire X est d'espérance $67,5$ et d'écart-type proche de $6,093$.

Il est ainsi inutile de déterminer les probabilités de toutes les valeurs que peut prendre X pour calculer son espérance et son écart-type. Ce serait d'ailleurs fastidieux, car X peut prendre au total 151 valeurs différentes, de 0 à 150 !

3 Somme de variables aléatoires identiques

Plus généralement, nous allons voir dans cette partie comment calculer l'espérance et la variance d'une variable aléatoire X qui s'écrit comme somme de variables aléatoires indépendantes suivant une même loi de probabilité.

Comme dans la deuxième partie, nous allons d'abord donner les définitions et les propriétés, avant de les appliquer sur un exemple.

a. Définition et propriétés

Définition

Échantillon de taille n d'une loi de probabilité :

Soit n un entier naturel non nul.

On appelle échantillon de taille n d'une loi de probabilité une liste $(X_1,\ X_2,\ \ldots,\ X_n)$ de variables aléatoires indépendantes suivant toutes cette loi.

Exemple

On lance 4 dés équilibrés et on note X_i (avec $1 \leq i \leq 4$) la variable aléatoire correspondant au résultat du dé numéro i.

Les dés sont équilibrés, donc ces variables aléatoires suivent la même loi de probabilité, donnée par le tableau suivant :

x_i	1	2	3	4	5	6
$p(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

➜ La liste $(X_1,\ X_2,\ X_3,\ X_4)$ est donc un échantillon, de taille 4, de variables aléatoires associé à la cette loi.

À retenir

Soit $(X_1,\ X_2,\ \dots,\ X_n)$ un échantillon de taille n d'une loi de probabilité.

- On note S_n la **somme** de ces variables aléatoires et :

$$S_n = X_1 + X_2 + \dots + X_n$$

- On note M_n la **moyenne** de ces variables aléatoires et :

$$\begin{aligned} M_n &= \frac{X_1 + X_2 + \dots + X_n}{n} \\ &= \frac{S_n}{n} \end{aligned}$$

Propriété

Soit $(X_1,\ X_2,\ \dots,\ X_n)$ un échantillon de taille n d'une loi de probabilité.
Soit E, V et σ respectivement l'espérance, la variance et l'écart-type de la variable suivant cette loi de probabilité.
Nous avons alors :

- $E(S_n) = nE$;
- $V(S_n) = nV$;
- $\sigma(S_n) = \sqrt{n}\,\sigma$.

Démonstration

Comme les n variables aléatoires suivent la même loi, elles ont donc la même espérance, que nous notons E.

➜ Par la propriété de la linéarité de l'espérance, on obtient :

$$\begin{aligned} E(S_n) &= E(X_1 + X_2 + \ldots + X_n) \\ &= E(X_1) + E(X_2) + \ldots + E(X_n) \\ &= nE \end{aligned}$$

Les n variables aléatoires ont aussi la même variance et le même écart-type, que nous notons respectivement V et $\sigma = \sqrt{V}$. Les variables aléatoires sont indépendantes.

➜ On peut donc utiliser la propriété d'additivité de la variance :

$$\begin{aligned} V(S_n) &= V(X_1 + X_2 + \ldots + X_n) \\ &= V(X_1) + V(X_2) + \ldots + V(X_n) \\ &= nV \end{aligned}$$

$$\begin{aligned} \sigma(S_n) &= \sqrt{nV} \\ &= \sqrt{n}\sqrt{V} \text{ [} V \text{ et } n \text{ sont positifs]} \\ &= \sqrt{n}\,\sigma \end{aligned}$$

Propriété

Soit $(X_1,\ X_2,\ \ldots,\ X_n)$ un échantillon de taille n d'une loi de probabilité.
Soit E, V et σ respectivement l'espérance, la variance et l'écart-type de la variable suivant cette loi de probabilité.
Nous avons alors :

- $E(M_n) = E$;
- $V(M_n) = \frac{V}{n}$;
- $\sigma(M_n) = \frac{\sigma}{\sqrt{n}}$.

Démonstration

Nous allons nous servir des propriétés que nous venons de voir pour S_n :

$$
\begin{aligned}
E(M_n) &= E\left(\frac{S_n}{n}\right) \\
&= \frac{1}{n} \times E(S_n) \text{ [par linéarité de l'espérance]} \\
&= \frac{1}{n} \times n \times E \text{ [car } E(S_n) = nE\text{]} \\
&= E
\end{aligned}
$$

$$
\begin{aligned}
V(M_n) &= V\left(\frac{S_n}{n}\right) \\
&= \frac{1}{n^2} \times V(S_n) \text{ [car } V(aX) = a^2V(X)\text{]} \\
&= \frac{1}{n^2} \times n \times V \text{ [car } V(S_n) = nV\text{]} \\
&= \frac{V}{n}
\end{aligned}
$$

$$
\begin{aligned}
\sigma(M_n) &= \sqrt{\frac{V}{n}} \\
&= \frac{\sqrt{V}}{\sqrt{n}} \text{ [}V\text{ et }n\text{ sont positifs]} \\
&= \frac{\sigma}{\sqrt{n}}
\end{aligned}
$$

b. Exemple

Considérons un nouveau jeu, avec une roue (non truquée) qui peut tourner sur elle-même et une aiguille (qui, elle, est fixe). Cette roue est un disque divisé en trois portions :

- la portion noire représente un quart de l'aire du disque ;
- la blanche représente aussi un quart de l'aire du disque ;
- la verte représente la moitié de l'aire du disque.

Contre une somme de 2 euros, le joueur peut lancer la roue à 4 reprises.
À chaque lancer de roue, les conditions sont les suivantes :

- si l'aiguille indique la portion noire (événement N et $p(N) = \frac{1}{4}$), le joueur perd 5 euros ;
- si l'aiguille indique la portion blanche (événement B et $p(B) = \frac{1}{4}$), le joueur ne gagne ni ne perd rien ;
- si l'aiguille indique la portion verte (événement V et $p(V) = \frac{1}{2}$), le joueur gagne 3 euros.

Définissons la variable aléatoire X qui compte le gain algébrique à la fin du jeu, c'est-à-dire après 4 lancers.

1. **Exprimons X comme somme de variables aléatoires.**

Nous voyons que X représente le gain après 4 répétitions d'une même expérience : le lancer d'une roue.

Soit X_i $(1 \leq i \leq 4)$ la variable aléatoire qui donne le gain algébrique obtenu après le lancer i.
Nous pouvons alors écrire :

$$\begin{aligned} X &= \sum_{i=1}^{4} X_i \\ &= X_1 + X_2 + X_3 + X_4 \end{aligned}$$

➜ Ces 4 variables aléatoires sont indépendantes et suivent la même loi de probabilité.

2. **Déterminons cette dernière, ainsi que ses indicateurs.**

Considérons la variable aléatoire $Y = X_1 = X_2 = X_3 = X_4$.

➜ Sa loi de probabilité est simple à déterminer et elle est donnée par le tableau suivant :

y_i	-5	0	3
$p(Y = y_i)$	$0,25$	$0,25$	$0,5$

Nous calculons maintenant son espérance, sa variance et son écart-type :

$$\begin{aligned} E(Y) &= -5 \times 0,25 + 0 \times 0,25 + 3 \times 0,5 \\ &= 0,25 \end{aligned}$$

$$\begin{aligned} V(Y) &= 0,25 \times (-5 - 0,25)^2 + 0,25 \times (0 - 0,25)^2 + 0,5 \times (3 - 0,25)^2 \\ &= 10,6875 \end{aligned}$$

$$\begin{aligned} \sigma(Y) &= \sqrt{V(Y)} \\ &= \sqrt{10,875} \approx 3,27 \end{aligned}$$

3. **Déterminons les indicateurs de X.**

Pour cela, nous utilisons tout simplement (et rapidement) les propriétés que nous venons de voir, avec $n = 4$:

$$\begin{aligned} E(X) &= nE(Y) \\ &= 4 \times 0,25 \\ &= 1 \end{aligned}$$

$$\begin{aligned} V(X) &= nV(Y) \\ &= 42,75 \end{aligned}$$

$$\begin{aligned} \sigma(Y) &= \sqrt{n}\,\sigma(Y) \\ &= 2\sqrt{10,875} \approx 6,54 \end{aligned}$$

C **Donnons, pour terminer, quelques interprétations simple de ces résultats.**

L'espérance de X est égale à 1. Nous pourrions commettre l'erreur de considérer le jeu comme à l'avantage du joueur.
Mais, attention, X ne prend pas en compte le prix à payer pour pouvoir jouer ! Pour se faire une meilleure idée, il conviendrait donc d'étudier la variable aléatoire $Z = Y - 2$, qui retranche au gain final le droit payé au début.
Par linéarité de l'espérance, nous obtenons :

$$\begin{aligned} E(Z) &= E(X - 2) \\ &= E(X) - 2 \\ &= 1 - 2 \\ &= -1 \end{aligned}$$

→ Ainsi, en réalité, le jeu est en défaveur du joueur !

Par ailleurs, nous avons trouvé pour X un écart-type relativement important (pour Z aussi, car $\sigma(Z) = \sigma(X)$).

→ Les valeurs sont assez dispersées autour de l'espérance.

En effet, par exemple, si nous tombons à 4 reprises sur la portion noire, notre perte finale s'élèvera à $20 + 2 = 22 \text{ euros}$. Et si nous tombons à 4 reprises sur la portion verte, notre gain final s'élèvera à $12 - 2 = 10 \text{ euros}$.

Conclusion :

Dans ce cours, nous avons vu comment définir une somme de variables aléatoires. Nous avons également vu que les calculs de l'espérance et de la variance d'une variable aléatoire peuvent être simplifiés, par transformation affine ou en la décomposant comme somme de variables aléatoires indépendantes.
Et c'est grâce à cette décomposition que nous avons pu calculer l'espérance et la variance de la loi binomiale.