Somme de variables aléatoires

1. Prenons l'exemple suivant:
Une urne contient cinq boules numérotées de 1 à 5 et une deuxième urne contient trois boules numérotées de 1 à 3. On tire au hasard une boule de la première urne et une boule de la deuxième urne.
La variable aléatoire X est égale au chiffre de la boule de la première urne et la variable aléatoire Y est égale au chiffre de la boule de la deuxième urne.
Par exemple, l'évènement $(X = 1) \cap (Y = 2)$ signifie qu'on a tiré la boule numérotée 1 dans la première urne et la boule numérotée 2 dans la deuxième urne.
On définit alors la variable aléatoire X + Y (somme des deux variables aléatoires).
X + Y prend les valeurs 2, 3, 4, 5, 6, 7, 8.
En effet, on a par exemple $X + Y = 3$ avec $(X = 1) \cap (Y = 2)$ ou $(X = 2) \cap (Y = 1)$.
Pour calculer par exemple, la probabilité de l'évènement $X + Y = 3$, on effectue le calcul :
$P(X + Y = 3) = P(X = 1 \cap Y = 2) + P(X = 2 \cap Y = 1)$.
Si de plus, les variables aléatoires X et Y sont indépendantes, alors on a :
$P(X + Y = 3) = P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1) = \dfrac{1}{5} \times \dfrac{1}{3} + \dfrac{1}{5} \times \dfrac{1}{3} = \dfrac{2}{15}$.

2. <u>Définition :</u> Soit X et Y deux variables aléatoires. La loi de probabilité de la variable aléatoire somme X + Y est donnée par : $P(X + Y = k) = \sum_{i+j=k} P((X = i) \cap (Y = j))$.
Si, de plus, les variables aléatoires X et Y sont indépendantes, alors on a :
$P(X + Y = k) = \sum_{i+j=k} P(X = i) \times P(Y = j)$.

<u>3. Espérance et variance de combinaisons linéaires de variables aléatoires :</u>
Soient X et Y des variables aléatoires et $a \in \mathbb{R}$ et $b \in \mathbb{R}$;
a) Linéarité de l'espérance :
$E(aX + b) = aE(X) + b$;
$E(X + Y) = E(X) + E(Y)$.

b) Variance et écart-type :
$V(aX + b) = a^2V(X)$; $\sigma(aX + b) = |a|\sigma(X)$;
Si X et Y sont deux variables aléatoires indépendantes, alors $V(X + Y) = V(X) + V(Y)$.

<u>Démonstration :</u> a) Si X prend les valeurs x_1 , x_2 , x_3 , ... , x_m , alors $E(X) = \sum_{i=1}^{i=m} x_i\, p(X=x_i)$;
de plus, $aX + b$ prend les valeurs $ax_1 + b$, $ax_2 + b$, $ax_3 + b$, ... , $ax_m + b$,
donc $E(aX + b) = \sum_{i=1}^{i=m} (ax_i+b)\, p(aX+b=ax_i+b) = \sum_{i=1}^{i=m} (ax_i+b)\, p(X=x_i) = \sum_{i=1}^{i=m} [ax_i\, p(X=x_i)+b\, p(X=x_i)] =$
$a\sum_{i=1}^{i=m} x_i\, p(X=x_i) + b\sum_{i=1}^{i=m} p(X=x_i) = aE(X) + b$;

Si Y prend les valeurs y_1 , y_2 , y_3 , ... , y_s , alors $E(Y) = \sum_{i=1}^{i=s} y_i\, p(Y=y_i)$;
la variable aléatoire X + Y prend les valeurs $z_k = x_i + y_j$ avec i dans $\{1, 2, ... , m\}$ et j dans $\{1, 2, ... , s\}$;
$E(X + Y) = \sum_{k=1}^{k=r} z_k\, p(X+Y=z_k) = \sum_{x_i+y_j=z_k} (x_i+y_j)\, p(X=x_i \cap Y=y_j) =$
$\sum_{x_i+y_j=z_k} x_i\, p(X=x_i \cap Y=y_j) + \sum_{x_i+y_j=z_k} y_j\, p(X=x_i \cap Y=y_j) = \sum_{i=1}^{m} x_i \sum_{j=1}^{s} p(X=x_i \cap Y=y_j) + \sum_{j=1}^{s} y_j \sum_{i=1}^{m} p(X=x_i \cap Y=y_j)$
$= \sum_{i=1}^{m} x_i\, p(X=x_i) + \sum_{j=1}^{s} y_j\, p(Y=y_j) = E(X) + E(Y)$.

b) Si X prend les valeurs x_1 , x_2 , x_3 , ... , x_m , alors aX prend les valeurs ax_1 , ax_2 , ax_3 , ... , ax_m ;

$V(aX) = \sum_{i=1}^{i=m} (ax_i - E(aX))^2 p(aX = ax_i) = \sum_{i=1}^{i=m} (ax_i - aE(X))^2 p(X = x_i) = \sum_{i=1}^{i=m} a^2 (x_i - E(X))^2 p(X = x_i) =$

$a^2 \sum_{i=1}^{i=m} (x_i - E(X))^2 p(X = x_i) = a^2 V(X)$;

Si X prend les valeurs x_1 , x_2 , x_3 , ... , x_m , alors X + b prend les valeurs $x_1 + b$, $x_2 + b$, $x_3 + b$, ... , $x_m + b$;

$V(X + b) = \sum_{i=1}^{i=m} (x_i + b - E(X+b))^2 p(X + b = x_i + b) = \sum_{i=1}^{i=m} (x_i + b - (E(X) + b))^2 p(X = x_i) = \sum_{i=1}^{i=m} (x_i - E(X))^2 p(X = x_i)$

$= V(X)$;

Donc $V(aX + b) = a^2 V(X)$;

<u>Exemple :</u> 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 2 €, si on tombe sur « face », on perd 2 €.
- La 2ème partie consiste à lancer un dé à 6 faces. Si on tombe sur un chiffre pair, on gagne 4 €, si on tombe sur le « 3 » ou le « 5 », on perd 2 €, et si on tombe sur le « 1 », on perd 6 €.

La variable aléatoire X désigne les gains à la 1ère partie, la variable aléatoire Y désigne les gains à la 2ème 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 :

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

Y\X	2	− 2
4	6	2
− 2	0	− 4
− 6	− 4	− 8

Ainsi, on a :

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

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

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

x_i	− 8	− 4	0	2	6
$P(S = x_i)$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$

D'où $E(S) = E(X + Y) = -8 \times \frac{1}{12} - 4 \times \frac{1}{4} + 0 \times \frac{1}{6} + 2 \times \frac{1}{4} + 6 \times \frac{1}{4} = \frac{1}{3}$.

En fait, $E(X) = 0$ et $E(Y) = 4 \times \frac{1}{2} - 2 \times \frac{1}{3} - 6 \times \frac{1}{6} = \frac{1}{3}$; on trouve bien $E(S) = E(X) + E(Y)$.

$V(S) = \sum_{i=1}^{i=n} (x_i - E(S))^2 p(S = x_i) =$

$(-8 - \frac{1}{3})^2 \times \frac{1}{12} + (-4 - \frac{1}{3})^2 \times \frac{1}{4} + (0 - \frac{1}{3})^2 \times \frac{1}{6} + (2 - \frac{1}{3})^2 \times \frac{1}{4} + (6 - \frac{1}{3})^2 \times \frac{1}{4} = \frac{173}{9} \simeq 19{,}22$

et $\sigma(S) = \sqrt{\frac{173}{9}} \simeq 4{,}38$.

On a $V(X) = \sum_{i=1}^{i=n} (x_i - E(X))^2 p(X = x_i) = (-2 - 0)^2 \times \frac{1}{2} + (2 - 0)^2 \times \frac{1}{2} = 4$ et $\sigma(X) = 2$;

$V(Y) = \sum_{i=1}^{i=n} (y_i - E(Y))^2 p(Y = y_i) = (-6 - \frac{1}{3})^2 \times \frac{1}{6} + (-2 - \frac{1}{3})^2 \times \frac{1}{3} + (4 - \frac{1}{3})^2 \times \frac{1}{2} = \frac{137}{9}$ et $\sigma(Y) \simeq 3{,}9$;

On retrouve bien $V(S) = V(X) + V(Y) = 4 + \frac{137}{9} = \frac{36+137}{9} = \frac{173}{9}$.

4. Application à la loi binomiale :

Soit X une variable aléatoire qui suit la loi de Bernoulli de paramètre p. On répète n fois cette expérience.
Pour tout i de 1 à n, on note X_i la variable aléatoire associée pour la i-ième répétition ;
la variable aléatoire X_i suit la loi de Bernoulli de paramètre p.
On peut considérer alors que la liste $(X_1, X_2, X_3, \ldots, X_n)$ forment un échantillon de taille n de variables aléatoires suivant la loi de Bernoulli de paramètre n. Ces variables aléatoires sont indépendantes.
On pose : $S_n = X_1 + X_2 + X_3 + \ldots + X_n$, alors S_n suit la loi binomiale de paramètres n et p.
On a $E(S_n) = E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n) = np$,
et $V(S_n) = V(X_1 + X_2 + \ldots + X_n) = V(X_1) + V(X_2) + \ldots + V(X_n) = np(1 - p)$.

Exemple : On étudie la fiabilité d'un test virologique. On appelle X la variable aléatoire égale à 1 si le test est fiable et 0 dans le cas contraire. Le fabricant précise que le test est fiable dans 98 % des cas.
Dans ce cas, la variable aléatoire X suit la loi de Bernoulli de paramètre 0,98.

On effectue les tests sur un échantillon de 100 tests prélevés au hasard dans le stock du fabricant.
Pour tout i allant de 1 à n, on note X_i la variable aléatoire associée qui prend la valeur 1 si le test est fiable et 0 dans le cas contraire pour le i-ième test de l'échantillon ;
la variable aléatoire X_i suit la loi de Bernoulli de paramètre 0,98.
La liste $(X_1, X_2, X_3, \ldots, X_n)$ forment un échantillon de taille 100 de variables aléatoires suivant la loi de Bernoulli de paramètre 0,98.
On pose : $S_n = X_1 + X_2 + X_3 + \ldots + X_n$, alors S_n suit la loi binomiale de paramètres $n = 100$ et $p = 0{,}98$.

On a l'espérance mathématique de S_n est $E(S_n) = E(X_1) + E(X_2) + \ldots + E(X_n) = np = 98$,
et la variance $V(S_n) = V(X_1) + V(X_2) + \ldots + V(X_n) = np(1 - p) = 1{,}96$ et l'écart-type $\sigma(S_n) = \sqrt{1{,}96} = 1{,}4$.