

CHAPITRE 14 - SOMME DE VARIABLES ALÉATOIRES



Siméon Denis POISSON est né le 27 juin 1781 à Pithiviers. Son père y a été envoyé occuper un poste administratif peu important après sa carrière de soldat, ce qu'il a ressenti comme une grande frustration. Après la Révolution, il est nommé à la tête du gouvernement local, et souhaite alors aider son fils, qui est pétri de talents, à trouver une bonne situation. Ainsi, Siméon Denis est d'abord envoyé chez son oncle, chirurgien à Fontainebleau. Très vite, il apparaît cependant qu'il n'est pas fait pour la médecine. D'abord, ceci ne l'attire guère. Ensuite, il est très peu doué de ses mains, et les premiers patients qu'il opère décèdent quelques heures après l'intervention.

Après avoir renoncé à une carrière médicale, Poisson étudie à L'École Centrale de Fontainebleau, puis il réussit (en première position !) le concours d'entrée à L'École Polytechnique en 1798. Il y est un étudiant brillant, hormis en géométrie descriptive où son manque d'habileté manuelle l'empêche de réaliser des figures correctes. Dès 1800, il écrit deux mémoires importants, l'un sur la méthode d'élimination de BÉZOUT, l'autre sur les équations aux différences finies. Ils sont si bons que le second est publié dans le Recueil des savants étrangers, ce qui constitue un honneur exceptionnel pour un homme aussi jeune, et qu'il devient dès son diplôme obtenu répétiteur à L'École Polytechnique, sous la recommandation de LAPLACE. Il accédera rapidement au statut de professeur suppléant en 1802, puis complet en 1806, en remplacement de FOURIER que Napoléon a nommé Préfet à Grenoble.

Les travaux de Poisson sont nombreux (près de 400 publiés) et touchent surtout aux mathématiques appliquées. Il étudie le mouvement du pendule, les petites perturbations des mouvements planétaires. Pour les aspects plus orientés vers la physique, on lui doit les lois de l'électrostatique, et une définition de l'électricité comme un fluide où les éléments semblables se repoussent et les éléments contraires s'attirent (en 1812). Pour les aspects plus orientés vers les mathématiques, il est à l'origine de travaux sur les séries de Fourier qui préfigurent ceux de DIRICHLET ; il écrivit aussi en 1837 un important mémoire sur les probabilités, *Recherches sur la probabilité des jugements en matières criminelles et matière civile*, dans lequel apparaît la distribution qui porte désormais son nom (la distribution de Poisson décrit la probabilité qu'un événement ait lieu durant un intervalle de temps donné, pourvu que la probabilité de réalisation d'un événement est très faible, mais que le nombre d'essais est très grand). Ce traité ne fut pas tellement remarqué par les contemporains de Poisson, mais eut une grande influence par la suite.

I. Variables aléatoires $X + Y$ et aX

X et Y sont deux variables aléatoires définies sur l'univers d'une expérience aléatoire.

X prend les valeurs $x_1, x_2, \dots, x_n$ et Y prend les valeurs $y_1, y_2, \dots, y_m$ (avec n et m entiers naturels non nuls).

A. Variable aléatoire $X + Y$

Définition 1.

- La variable aléatoire $X + Y$ prend toutes les valeurs possibles $x_i + y_j$ avec $1 \leq i \leq n$ et $1 \leq j \leq m$.
- Loi de probabilité de $X + Y$: pour toute valeur w prise par $X + Y$, $P(X + Y = w)$ est la somme de toutes les probabilités $P(\{X = x_i\} \cap \{Y = y_j\})$ où $x_i + y_j = w$.

Exemple 1.

En période de réglage de ses machines, le contrôleur qualité d'une usine effectue de nombreux prélèvements dans la production et relève, pour chaque pièce, si elle a un défaut A , un défaut B . Compte tenu du grand nombre de pièces relevées, les fréquences indiquées dans ce tableau peuvent être assimilées à des probabilités.

X (resp. Y) est la variable aléatoire qui compte le nombre de défauts A (resp. B). Dans le tableau figurent en rouge les valeurs prises par la variable aléatoire somme $S = X + Y$.

Voici ci-dessous sa loi de probabilité.

s	0	1	2	3
$P(X = s)$	0,07	0,33	0,37	0,23

$Y \backslash X$	0	1	2	Loi de X
0	0,07	0,18	0,17	0,42
1	0,15	0,20	0,23	0,58
Loi de Y	0,22	0,38	0,40	1

B. Variable aléatoire aX

Définition 2.

a désigne un nombre réel différent de 0.

- La variable aléatoire aX prend toutes les valeurs possibles $a \times x_i$ avec $1 \leq i \leq n$.
- Loi de probabilité de aX : pour toute valeur w prise par aX , $P(aX = w)$ est la somme de toutes les probabilités $P(\{X = x_i\})$ où $a \times x_i = w$.

Exemple 2.

On lance un dé équilibré à six faces numérotées de 1 à 6. X est la variable aléatoire qui donne le numéro obtenu et Y celui qui donne le double de ce numéro. Alors $Y = 2X$.

C. Linéarité de l'espérance

Propriété 1 (admise).

- $E(X + Y) = E(X) + E(Y)$, et en particulier, $E(X + \text{constante}) = E(X) + \text{constante}$;
- $E(aX) = aE(X)$, avec a un nombre réel non nul.

Exemple 3.

On reprend l'exemple du paragraphe A..

$E(X) = 0,42 \times 0 + 0,58 \times 1 = 0,58$ et $E(Y) = 0,22 \times 0 + 0,38 \times 1 + 0,4 \times 2 = 1,18$.

Donc $E(X) + E(Y) = 1,76$.

$E(X + Y) = 0,07 \times 0 + 0,33 \times 1 + 0,37 \times 2 + 0,23 \times 3 = 1,76$.

On constate que $E(X + Y) = E(X) + E(Y)$.

II. Variables aléatoires indépendantes

A. Succession d'épreuves aléatoires indépendantes

On effectue successivement deux épreuves aléatoires indépendantes l'une de l'autre.

Ainsi, l'issue de la première épreuve n'influence pas l'issue de la seconde.

On note X (resp. Y) la variable aléatoire qui donne le résultat de la première (resp. seconde) épreuve.

On dit que les variables aléatoires X et Y sont **indépendantes**.

Conséquence :

Les événements $\{X = x_i\}$ et $\{Y = y_j\}$ sont indépendants, donc :

$$P(\{X = x_i\} \cap \{Y = y_j\}) = P(X = x_i) \times P(Y = y_j).$$

Exemple 4.

On lance successivement deux dés équilibrés, l'un à quatre faces numérotées de 1 à 4 et l'autre à six faces numérotées 0, 3, 3, 6, 6 et 6.

X (resp. Y) est la variable aléatoire qui donne le numéro obtenu avec le premier (resp. second) dé.

Les variables aléatoires X et Y sont indépendantes, car les deux lancers le sont.

Ainsi, par exemple, $P(\{X = 1\} \cap \{Y = 3\}) = P(X = 1) \times P(Y = 3) = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$.

B. Variance

Propriété 2 (admise).

X et Y sont deux variables aléatoires définies sur un même univers et a est un nombre réel différent de 0.

Si X et Y sont **indépendantes**, alors :

- $V(X + Y) = V(X) + V(Y)$, et en particulier, $V(X + \text{constante}) = V(X)$;
- $V(aX) = a^2 V(X)$.

Exemple 5.

X et Y sont deux variables aléatoires indépendantes telles que $V(X) = 1,25$ et $V(Y) = 5$.

Donc, $V(X + Y) = 1,25 + 5 = 6,25$ et $V(3X) = 3^2 \times 1,25 = 11,25$.

C. Application à la loi binomiale

Propriété 3.

X est une variable aléatoire qui suit la **loi binomiale** $\mathcal{B}(n ; p)$.

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

Démonstration.

On note X_i (avec $1 \leq i \leq n$) la variable aléatoire qui prend la valeur 1 en cas de succès à la i -ème épreuve et 0 sinon.

Chaque variable aléatoire X_i suit la loi de Bernoulli de paramètre p .

$$E(X_i) = (1 - p) \times 0 + p \times 1 = p \text{ et}$$

$$V(X_i) = (1 - p)(0 - p)^2 + p(1 - p)^2 = (1 - p)(p^2 + p(1 - p)) = p(1 - p).$$

Or, $X = X_1 + X_2 + \dots + X_n$ et $X_1, X_2, \dots, X_n$ sont indépendantes.

En généralisant les propriétés de l'espérance et de la variance d'une somme de deux variables au cas de n variables, on obtient :

- $E(X) = E(X_1) + E(X_2) + \dots + E(X_n) = nE(X_i) = np$;
- $V(X) = V(X_1) + V(X_2) + \dots + V(X_n) = nV(X_i) = np(1 - p)$.

On en déduit que $\sigma(X) = \sqrt{np(1 - p)}$.

k	0	1
$P(X_i = k)$	$1 - p$	p

III. Somme et moyenne d'un échantillon

A. Échantillon de taille n d'une loi de probabilité

Définition 3.

Un échantillon de taille n d'une loi de probabilité est une liste $(X_1 ; X_2 ; \dots ; X_n)$ de n variables aléatoires **indépendantes** et **identiques** qui suivent toutes cette loi.

Exemple 6.

Laetitia prend le même train cinq jours de la semaine. On admet que la variable aléatoire X qui compte le nombre de retards suit la loi binomiale $\mathcal{B}(5 ; 0,1)$. En répétant cette expérience pendant 8 semaines, on construit un échantillon $(X_1 ; X_2 ; X_3 ; X_4 ; X_5 ; X_6 ; X_7 ; X_8)$ de taille 8 de cette loi de probabilité.

B. Somme d'un échantillon

Définition 4.

On considère un échantillon $(X_1 ; X_2 ; \dots ; X_n)$ de taille n de la loi de probabilité suivie par une variable aléatoire X .

La **somme** de cet échantillon est la variable aléatoire $S_n = X_1 + X_2 + \dots + X_n$.

Propriété 4.

S_n est la somme d'un échantillon de taille n de la loi de probabilité suivie par une variable aléatoire X .

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

Démonstration.

- $E(S_n) = E(X_1) + \dots + E(X_n)$. Or les variables aléatoires $X_1, \dots, X_n$ suivent toutes la même loi que X , et donc $E(X_1) = E(X_2) = \dots = E(X_n) = E(X)$. On en déduit que $E(S_n) = nE(X)$.
- Les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, donc $V(S_n) = V(X_1) + \dots + V(X_n) = nV(X)$. On en déduit que $\sigma(S_n) = \sqrt{n}\sigma(X)$.

C. Moyenne d'un échantillon

Définition 5.

On considère un échantillon $(X_1 ; \dots ; X_n)$ de taille n de la loi de probabilité suivie par une variable aléatoire X . La **moyenne** de cet échantillon est la variable aléatoire $M_n = \frac{S_n}{n}$.

Propriété 5.

M_n est la moyenne d'un échantillon de taille n de la loi de probabilité suivie par une variable aléatoire X .

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

Démonstration.

On sait que pour tout réel $a \neq 0$, $E(aX) = aE(X)$ et $V(aX) = a^2V(X)$.

Donc $E(M_n) = E\left(\frac{1}{n}S_n\right) = \frac{1}{n}E(S_n) = \frac{1}{n} \times nE(X) = E(X)$.

Et $V(M_n) = V\left(\frac{1}{n}S_n\right) = \frac{1}{n^2}V(S_n) = \frac{1}{n^2} \times nV(X) = \frac{V(X)}{n}$.

On en déduit que $\sigma(M_n) = \frac{\sqrt{V(X)}}{\sqrt{n}} = \frac{\sigma(X)}{\sqrt{n}}$.