Chapitre 10
Somme de variables aléatoires

I. Somme de deux variables aléatoires

1) Variable aléatoire

Définitions :

- L'ensemble des issues d'une expérience aléatoire s'appelle l'**univers** de l'expérience.
  On le note Ω.
- Une **variable aléatoire réelle** est une fonction définie sur Ω et à valeurs dans $\mathbb{R}$.
  On la note X.

On a donc
$$\begin{array}{rl} X : \Omega & \to \mathbb{R} \\ \omega & \mapsto X(\omega) \end{array}$$

Exemple :

On lance une pièce de monnaie. Si on obtient pile, on gagne 5 € et si on obtient face, on gagne 2 €.

On peut alors définir une variable aléatoire X correspondant au gain obtenu en euro.

X est défini sur l'univers $\Omega = \{\text{pile, face}\}$.

On a alors $X(\text{pile}) = 5$ et $X(\text{face}) = 2$.

X peut prendre deux valeurs : 5 et 2.

2) Transformation affine

Définition :

Soit X une variable aléatoire définie sur l'univers Ω et a et b deux nombres réels.

On peut définir une variable aléatoire réelle Y telle que, pour tout élément $\omega \in \Omega$,

$$Y(\omega) = aX(\omega) + b$$

On note $Y = aX + b$.

Exemple :

- On lance un dé équilibré à six faces et on joue au jeu suivant : le nombre de points obtenus est le résultat du dé multiplié par 5 auquel on ajoute 3.

  En notant respectivement X et Y les variables aléatoires correspondant au résultat du dé et aux points obtenus, on a alors $Y = 5X + 3$.
- X est une variable aléatoire qui suit la loi binomiale de paramètres $n = 2$ et $p = 0{,}2$.

  $Y = 4X - 0{,}6$ est une variable aléatoire dont les valeurs possibles sont $\{-0{,}6 ; 3{,}4 ; 7{,}4\}$.

3) Somme

Définition :

Soient X et Y, deux variables aléatoires définies sur l'univers Ω.

On peut définir une variable aléatoire Z sur Ω telle que, pour tout élément $\omega \in \Omega$,

$$Z(\omega) = X(\omega) + Y(\omega).$$

Cette variable aléatoire est appelée **somme des variables aléatoires X et Y**.

On note Z = X + Y.

Exemples :

- On lance cinq dés équilibrés et on compte la somme des nombres obtenus.

  Soit X la variable aléatoire correspondant à cette somme.

  Alors on peut écrire X sous la forme $X = X_1 + X_2 + \ldots + X_5$ où,

  pour tout $k \in \{1 ; 2 ; 3 ; 4 ; 5\}$, X_k correspond au résultat du dé numéro k.

  L'ensemble des valeurs prises par X est $\{5 ; 6 ; 7 ; 8 ; 9 ; \ldots ; 30\}$.

  On remarque que $X \neq 5X_1$.
- X est une variable aléatoire qui suit la loi binomiale de paramètres $n = 200$ et $p = 0,2$ et Y une variable aléatoire qui suit la loi binomiale de paramètres $n = 100$ et $p = 0,5$.

  X + Y est une variable aléatoire dont les valeurs possibles sont $\{0 ; 1 ; 2 ; \ldots ; 300\}$.
- On lance 20 fois une pièce de monnaie et on note X la variable aléatoire comptant le nombre de pile obtenu.

  On peut écrire la variable aléatoire X sous la forme $X = X_1 + X_2 + \ldots + X_{20}$ où,

  pour tout $k \in \{1 ; 2 ; \ldots ; 20\}$, $X_k = 1$ si on a obtenu pile au k^e lancer et $X_k = 0$ si on a obtenu face au k^e lancer.

II. Caractéristiques des variables aléatoires

Dans cette partie, on considère une variable aléatoire X définie sur $\Omega = \{\omega_1 ; \omega_2 ; \ldots ; \omega_r\}$ et on note $\{x_1 ; x_2 ; \ldots ; x_s\}$ l'ensemble des valeurs prises par X où r et s sont des entiers naturels non nuls.

$$E(X)=\sum_{i=1}^{s} x_i \times p(X=x_i) \quad \text{et} \quad V(X)=\sum_{i=1}^{s} (x_i - E(X))^2 \times p(X=x_i)$$

1) Espérance

Propriété :

En reprenant les notations précédentes, on a $E(X)=\sum_{j=1}^{r} X(\omega_j)\, p(\{\omega_j\})$.

Remarque :

Dans cette propriété, l'espérance s'écrit en fonction des issues ω_i de l'expérience aléatoire et non en fonctions des valeurs x_i .

Exemple :

On jette un dé cubique équilibré, on gagne 2 € si on obtient un nombre pair et on perd 6 € si on obtient un nombre impair.

L'espérance de la variable aléatoire X correspondant au gain remporté s'élève à :

$$E(X)=X(1)\times p(\{1\})+\ldots+X(6)\times p(\{6\})=(-6)\times\frac{1}{6}+\ldots+2\times\frac{1}{6} = -2\text{ €}$$

Propriété :

Soit X et Y deux variables aléatoires définies sur le même univers Ω. Alors :

E(X + Y) = E(X) + E(Y)

Démonstration :

Soient X et Y deux variables aléatoires définies sur Ω. Soit Z la variable aléatoire définie sur Ω par Z = X + Y.

On a alors E(X + Y) = E(Z) = $\sum_{j=1}^{r} Z(\omega_j)\,p(\{\omega_j\})$ et donc E(X + Y) = $\sum_{j=1}^{r}(X+Y)(\omega_j)\,p(\{\omega_j\})$.

On a, par ailleurs, (X + Y)(ω_j) = X(ω_j) + Y(ω_j) .

Donc E(X + Y) = $\sum_{j=1}^{r} X(\omega_j)\,p(\{\omega_j\})+\sum_{j=1}^{r} Y(\omega_j)\,p(\{\omega_j\})$. D'où E(X + Y) = E(X) + E(Y).

Remarque :

Cette propriété permet de déterminer l'espérance de X + Y simplement à l'aide de celles de X et Y (donc sans la connaissance de la loi de probabilité de X + Y).

Exemple :

X est une variable aléatoire qui suit la loi binomiale de paramètres n = 200 et p = 0,2 et Y une variable aléatoire qui suit la loi binomiale de paramètres n = 100 et p = 0,5.

E(X + Y) = E(X) + E(Y) = 200 × 0,2 + 100 × 0,5 = 90

Propriété :

Soit X une variable aléatoire et Y la variable aléatoire définie par Y = aX + b, où a et b sont deux réels. Alors :

E(Y) = E(aX + b) = aE(X) + b

Démonstration :

Si a = 0, on a E(0X + b) = E(b) = b (car Y prend la valeur b et p(X = b) = 1) et 0 × E(X) + b = b

Si $a \neq 0$, en notant x_1 ; x_2 ; … ; x_s les valeurs prises par X, alors aX + b prend les valeurs $ax_1 + b$; $ax_2 + b$; … ; $ax_s + b$.

Par définition, E(aX + b) = $\sum_{i=1}^{s}(ax_i+b)\,p(aX+b=ax_i+b)$.

Or $aX + b = ax_i + b$, si et seulement si, $X = x_i$, donc $p(aX+b=ax_i+b)=p(X=x_i)$.

Ainsi $E(aX + b) = \sum_{i=1}^{s}(ax_i+b)\,p(X=x_i)=\sum_{i=1}^{s}ax_i\,p(X=x_i)+\sum_{i=1}^{s}b\,p(X=x_i)$

$E(aX + b) = a\times\sum_{i=1}^{s}x_i\,p(X=x_i)+b\times\sum_{i=1}^{s}p(X=x_i) = aE(X) + b$

Exemple :

X est une variable aléatoire qui suit la loi binomiale de paramètres $n = 2$ et $p = 0,2$ et soit

Y = 4X – 0,6.

E(Y) = E(4X – 0,6) = 4 E(X) – 0,6 = 4 × (2 × 0,2) – 0,6 = 1

2) Variance

Propriété :

Soit X une variable aléatoire et Y la variable aléatoire définie par $Y = aX + b$, où a et b sont deux réels. Alors :

$$V(Y) = V(aX + b) = V(aX) = a^2\,V(X)$$

Démonstration :

Si $a = 0$, on a $V(0X + b) = V(b) = b$ (car Y prend la valeur b et $E(Y) = b$ donc $V(Y) = 0$)

et $0^2 \times V(Y) = b$

Si $a \neq 0$, en notant x_1 ; x_2 ; … ; x_s les valeurs prises par X, alors $aX + b$ prend les valeurs $ax_1+ b$; $ax_2 + b$; … ; $ax_s + b$.

Par définition, $V(aX + b) = \sum_{i=1}^{s}(ax_i+b-E(aX+b))^2\,p(aX+b=ax_i+b)$.

Or $aX + b = ax_i + b$, si et seulement si, $X = x_i$, donc $p(aX+b=ax_i+b)=p(X=x_i)$.

Ainsi $V(aX + b) = \sum_{i=1}^{s}(ax_i+b-aE(X)+b)^2\,p(X=x_i) = \sum_{i=1}^{s}(ax_i-aE(X))^2\,p(X=x_i)$

$V(aX + b) = \sum_{i=1}^{s}a^2(x_i-E(X))^2\,p(X=x_i)=a^2\times\sum_{i=1}^{s}(x_i-E(X))^2\,p(X=x_i) = a^2V(X)$

Remarque :

L'écart type $\sigma(aX + b)$ vérifie $\sigma(aX + b) = |a|\sigma(X)$.

Exemple :

X est une variable aléatoire qui suit la loi binomiale de paramètres $n = 2$ et $p = 0,2$ et soit

Y = 4X – 0,6.

$V(Y) = V(4X – 0,6) = V(4X) = 4^2 \times V(X) = 16 \times (2 \times 0,2 \times 0,8) = 5,12$

Définitions :

Soient $X_1, X_2, \dots X_n$, n variables aléatoires à valeurs respectivement dans $E_1, E_2, \dots E_n$.

On dit que $X_1, X_2, \dots X_n$ sont **indépendantes** lorsque, pour tout $x_1 \in E_1$, $x_2 \in E_2$, … , $x_n \in E_n$:

$$p(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \dots \cap \{X_n = x_n\}) = p(X_1 = x_1) \times p(X_2 = x_2) \times \dots \times p(X_n = x_n)$$

Remarque :

Si les variables $X_1, X_2, \dots X_n$ sont deux à deux indépendantes, on ne peut pas en conclure que $X_1, X_2, \dots X_n$ sont mutuellement indépendantes.

Propriété :

Si X et Y sont deux variables aléatoires **indépendantes** définies sur Ω, alors :

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

Remarques :

- Dans le cas où les expériences ne sont pas indépendantes, il se peut que :

$$V(X + Y) \neq V(X) + V(Y).$$

- Si $X_1, X_2, \dots X_n$ sont n variables aléatoires indépendantes définies sur Ω, alors :

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

III. Applications

1) Application à la loi binomiale

Dans cette partie, n désigne un entier naturel supérieur ou égal à 1 et p un nombre réel appartenant à l'intervalle [0 ; 1].

Définition :

Deux variables aléatoires sont dites **identiquement distribuées** lorsqu'elles ont la même loi de probabilité.

Remarque :

Deux variables aléatoires identiquement distribuées peuvent être ou ne pas être indépendantes.

Propriété :

Toute variable aléatoire suivant une loi binomiale peut s'écrire comme une somme de variables aléatoires de Bernoulli indépendantes et identiquement distribuées.

Propriétés :

Si X suit la loi binomiale de paramètres n et p, alors :

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

Démonstrations :

- Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p.

  Alors, il existe n variables aléatoires de Bernoulli de paramètre p telles que :

  $X = X_1 + X_2 + \dots + X_n$.

  Ainsi pour tout $k \in \{1 ; 2 ; \dots ; n\}$, $E(X_k) = p$ et $V(X_k) = p(1-p)$.

  Or, $E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = p + p + \dots + p = np$.
- Les variables aléatoires X_1, X_2, … X_n étant indépendantes, par définition de schéma de Bernoulli, on a :

  $V(X) = V(X_1) + V(X_2) + \dots + V(X_n) = p(1-p) + p(1-p) + \dots + p(1-p) = np(1-p)$.
- $\sigma(X) = \sqrt{V(x)} = \sqrt{np(1-p)}$.

Exemple :

Soit X une variable aléatoire suivant une loi binomiale de paramètres $n = 20$ et $p = 0{,}2$.

On a $E(X) = np = 20 \times 0{,}2 = 4$.

De plus $V(X) = np(1-p) = 20 \times 0{,}2 \times 0{,}8 = 3{,}2$.

Donc $\sigma(X) = \sqrt{V(x)} = \sqrt{3{,}2} \approx 1{,}789$

2) Échantillons de *n* variables aléatoires identiques et indépendantes

On considère un entier naturel $n \geqslant 1$ et X_1, X_2, … X_n, n variables aléatoires définies sur Ω supposées indépendantes et identiquement distribuées.

On note $S_n = X_1 + X_2 + \dots + X_n$ la somme de ces n variables aléatoires et $M_n = \dfrac{X_1 + X_2 + \dots + X_n}{n}$ la moyenne de ces n variables aléatoires.

Propriété :

Pour tout $k \in \{1 ; 2 ; \dots ; n\}$, on a :

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

Démonstrations :

- La linéarité de l'espérance donne $E(S_n) = E(X_1) + E(X_2) + \dots + E(X_n)$. Or ces variables aléatoires suivent la même loi. Elles ont donc la même espérance.

  D'où, pour tout $k \in \{1 ; 2 ; \dots ; n\}$, $E(S_n) = nE(X_k)$.
- De la même manière, les variables aléatoires X_1, X_2, … X_n étant supposées indépendantes, on obtient, pour tout $k \in \{1 ; 2 ; \dots ; n\}$, $V(S_n) = V(X_1) + V(X_2) + \dots + V(X_n) = nV(X_k)$.

  Enfin, on a $\sigma(S_n) = \sqrt{n}\,\sigma(X_k)$.

Remarque :

Cette propriété généralise les résultats obtenus sur la loi binomiale en considérant, dans ce cas, la variable aléatoire X comme somme de variables aléatoires de Bernoulli indépendantes de paramètres p.

Propriété :

Pour tout $k \in \{1 ; 2 ; \dots ; n\}$, on a :

- $E(M_n) = E(X_k)$
- $V(M_n) = \dfrac{V(X_k)}{n}$ et $\sigma(M_n) = \dfrac{\sigma(X_k)}{\sqrt{n}}$.

Démonstrations :

- Soit $k \in \{1 ; 2 ; \dots ; n\}$, la linéarité de l'espérance et la propriété précédente donnent $E(M_n) = E\left(\dfrac{X_1+X_2+\ldots+X_n}{n}\right)=E\left(\dfrac{S_n}{n}\right)=\dfrac{1}{n}E(S_n)=\dfrac{1}{n}\times n\,E(X_k) = E(X_k)$.
- Par ailleurs, pour tout $a \in \mathbb{R}$, $V(aS_n) = a^2V(S_n)$.

  En combinant cette égalité au résultat de la propriété précédente,

  $V(M_n) = V\left(\dfrac{S_n}{n}\right)=\dfrac{1}{n^2}V(S_n)=\dfrac{1}{n^2}n\,V(X_k)=\dfrac{V(X_k)}{n}$. On obtient ensuite $\sigma(M_n) = \dfrac{\sigma(X_k)}{\sqrt{n}}$.

Remarques :

- $E(M_n)$ peut s'interpréter comme ceci : en prenant un grand nombre de fois des échantillons de taille n et en calculant, à chaque fois la moyenne de l'échantillon obtenu, la moyenne théorique de ces résultats est égale à $E(X_k)$.
- $V(M_n) = \dfrac{V(X_k)}{n}$ montre que la variance diminue quand la taille de l'échantillon augmente. Elle quantifie la fluctuation d'échantillonnage, c'est-à-dire l'écart moyen entre les valeurs prises par la variable aléatoire et son espérance.

Exemple :

Soit X la variable aléatoire qui, à chaque paquet de chips issue d'une chaîne de production, associe sa masse en grammes. On note X_i la variable aléatoire qui, à chaque lot de 3 paquets de chips, associe la masse du i-ème paquet.

Les variables aléatoires X_1, X_2 et X_3 sont indépendantes et suivent la même loi que X, donc (X_1, X_2, X_3) est un échantillon de taille 3 de la loi de X.

La variable aléatoire somme $S_3 = X_1 + X_2 + X_3$ associe, à chaque lot sa masse en grammes.

La variable aléatoire moyenne $M_3 = \dfrac{X_1+X_2+X_3}{3}$ associe, à chaque lot de 3 paquets, la masse moyenne d'un paquet.

$E(M_3) = E(X)$ et $V(M_3) = \dfrac{1}{3}V(X)$.