Chapitre XXV : Couples de variables aléatoires

Dans tout ce chapitre, on note $(\Omega, \mathbb{P})$ un espace probabilisé fini.

I Couples de variables aléatoires

I.1 Définition

Définition I.1

On appelle **couple de variables aléatoires réelles** sur $(\Omega, \mathbb{P})$ toute application définie par

$$Z \quad : \quad \begin{aligned} \Omega &\to \mathbb{R}^2 \\ \omega &\mapsto (X(\omega), Y(\omega)), \end{aligned}$$

où X et Y sont deux variables aléatoires réelles sur $(\Omega, \mathbb{P})$. On note alors $Z = (X, Y)$.

Remarque 1 :

- L'univers image de Z est le sous-ensemble de $\mathbb{R}^2$ donné par $Z(\Omega) = \left\{ (X(\omega), Y(\omega)) \in \mathbb{R}^2 \mid \omega \in \Omega \right\}$.
- Z est une variable aléatoire avec $E = \mathbb{R}^2$.

Exemple 2 : On lance deux dés équilibrés à 4 faces. On note X le plus petit résultat des deux dés et Y le plus grand résultat des deux dés. Déterminer l'univers image de Z.

Exemple 3 : Une urne contient 1 boule rouge, 2 boules vertes et 3 boules bleues. On tire simultanément trois boules et on note V le nombre de boules vertes obtenues et B le nombre de boules bleues obtenues. Déterminer l'univers image de $Z = (V, B)$.

Proposition I.2

Soit $Z = (X, Y)$ un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$. On note $X(\Omega) = \{x_1, \ldots, x_n\}$ et $Y(\Omega) = \{y_1, \ldots, y_p\}$. Alors, la famille

$$((X = x_i) \cap (Y = y_j))_{(i,j) \in [\![1;n]\!] \times [\![1;p]\!]}$$

forme un système complet d'évènements (incompatibles).

Démonstration. Posons pour tout $(i, j) \in [\![1;n]\!] \times [\![1;p]\!]$, $A_{i,j} = (X = x_i) \cap (Y = y_j)$.
Soient (i, j) et (k, l) deux couples d'indices de $[\![1;n]\!] \times [\![1;p]\!]$. On suppose $(i, j) \neq (k, l)$. Premier cas $i \neq k$. Alors $(X = x_i) \cap (X = x_k) = \emptyset$ (cf proposition II.7 chap22). Par suite

$$A_{i,j} \cap A_{k,l} = ((X = x_i) \cap (X = x_k)) \cap ((Y = y_j) \cap (Y = y_l)) = \emptyset.$$

On procède de même si $j \neq l$. Donc les évènements $(A_{i,j})_{(i,j) \in [\![1;n]\!] \times [\![1;p]\!]}$ sont disjoints et donc incompatibles.
De plus, on a

$$\begin{aligned} \bigsqcup_{(i,j) \in [\![1;n]\!] \times [\![1;p]\!]} A_{i,j} = \bigsqcup_{1 \leqslant i \leqslant n} \left(\bigsqcup_{1 \leqslant j \leqslant p} A_{i,j} \right) &= \bigsqcup_{1 \leqslant i \leqslant n} \left(\bigsqcup_{1 \leqslant j \leqslant p} ((X = x_i) \cap (Y = y_j)) \right) \\ &= \bigsqcup_{1 \leqslant i \leqslant n} \left((X = x_i) \cap \left(\bigsqcup_{1 \leqslant j \leqslant p} (Y = y_j) \right) \right) \\ &\underset{\text{prop II.7 chap22}}{=} \bigsqcup_{1 \leqslant i \leqslant n} ((X = x_i) \cap \Omega) \\ &= \bigsqcup_{1 \leqslant i \leqslant n} (X = x_i) \\ &= \Omega. \end{aligned}$$

On a donc bien démontré que la famille $(A_{i,j})_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]}$ forme un système complet d'évènements. □

Remarque 4 :

- Il est courant qu'il existe $(i,j)\in[\![1;n]\!]\times[\![1;p]\!]$, tel que $(X=x_i)\cap(Y=y_j)$ soit négligeable. Dans ce cas, il est possible de réduire l'univers image $Z(\Omega)$ qui sera alors distinct de $X(\Omega)\times Y(\Omega)$, voir l'exemple 2.
- En particulier, on a toujours
$$\sum_{i=1}^{n}\sum_{j=1}^{p}\mathbb{P}\left((X=x_i)\cap(Y=y_j)\right)=1.$$
- L'évènement $(X=x_i)\cap(Y=y_j)$ se note souvent $(X=x_i,Y=y_j)$.

Corollaire I.3

Soit $Z=(X,Y)$ un couple de variables aléatoires réelles sur $(\Omega,\mathbb{P})$. Pour tout $A\in X(\Omega)$ et tout $B\in Y(\Omega)$, on a
$$\mathbb{P}(Z\in A\times B)=\mathbb{P}\left((X\in A)\cap(Y\in B)\right)=\sum_{a\in A}\sum_{b\in B}\mathbb{P}\left((X=a)\cap(Y=b)\right).$$

I.2 Loi conjointe, loi marginale

Définition I.4

Soient $Z=(X,Y)$ un couple de variables aléatoires réelles sur $(\Omega,\mathbb{P})$. Alors la loi $\mathbb{P}_Z=\mathbb{P}_{(X,Y)}$ du couple donnée par
$$\begin{array}{rcl}\mathbb{P}_{(X,Y)} & : & Z(\Omega)\to[0;1]\\ & & (x,y)\mapsto\mathbb{P}\left((X=x)\cap(Y=y)\right)\end{array}$$
est appelée **la loi conjointe** de X et de Y.

Exemple 5 : On reprend l'exemple 2. La loi du couple est alors donnée par les probabilités suivantes :

$x_i\backslash y_i$	1	2	3	4
1	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$
2	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$
3	0	0	$\frac{1}{16}$	$\frac{2}{16}$
4	0	0	0	$\frac{1}{16}$

Définition I.5

Soit $Z=(X,Y)$ un couple de variables aléatoires réelles de $(\Omega,\mathbb{P})$. On appelle **lois marginales** de Z les lois des variables aléatoires X et Y.

Proposition I.6

Soit $Z=(X,Y)$ un couple de variables aléatoires réelles de $(\Omega,\mathbb{P})$. On suppose que $Z(\Omega)=\{x_1,\dots,x_n\}\times\{y_1,\dots,y_p\}$. Alors les lois marginales de Z sont données par
$$\forall i\in[\![1;n]\!],\quad \mathbb{P}(X=x_i)=\sum_{j=1}^{p}\mathbb{P}\left(Z=(x_i,y_j)\right)=\sum_{j=1}^{p}\mathbb{P}\left((X=x_i)\cap(Y=y_j)\right)$$
$$\forall j\in[\![1;p]\!],\quad \mathbb{P}(Y=y_j)=\sum_{i=1}^{n}\mathbb{P}\left(Z=(x_i,y_j)\right)=\sum_{i=1}^{n}\mathbb{P}\left((X=x_i)\cap(Y=y_j)\right).$$

Démonstration. Puisque $Z(\Omega) = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_p\}$, on en déduit que $Y(\Omega) = \{y_1, \dots, y_p\}$ (éventuellement certains évènements $(Y = y_j)$ sont négligeables). Donc par la proposition II.7 chap22, la famille $((Y = x_j))_{j \in [\![1;p]\!]}$ forme un système complet d'évènements (incompatibles). Donc par la formule des probabilités totales pour tout $i \in [\![1;n]\!]$,

$$\mathbb{P}(X = x_i) = \sum_{j=1}^{p} \mathbb{P}((X = x_i) \cap (Y = y_j)).$$

De même pour la loi marginale de Y.

$\square$

Remarque 6 : Si les probabilités de Z sont rangés dans un tableau, il suffit de sommer les lignes ou les colonnes pour obtenir la loi marginale de X ou de Y.

Exemple 7 : On reprend l'exemple 2. On obtient que

$$\sum_{j=1}^{4} \mathbb{P}((X = 2) \cap (Y = j)) = 0 + \frac{1}{16} + \frac{2}{16} + \frac{2}{16} = \frac{5}{16}.$$

Plus généralement, on a

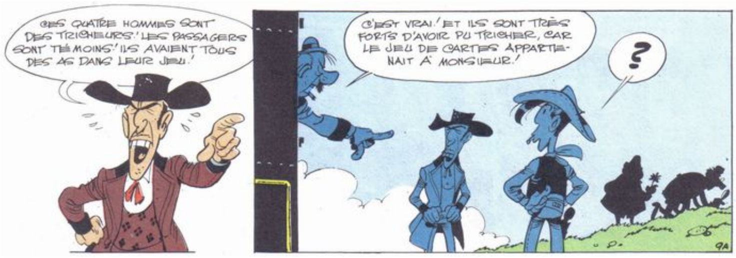
$x_i \backslash y_i$	1	2	3	4	$\mathbb{P}(X = x_i)$
1	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{7}{16}$
2	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{5}{16}$
3	0	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$
4	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$
$\mathbb{P}(Y = y_j)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$	1

Anti-Proposition I.7

Attention, la loi conjointe permet de déterminer les lois marginales mais la réciproque est fausse en générale et les lois marginales ne suffisent pas à déterminer entièrement la loi conjointe (qui contient l'information de comment se comporte X connaissant Y ou réciproquement).

Exemple 8 : On vérifie qu'une variable aléatoire $Z' = (X', Y')$ dont la loi est donnée par les probabilités ci-dessous est distincte de la loi de Z de l'exemple 2 et pourtant possède exactement les mêmes lois marginales.

Z'	1	2	3	4	$\mathbb{P}(X' = x_i)$
1	$\frac{1}{16}$	0	$\frac{4}{16}$	$\frac{2}{16}$	$\frac{7}{16}$
2	0	$\frac{3}{16}$	0	$\frac{2}{16}$	$\frac{5}{16}$
3	0	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$
4	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$
$\mathbb{P}(Y' = y_j)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$	1



I.3 Loi conditionnelle

Définition I.8

Soit (X, Y) un couple de variables aléatoires réelles.

- Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$, l'application
$$\begin{aligned} \mathbb{P}_{(Y=y)} \quad : \quad & X(\Omega) \to [0;1] \\ & x \mapsto \mathbb{P}(X = x \mid Y = y) = \frac{\mathbb{P}((X = x) \cap (Y = y))}{\mathbb{P}(Y = y)}, \end{aligned}$$
est appelée la **loi conditionnelle de X sachant** $(Y = y)$.
- Pour tout $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$, l'application
$$\begin{aligned} \mathbb{P}_{(X=x)} \quad : \quad & Y(\Omega) \to [0;1] \\ & y \mapsto \mathbb{P}(Y = y \mid X = x) = \frac{\mathbb{P}((X = x) \cap (Y = y))}{\mathbb{P}(X = x)}, \end{aligned}$$
est appelée la **loi conditionnelle de Y sachant** $(X = x)$.

Remarque 9 : Les lois conditionnelles définissent des probabilités sur $X(\Omega)$ et $Y(\Omega)$ respectivement.

Exemple 10 : On reprend l'exemple 2. Déterminer la loi conditionnelle de X sachant $(Y = 3)$.

Remarque 11 : En reprenant les notations de la proposition I.6, si $\forall i \in [\![1;n]\!]$, $\mathbb{P}(X = x_i) \neq 0$ et si $\forall j \in [\![1;p]\!]$, $\mathbb{P}(Y = y_j) \neq 0$, alors par la formule des probabilités totales, on a
$$\begin{aligned} \forall i \in [\![1;n]\!], \quad \mathbb{P}(X = x_i) &= \sum_{j=1}^{p} \mathbb{P}(X = x_i \mid Y = y_j)\, \mathbb{P}(Y = y_j) \\ \forall j \in [\![1;p]\!], \quad \mathbb{P}(Y = y_j) &= \sum_{i=1}^{n} \mathbb{P}(Y = y_j \mid X = x_i)\, \mathbb{P}(X = x_i). \end{aligned}$$

Définition I.9

Soit (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$. On dit que X et Y sont **indépendantes** si
$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x)\, \mathbb{P}(Y = y).$$

Remarque 12 :

1. Autrement dit, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,
$$\mathbb{P}_{(X,Y)}(\{(x, y)\}) = \mathbb{P}_X(\{x\})\, \mathbb{P}_Y(\{y\}).$$
2. Si X et Y sont indépendantes alors la loi conjointe s'obtient à l'aide des lois marginales (mais uniquement dans ce cas).
3. Si X et Y sont indépendantes alors les lois conditionnelles coïncident avec les lois marginales : pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$,
$$\forall x \in X(\Omega), \quad \mathbb{P}(X = x \mid Y = y) = \mathbb{P}(X = x)$$

et de même pour tout $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$, on a

$$\forall y \in Y(\Omega), \quad \mathbb{P}(Y = y \mid X = x) = \mathbb{P}(Y = y).$$

Exemple 13 : On considère une urne contenant n boules numérotées de 1 à n.

1. On tire successivement deux boules. On note X le numéro de la première boule et Y le numéro de la seconde boule. Les variables X et Y sont-elles indépendantes si le tirage est avec remise ?
2. Même question si le tirage est sans remise.

Proposition I.10

Soient (X, Y) un couple de variables aléatoires réelles indépendantes. Alors pour tout $(A, B) \in \mathscr{P}(X(\Omega)) \times \mathscr{P}(Y(\Omega))$, les évènements $(X \in A)$ et $(Y \in B)$ sont indépendants :

$$\mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B).$$

Démonstration. Par le corollaire I.3, on a

$$\begin{aligned}
\mathbb{P}((X \in A) \cap (Y \in B)) &= \sum_{a \in A} \sum_{b \in B} \mathbb{P}((X = a) \cap (Y = b)) \\
&= \sum_{a \in A} \sum_{b \in B} \mathbb{P}(X = a)\,\mathbb{P}(Y = b) && \textit{par indépendance} \\
&= \sum_{a \in A} \mathbb{P}(X = a) \sum_{b \in B} \mathbb{P}(Y = b) \\
&= \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B).
\end{aligned}$$

□

Proposition I.11

Soient (X, Y) un couple de variables aléatoires réelles indépendantes. Soient $f \in \mathscr{F}(\mathbb{R}, \mathbb{R})$ et $g \in \mathscr{F}(\mathbb{R}, \mathbb{R})$. Alors $f(X)$ et $g(Y)$ sont deux variables aléatoires indépendantes.

Démonstration. *Esquisse.* Soient $x \in f(X(\Omega))$ et $y \in g(Y(\Omega))$. On pose $A = f^{-1}(\{x\}) \cap X(\Omega)$ et $B = g^{-1}(\{y\}) \cap Y(\Omega)$ alors

$$\mathbb{P}((f(X) = x) \cap (g(Y) = y)) = \mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B) = \mathbb{P}(f(X) = x)\,\mathbb{P}(g(Y) = y).$$

□

I.4 Cas de la somme et vecteur aléatoire

Définition I.12

Soient $Z = (X, Y)$ un couple de variable aléatoires réelles sur $(\Omega, \mathbb{P})$ et $\varphi \in \mathscr{F}(\mathbb{R}^2, \mathbb{R})$ une application définie sur $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. Alors on définit $\varphi(Z)$ par

$$\varphi(Z) \quad : \quad \begin{aligned} &\Omega \to \mathbb{R} \\ &\omega \mapsto \varphi(Z)(\omega) = \varphi(Z(\omega)) = \varphi(X(\omega), Y(\omega)). \end{aligned}$$

Notamment si $\forall (x, y) \in \mathbb{R}^2$, $\varphi(x, y) = x + y$, on obtient $\varphi(Z) = X + Y$.

Proposition I.13 (loi de la somme)

Soit (X, Y) un couple de variable aléatoires réelles sur $(\Omega, \mathbb{P})$. La loi de $S = X + Y$ est donnée par

$$\forall s \in S(\Omega), \quad \mathbb{P}(S = s) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ s = x + y}} \mathbb{P}((X = x) \cap (Y = y)).$$

Exemple 14 : On lance deux dés équilibrés à 4 faces. On note X le résultat du premier dé et Y le résultat du second dé. Déterminer la loi de $S = X + Y$.

Définition I.14

Soient $X_1, \dots, X_n$, n variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

1. On définit alors **le vecteur aléatoire** $Z = (X_1, \dots, X_n)$ la variable aléatoire définie par
$$Z \quad : \quad \begin{aligned} \Omega &\to \mathbb{R}^n \\ \omega &\mapsto (X_1(\omega), \dots, X_n(\omega)) . \end{aligned}$$
2. On appelle loi conjointe de $(X_1, \dots, X_n)$ l'application
$$P_{X_1,\dots,X_n} \quad : \quad \begin{aligned} X_1(\Omega) \times \cdots \times X_n(\Omega) &\to [0; 1] \\ (x_1, \dots, x_n) &\mapsto \mathbb{P}((X_1 = x_i) \cap \cdots \cap (X_n = x_n)) . \end{aligned}$$
3. Les lois marginales de $(X_1, \dots, X_n)$ sont les lois de X_1, de X_2, ..., de X_n.

Définition I.15

Soient $X_1, \dots, X_n$, n variables aléatoires réelles sur $(\Omega, \mathbb{P})$. On dit que les variables aléatoires $X_1, \dots, X_n$ sont **mutuellement indépendantes** ou simplement **indépendantes** si et seulement si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$, les évènements $(X_1 = x_1), \dots, (X_n = x_n)$ sont indépendants :
$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega), \quad \mathbb{P}((X_1 = x_1) \cap \cdots \cap (X_n = x_n)) = \mathbb{P}(X_1 = x_1) \times \cdots \times \mathbb{P}(X_n = x_n) .$$

Proposition I.16 (Lemme des coalitions)

Soient $X_1, \dots, X_p, X_{p+1}, \dots, X_n$, n variables aléatoires réelles indépendantes sur $(\Omega, \mathbb{P})$. Alors

1. Les variables aléatoires $X_1, \dots, X_p$ sont indépendantes.
2. Si $f : \mathbb{R}^p \to \mathbb{R}$ et $g : \mathbb{R}^{n-p} \to \mathbb{R}$ alors les variables aléatoires $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.
3. Si $f_1 : \mathbb{R} \to \mathbb{R}$, ..., $f_n : \mathbb{R} \to \mathbb{R}$ alors les variables aléatoires $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes.

Exemple 15 : Si X, Y, Z, T sont des variables aléatoires réelles indépendantes alors

1. $X^2 + Y$ et ZT sont indépendantes.
2. $X, 2Y + \mathrm{e}^Z, T^3$ sont indépendantes.

II Espérance, Variance

II.1 Espérance

Définition II.1

Soit X une variable aléatoire réelle sur $(\Omega, \mathbb{P})$. On note $X(\Omega) = \{x_1, \dots, x_n\}$. On appelle **espérance mathématique** de X le réel
$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i) .$$

Interprétation. L'espérance de X est la moyenne sur les valeurs x_i prises par X pondérées par la fréquence d'apparition de ces valeurs $p_i = \mathbb{P}(X = x_i)$.

Exemple 16 : On note X le résultat d'un dé équilibré à six faces. Calculer $\mathbb{E}(X)$.

Proposition II.2

Soient (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$. Si X et Y ont la même loi alors, $\mathbb{E}(X) = \mathbb{E}(Y)$.

Démonstration. Si $\mathbb{P}_X = \mathbb{P}_Y$ alors $X(\Omega) = Y(\Omega)$ et pour tout $x \in X(\Omega)$, $\mathbb{P}(X = x) = \mathbb{P}_X(\{x\}) = \mathbb{P}_Y(\{x\}) = \mathbb{P}(Y = x)$. Par suite,

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) = \sum_{x \in Y(\Omega)} x\mathbb{P}(Y = x) = \mathbb{E}(Y).$$

□

Théorème II.3 (Théorème de Transfert)

Soit X une variable aléatoire réelle sur $(\Omega, \mathbb{P})$ et $\varphi \in \mathscr{F}(\mathbb{R}, \mathbb{R})$ une fonction réelle. Alors

$$\mathbb{E}(\varphi(X)) = \sum_{x \in X(\Omega)} \varphi(x)\,\mathbb{P}(X = x).$$

Remarque 17 :

- Une espérance peut-être vue comme l'intégrale. Pour rappel, si $\varphi \in \mathscr{E}([a;b])$, on avait alors $I_{[a;b]}(\varphi) = \sum_{i=1}^{n} \varphi(x_i)(x_{i+1} - x_i)$. Ici la largeur du rectangle est remplacée par $p_i = \mathbb{P}(X = x_i)$ le poids associé à x_i et la hauteur du rectangle reste la valeur de $\varphi(x_i)$.
- Cette proposition est intuitive mais c'est bien un théorème et non la définition de $\mathbb{E}(\varphi(X))$.

Démonstration. Soit $Y = \varphi(X)$. Alors par définition de l'intégrale,

$$\mathbb{E}(\varphi(X)) = \mathbb{E}(Y) = \sum_{y \in Y(\Omega)} y\mathbb{P}(Y = y) = \sum_{y \in Y(\Omega)} y\mathbb{P}(\varphi(X) = y).$$

Donc par la proposition III.5 chap22,

$$\begin{aligned}
\mathbb{E}(\varphi(X)) &= \sum_{y \in Y(\Omega)} y \sum_{\substack{x \in X(\Omega) \\ \varphi(x) = y}} \mathbb{P}(X = x) \\
&= \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ \varphi(x) = y}} y\mathbb{P}(X = x) \\
&= \sum_{x \in X(\Omega)} \sum_{\substack{y \in Y(\Omega) \\ \varphi(x) = y}} \varphi(x)\mathbb{P}(X = x) \\
&= \sum_{x \in X(\Omega)} \left(\varphi(x)\mathbb{P}(X = x) \sum_{\substack{y \in Y(\Omega) \\ \varphi(x) = y}} 1 \right).
\end{aligned}$$

Or à x fixé il existe un unique y tel que $\varphi(x) = y$. Donc $\sum_{y \in Y(\Omega), \varphi(x) = y} 1 = 1$. Conclusion,

$$\mathbb{E}(\varphi(X)) = \sum_{x \in X(\Omega)} \varphi(x)\mathbb{P}(X = x).$$

□

Exemple 18 : On reprend l'exemple 26 du chap22 : on lance un dé rouge et un dé vert tous les deux à 4 faces et équilibrés. On note X le résultat du dé rouge moins le résultat du dé vert et $Y = |X|$. Calculer l'espérance de Y.

Théorème II.4 (théorème de transfert pour deux variables aléatoires

Soit (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$ et $\varphi : \mathbb{R}^2 \to \mathbb{R}$. Alors

$$\mathbb{E}(\varphi(X, Y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \varphi(x, y)\,\mathbb{P}((X = x) \cap (Y = y)).$$

Remarque 19 : On peut étendre ce théorème à n variables aléatoires.

Proposition II.5

Soient (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

1. *Linéarité.* Pour tout $(a, b) \in \mathbb{R}^2$, on a $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
2. *Positivité.* Si pour tout $\omega \in \Omega$, $X(\omega) \geqslant 0$, autrement dit si $X(\Omega) \subseteq \mathbb{R}_+$, alors $\mathbb{E}(X) \geqslant 0$.
3. *Croissance.* Si pour tout $\omega \in \Omega$, $X(\omega) \geqslant Y(\omega)$ alors $\mathbb{E}(X) \geqslant \mathbb{E}(Y)$.
4. *Inégalité triangulaire.* On a $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$.

Démonstration.

1. Soit $(a, b) \in \mathbb{R}^2$. Par le théorème de transfert,

$$\mathbb{E}(aX + bY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (ax + by)\, \mathbb{P}((X = x) \cap (Y = y)).$$

Donc

$$\begin{aligned}
\mathbb{E}(aX + bY) &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (ax + by)\, \mathbb{P}((X = x) \cap (Y = y)) \\
&= \sum_{x \in X(\Omega)} \left(ax \sum_{y \in Y(\Omega)} \mathbb{P}((X = x) \cap (Y = y)) + \sum_{y \in Y(\Omega)} by\mathbb{P}((X = x) \cap (Y = y)) \right) \\
&\underset{\text{prop I.6}}{=} \sum_{x \in X(\Omega)} ax\mathbb{P}(X = x) + \sum_{y \in Y(\Omega)} by \sum_{x \in X(\Omega)} \mathbb{P}((X = x) \cap (Y = y)) \\
&\underset{\text{prop I.6}}{=} a \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) + b \sum_{y \in Y(\Omega)} y\mathbb{P}(Y = y) \\
&= a\mathbb{E}(X) + b\mathbb{E}(Y).
\end{aligned}$$

2. Si $X \geqslant 0$, alors pour tout $x \in X(\Omega)$, on a $x \geqslant 0$. Donc

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} \underbrace{x}_{\geqslant 0} \underbrace{\mathbb{P}(X = x)}_{\geqslant 0} \geqslant 0.$$

3. Si $X \geqslant Y$, alors $X - Y \geqslant 0$. Donc par le point 2, $\mathbb{E}(X - Y) \geqslant 0$. Ainsi, par le point 1, $\mathbb{E}(X) - \mathbb{E}(Y) \geqslant 0$ i.e. $\mathbb{E}(X) \geqslant \mathbb{E}(Y)$.
4. Par l'inégalité triangulaire sur une somme finie, on a d'une part,

$$|\mathbb{E}(X)| = \left| \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) \right| \leqslant \sum_{x \in X(\Omega)} |x\mathbb{P}(X = x)| = \sum_{x \in X(\Omega)} |x|\, \mathbb{P}(X = x).$$

D'autre part, par le théorème de transfert,

$$\sum_{x \in X(\Omega)} |x|\, \mathbb{P}(X = x) = \mathbb{E}(|X|).$$

Donc $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$.

□

Remarque 20 : Puisque l'espérance est une intégrale, il est normal de retrouver toutes les propriétés de l'intégrale, on peut également montrer un résultat de séparation, l'inégalité de Cauchy-Schwarz...

Remarque 21 :

1. Si Y est une variable constante égale à $c \in \mathbb{R}$, alors $\mathbb{E}(Y) = c$. En particulier,

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

2. Si $X_1, \dots, X_n$ sont n variables aléatoires alors

$$\mathbb{E}(X_1 + \dots + X_n) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_n).$$

Exemple 22 : On considère le jeu suivant. On mise 3 euros, on lance deux dés à quatre faces et on gagne alors autant d'euros que la moitié de la somme des deux dés. Quel est le gain moyen ? Le jeu est-il avantageux ?

II.2 Variance

Définition II.6

Soit X une variable aléatoire réelle $(\Omega, \mathbb{P})$. On définit **la variance** de X par

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right).$$

Remarque 23 :

- Si $X(\Omega) = \{x_1, \ldots, x_n\}$ alors, par le théorème de transfert,

$$\mathbb{V}(X) = \sum_{i=1}^{n} (x_i - \mathbb{E}(X))^2 \, \mathbb{P}(X = x_i) = \sum_{i=1}^{n} \left(x_i - \sum_{j=1}^{n} x_j \mathbb{P}(X = x_j) \right)^2 \mathbb{P}(X = x_i).$$

- Puisque $(X - \mathbb{E}(X))^2$ est une variable aléatoire positive alors par positivité de l'espérance, $\mathbb{V}(X) \geqslant 0$.
- Attention, à cause du carré, la variance n'est pas linéaire !

Interprétation. La variance est la moyenne pondérée des écarts de X à sa moyenne $\mathbb{E}(X)$ au carré. On dit que l'espérance est un indicateur de position tandis que la variance est un indicateur de dispersion. En particulier $\mathbb{V}(X) = 0$ si et seulement si $\mathbb{P}(X = \mathbb{E}(X)) = 1$.

Proposition II.7 (Formule de Koenig-Huygens)

Soit X une variable aléatoire réelle $(\Omega, \mathbb{P})$. Alors, on a

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2.$$

Démonstration. Par définition, on a

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2\right).$$

Par linéarité de l'espérance,

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - 2\mathbb{E}\left(X \underbrace{\mathbb{E}(X)}_{=\text{constante}}\right) + \mathbb{E}\left(\underbrace{\mathbb{E}(X)^2}_{=\text{constante}}\right) = \mathbb{E}\left(X^2\right) - 2\mathbb{E}(X)\,\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2.$$

□

Remarque 24 : Par le théorème de transfert, si $X(\Omega) = \{x_1, \ldots, x_n\}$ alors

$$\mathbb{V}(X) = \sum_{i=1}^{n} x_i^2 \mathbb{P}(X = x_i) - \left(\sum_{i=1}^{n} x_i \mathbb{P}(X = x_i) \right)^2.$$

Exemple 25 : Soit X le résultat d'un dé équilibré à six faces. Calculer la variance de X.

Exemple 26 : Soit S la somme de deux dés à 4 faces équilibrés. Déterminer la variance de S.

Définition II.8

- Soit X une variable aléatoire réelle sur $(\Omega, \mathbb{P})$. On définit **l'écart-type** de X par
$$\sigma = \sqrt{\mathbb{V}(X)}.$$
- Soit (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$. On définit **la covariance** de X et Y par
$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Proposition II.9

Soient (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$ et $(a, b) \in \mathbb{R}^2$.

1. $\mathrm{Cov}(X, X) = \mathbb{V}(X)$
2. $\mathbb{V}(aX + bY) = a^2\mathbb{V}(X) + 2ab\mathrm{Cov}(X, Y) + b^2\mathbb{V}(Y)$.
3. S'il existe $c \in \mathbb{R}$ telle que pour tout $\omega \in \Omega$, $X(\omega) = c$, alors $\mathbb{V}(X) = 0$.

Démonstration. Les deux premiers points sont laissés en exercice. Montrons le troisième point. Par la formule de Koenig-Huygens,
$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2 = \sum_{x \in X(\Omega)} x^2\mathbb{P}(X = x) - \left(\sum_{x \in X(\Omega)} x\mathbb{P}(X = x)\right)^2.$$
Or par hypothèse $X(\Omega) = \{c\}$ et $\mathbb{P}(X = c) = \mathbb{P}(\Omega) = 1$. Par conséquent,
$$\mathbb{V}(X) = c^2 - (c)^2 = 0.$$
□

Remarque 27 :

- En particulier la variance N'EST PAS linéaire.
- Si $Y = 1$ alors $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$.

Proposition II.10

Soit (X, Y) un couple de variables aléatoires réelles sur $(\Omega, \mathbb{P})$. Si X et Y sont indépendantes, alors

1. $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$,
2. $\mathrm{Cov}(X, Y) = 0$, on dit dans ce cas que X et Y sont décorrélées,
3. $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

Démonstration. Par le théorème de transfert,
$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xy\mathbb{P}((X = x) \cap (Y = y)) \\
&= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xy\mathbb{P}(X = x)\mathbb{P}(Y = y) \qquad \textit{par indépendance} \\
&= \sum_{x \in X(\Omega)} \left(x\mathbb{P}(X = x) \sum_{y \in Y(\Omega)} y\mathbb{P}(Y = y)\right) \\
&= \left(\sum_{x \in X(\Omega)} x\mathbb{P}(X = x)\right)\left(\sum_{y \in Y(\Omega)} y\mathbb{P}(Y = y)\right) \\
&= \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$
□

Remarque 28 :

1. ATTENTION la réciproque est fausse. Si X et Y sont décorrélés ou si $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ alors cela n'implique pas toujours que X et Y sont indépendants.
2. Si X et Y sont indépendantes alors $f(X)$ et $g(Y)$ sont indépendantes. Donc
$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)).$$

II.3 Fonction génératrice

Définition II.11

Soient $n \in \mathbb{N}$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$ à valeurs dans $[\![0;n]\!]$: $X(\Omega) \subseteq [\![0;n]\!]$. On appelle **fonction génératrice** de X la fonction G_X définie par
$$\forall t \in \mathbb{R}, \qquad G_X(t) = \mathbb{E}\left(t^X\right) = \sum_{k=0}^{n} t^k \mathbb{P}(X=k).$$

Proposition II.12

Soient X et Y deux variables aléatoires de même fonction génératrice : $G_X = G_Y$. Alors, $X \sim Y$.

Démonstration. Soient $n \in \mathbb{N}$ et X_1 et X_2 deux variables aléatoires à valeurs dans $[\![0;n]\!]$ de même fonction génératrice : $G_{X_1} = G_{X_2}$ (on change de notation pour parler ensuite de polynômes et ne pas confondre X et X). Alors,
$$\forall t \in \mathbb{N}, \qquad G_{X_1}(t) = \sum_{k=0}^{n} t^k \mathbb{P}(X_1=k) = \sum_{k=0}^{n} t^k \mathbb{P}(X_2=k) = G_{Y_1}(t).$$
Soient P et Q les polynômes noter $P = \sum_{k=0}^{n} \mathbb{P}(X_1=k) X^k$ et $Q = \sum_{k=0}^{n} \mathbb{P}(X_2=k) X^k$. Alors,
$$\forall t \in \mathbb{R}, \qquad P(t) = Q(t).$$
Les polynômes P et Q coïncidant sur une partie infinie sont nécessairement égaux. Donc
$$\sum_{k=0}^{n} \mathbb{P}(X_1=k) X^k = \sum_{k=0}^{n} \mathbb{P}(X_2=k) X^k.$$
Par unicité des coefficients d'un polynôme :
$$\forall k \in [\![0;n]\!], \qquad \mathbb{P}(X_1=k) = \mathbb{P}(X_2=k).$$
Conclusion, $X_1 \sim X_2$.

□

Proposition II.13

Soient $n \in \mathbb{N}$ et X une variable aléatoire à valeurs dans $[\![0;n]\!]$. Alors, G_X est $\mathscr{C}^\infty$ sur $\mathbb{R}$ et

1. $G_X(1) = 1$,
2. $\mathbb{E}(X) = G'_X(1)$,
3. $\mathbb{V}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$.

Démonstration.

1. On a
$$G_X(1) = \mathbb{E}\left(1^X\right) = \mathbb{E}(1) = 1.$$
2. De plus,
$$\forall t \in \mathbb{R}, \qquad G_X(t) = \sum_{k=0}^{n} t^k \mathbb{P}(X=k).$$

La fonction G_X est $\mathscr{C}^\infty$ sur $\mathbb{R}$ en tant que fonction polynomiale et

$$\forall t \in \mathbb{R}, \qquad G'_X(t) = \sum_{k=1}^{n} k t^{k-1} \mathbb{P}(X = k).$$

En particulier en $t = 1$,

$$G'_X(1) = \sum_{k=1}^{n} k \mathbb{P}(X = k) = \mathbb{E}(X).$$

3. De même,

$$G''_X(t) = \sum_{k=2}^{n} k(k-1) t^{k-2} \mathbb{P}(X = k).$$

Donc

$$G''_X(1) = \sum_{k=2}^{n} k(k-1) \mathbb{P}(X = k) = \sum_{k=0}^{n} k(k-1) \mathbb{P}(X = k) - 0.$$

Par le théorème de transfert,

$$\begin{aligned}
G''_X(1) = \mathbb{E}(X(X-1)) &= \mathbb{E}\left(X^2\right) - \mathbb{E}(X) && \text{par linéarité} \\
&= \underbrace{\mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2}_{=\mathbb{V}(X)} + \mathbb{E}(X)^2 - \mathbb{E}(X) \\
&= \mathbb{V}(X) + \mathbb{E}(X)^2 - \mathbb{E}(X) \\
&= \mathbb{V}(X) + \left(G'_X(1)\right)^2 - G'_X(1).
\end{aligned}$$

Conclusion,

$$\mathbb{V}(X) = G''_X(1) - \left(G'_X(1)\right)^2 + G'_X(1).$$

□

Proposition II.14

Soient X et Y deux variables aléatoires à valeurs dans un sous-ensemble de $\mathbb{N}$. Si X et Y sont indépendants, alors

$$G_{X+Y} = G_X \times G_Y.$$

Démonstration. Soient X et Y deux variables aléatoires à valeurs dans $[\![0;n]\!]$ et $[\![0;p]\!]$ respectivement. Supposons X et Y indépendantes et notons $Z = X + Y$. Alors, $Z(\Omega) = [\![0;n+p]\!]$. Soit $t \in \mathbb{R}$. Par définition,

$$G_{X+Y}(t) = \sum_{k=0}^{n+p} t^k \mathbb{P}(X + Y = k).$$

La famille $(X=i)_{i\in[\![0;n]\!]}$ forme un système complet d'évènements. Donc par la formule des probabilités totales :

$$\begin{aligned}
G_{X+Y}(t) &= \sum_{k=0}^{n+p} t^k \sum_{i=0}^{n} \mathbb{P}\left(X+Y=k \mid X=i\right)\mathbb{P}\left(X=i\right) \\
&= \sum_{k=0}^{n+p} t^k \sum_{i=0}^{n} \mathbb{P}\left(i+Y=k \mid X=i\right)\mathbb{P}\left(X=i\right) \\
&= \sum_{k=0}^{n+p} t^k \sum_{i=0}^{n} \mathbb{P}\left(Y=k-i \mid X=i\right)\mathbb{P}\left(X=i\right) \\
&= \sum_{k=0}^{n+p} t^k \sum_{i=0}^{n} \mathbb{P}\left(Y=k-i\right)\mathbb{P}\left(X=i\right) && \text{car } X \perp\!\!\!\perp Y \\
&= \sum_{i=0}^{n} \mathbb{P}\left(X=i\right) \sum_{k=0}^{n+p} t^k \mathbb{P}\left(Y=k-i\right) && \text{car la somme est triangulaire} \\
&= \sum_{i=0}^{n} \mathbb{P}\left(X=i\right) \sum_{k=i}^{p+i} t^k \mathbb{P}\left(Y=k-i\right) && \text{car } \mathbb{P}\left(Y=k-i\right)=0 \begin{cases} \text{si } k-i<0 \text{ i.e. } k<i \\ \text{ou} \\ \text{si } k-i>p \text{ i.e. } k>p+i \end{cases} \\
&= \sum_{i=0}^{n} \mathbb{P}\left(X=i\right) \sum_{j=0}^{p} t^{j+i} \mathbb{P}\left(Y=j\right) && \text{en posant, à } i \text{ fixé, } j=k-i \\
&= \left(\sum_{i=0}^{n} t^i \mathbb{P}\left(X=i\right)\right)\left(\sum_{j=0}^{p} t^j \mathbb{P}\left(Y=j\right)\right) && \text{car les variables sont séparées} \\
&= G_X(t)G_Y(t).
\end{aligned}$$

□

II.4 Inégalités probabilistes

Proposition II.15 (Inégalité de Markov)

Soit X une variable aléatoire réelle sur $(\Omega,\mathbb{P})$. On a alors pour tout $a>0$

$$\mathbb{P}\left(|X|\geqslant a\right) \leqslant \frac{\mathbb{E}\left(|X|\right)}{a}.$$

En particulier si X est **positive**, pour tout $a>0$

$$\mathbb{P}\left(X\geqslant a\right) \leqslant \frac{\mathbb{E}\left(X\right)}{a}.$$

Démonstration. Soit $a>0$. Notons $X_a=\{x\in X(\Omega) \mid |x|\geqslant a\}$. Alors, par la proposition III.5 du chap22, on a

$$a\mathbb{P}\left(|X|\geqslant a\right) = a\sum_{x\in X_a}\mathbb{P}\left(X=x\right) = \sum_{x\in X_a} a\mathbb{P}\left(X=x\right).$$

Or pour tout $x\in X_a$, $a\leqslant |x|$. Donc puisque $\mathbb{P}\left(X=x\right)$ est positif,

$$a\mathbb{P}\left(|X|\geqslant a\right) \leqslant \sum_{x\in X_a}|x|\,\mathbb{P}\left(X=x\right) = \sum_{x\in X(\Omega)}|x|\,\mathbb{P}\left(X=x\right) - \sum_{x\in X(\Omega)\setminus X_a}\underbrace{|x|\,\mathbb{P}\left(X=x\right)}_{\geqslant 0} \leqslant \sum_{x\in X(\Omega)}|x|\,\mathbb{P}\left(X=x\right).$$

Par le théorème de transfert, on conclut que

$$a\mathbb{P}\left(|X|\geqslant a\right) \leqslant \mathbb{E}\left(|X|\right).$$

□

Proposition II.16 (Inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle sur $(\Omega, \mathbb{P})$. On a alors pour tout $\varepsilon > 0$

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Démonstration. Par croissance de la fonction carré sur $\mathbb{R}_+$, on a $\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) = \mathbb{P}\left(|X - \mathbb{E}(X)|^2 \geqslant \varepsilon^2\right)$. On applique alors l'inégalité de Markov à la variable $Y = (X - \mathbb{E}(X))^2$ et au réel $a = \varepsilon^2$. On obtient bien que

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{E}\left((X - \mathbb{E}(X))^2\right)}{\varepsilon^2} = \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

□

Exemple 29 : Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires, toutes de même loi et indépendantes. Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$. Montrer que

$$\lim_{n\to+\infty} \mathbb{P}\left(\left|\overline{X_n} - \mathbb{E}(X_1)\right| \geqslant \varepsilon\right) = 0.$$

Ce résultat est appelé la loi faible des grands nombres. On dit que $\left(\overline{X_n}\right)_{n\in\mathbb{N}^*}$ converge *en probabilité* vers $\mathbb{E}(X_1)$.

III Complément sur les lois usuelles

III.1 Loi déterministe

Définition III.1

On dit qu'une variable aléatoire X sur $(\Omega, \mathbb{P})$ suit une **loi déterministe** ou certaine s'il existe $c \in \mathbb{R}$ tel que

$$X(\Omega) = \{c\} \qquad \text{et} \qquad \mathbb{P}(X = c) = 1.$$

Proposition III.2

Si X est déterministe, égale à c alors $\mathbb{E}(X) = c$, $\mathbb{V}(X) = 0$ et, si $c \in \mathbb{N}$, $\forall t \in \mathbb{R}$, $G_X(t) = t^c$.

III.2 Loi uniforme

Définition III.3

Soient $n \in \mathbb{N}^*$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$. On dit que X suit une **loi uniforme** sur $[\![1;n]\!]$ si et seulement si

$$X(\Omega) = [\![1;n]\!] \qquad \text{et} \qquad \forall k \in [\![1;n]\!],\quad \mathbb{P}(X = k) = \frac{1}{n}.$$

On note alors $X \sim \mathscr{U}([\![1;n]\!])$.

Interprétation. Cela correspond au résultat d'une expérience équiprobable contenant n issues.

Proposition III.4

Soient $n \in \mathbb{N}^*$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$ de loi uniforme sur $[\![1;n]\!]$, $X \sim \mathscr{U}([\![1;n]\!])$. Alors

$$\mathbb{E}(X) = \frac{n+1}{2}, \qquad \mathbb{V}(X) = \frac{(n-1)(n+1)}{12} \qquad \text{et} \qquad \forall t \in \mathbb{R},\ G_X(t) = \frac{1}{n}\sum_{k=1}^{n} t^k.$$

Démonstration. EXO !

□

Remarque 30 : En particulier, pour tout $t \in \mathbb{R} \setminus \{1\}$, $G_X(t) = \frac{t}{n}\frac{1-t^n}{1-t}$.

III.3 Loi de Bernoulli

Définition III.5

Soient $p \in [0;1]$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$. On dit que X suit une **loi de Bernoulli** de paramètre p si et seulement si

$$X(\Omega) = \{0;1\} \qquad \text{et} \qquad \mathbb{P}(X=0) = 1-p, \qquad \mathbb{P}(X=1) = p.$$

On note alors $X \sim \mathscr{B}(p)$ ou parfois $X \sim b(p)$.

Interprétation. Cela correspond au résultat d'une expérience à deux issues, dont la probabilité du succès est donnée par p et la probabilité de l'échec par $1-p$.

Proposition III.6

Soient $p \in [0;1]$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$ de loi de Bernoulli de paramètre p, $X \sim \mathscr{B}(p)$. Alors

$$\mathbb{E}(X) = p, \qquad , \mathbb{V}(X) = p(1-p) \qquad \text{et} \qquad \forall t \in \mathbb{R},\ G_X(t) = 1-p+pt.$$

Démonstration. EXO ! □

III.4 Loi binomiale

Définition III.7

Soient $n \in \mathbb{N}^*$, $p \in [0;1]$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$. On dit que X suit une **loi binomiale** de paramètre n et p si et seulement si

$$X(\Omega) = [\![0;n]\!] \qquad \text{et} \qquad \forall k \in [\![0;n]\!],\quad \mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note alors $X \sim \mathscr{B}(n,p)$.

Interprétation. Cela correspond au nombre de succès obtenus après n réalisations indépendantes d'une même expérience de Bernoulli de paramètre p.

Proposition III.8

Soient $n, m \in \mathbb{N}^*$, $p \in [0;1]$, $X \sim \mathscr{B}(n,p)$ et $Y \sim \mathscr{B}(m,p)$. On suppose que X et Y sont **indépendantes**. Alors

$$X + Y \sim \mathscr{B}(n+m, p).$$

Démonstration. Commençons par déterminer la fonction génératrice d'une binomiale. Soit $X \sim \mathscr{B}(n,p)$. Alors,

$$\forall t \in \mathbb{R}, \qquad G_X(t) = \sum_{k=0}^{n} t^k \mathbb{P}(X=k) = \sum_{k=0}^{n} t^k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pt)^k (1-p)^{n-k}.$$

On reconnaît un binôme de Newton. Donc

$$\forall t \in \mathbb{R}, \qquad G_X(t) = (1-p+pt)^n.$$

Soient maintenant $X \sim \mathscr{B}(n,p)$ et $Y \sim \mathscr{B}(m,p)$ indépendantes. Puisque X et Y sont indépendantes, par la proposition II.14

$$\forall t \in \mathbb{R}, \qquad G_{X+Y}(t) = G_X(t) G_Y(t).$$

Or on a vu que $G_X(t) = (1-p+pt)^n$ et de même $G_Y(t) = (1-p+pt)^m$. Ainsi,

$$\forall t \in \mathbb{R}, \qquad G_{X+Y}(t) = (1-p+pt)^n \times (1-p+pt)^m = (1-p+pt)^{n+m}$$

Donc G_{X+Y} est la fonction génératrice d'une loi binomiale de paramètre $n+m$ et p. Or par la proposition II.12 la fonction génératrice caractérise la loi. Conclusion,

$$X + Y \sim \mathscr{B}(n+m, p).$$

Version 2, pour les plus curieux. Puisque $X(\Omega) = [\![0;n]\!]$ et $Y(\Omega) = [\![0;m]\!]$, on constate que $(X+Y)(\Omega) = [\![0;n+m]\!]$. De plus pour tout $s \in [\![0;m+n]\!]$, on a par la formule des probabilités totales,

$$\begin{aligned}\mathbb{P}(X+Y=s) &= \sum_{k=1}^{m} \mathbb{P}(X+Y=s \mid Y=k)\,\mathbb{P}(Y=k)\\ &= \sum_{k=1}^{m} \mathbb{P}(X+k=s \mid Y=k)\,\mathbb{P}(Y=k)\\ &= \sum_{k=1}^{m} \mathbb{P}(X=s-k \mid Y=k)\,\mathbb{P}(Y=k)\,.\end{aligned}$$

Or X et Y sont indépendantes donc pour tout $k \in [\![1;m]\!]$, $\mathbb{P}(X=s-k \mid Y=k) = \mathbb{P}(X=s-k)$. Puisque $X(\Omega) = [\![1;m]\!]$, on a $\mathbb{P}(X=s-k)=0$ si $k>s$. Donc

$$\mathbb{P}(X+Y=s) = \sum_{k=1}^{s} \mathbb{P}(X=s-k)\,\mathbb{P}(Y=k) = \sum_{k=1}^{s} \binom{n}{s-k} p^{s-k}(1-p)^{n-s+k}\binom{m}{k}p^k(1-p)^{m-k}.$$

Ainsi,

$$\mathbb{P}(X+Y=s) = p^s(1-p)^{n+m-s}\sum_{k=1}^{s}\binom{n}{s-k}\binom{m}{k}.$$

Par la formule de Vandermonde (cf exo25 TD22), on conclut que

$$\forall s \in [\![0;m+n]\!], \quad \mathbb{P}(X+Y=s) = \binom{n+m}{s}p^s(1-p)^{n+m-s},$$

i.e. $X+Y \sim \mathscr{B}(n+m,p)$.

□

Proposition III.9

Soient $n \in \mathbb{N}^*$, $p \in [0;1]$ et $X_1,\dots,X_n$, n variables aléatoires **indépendantes** et de même loi, la loi de Bernoulli de paramètre p :

$$\forall i \in [\![1;n]\!], \qquad X_i \sim \mathscr{B}(p)\,.$$

Alors

$$X_1+X_2+\cdots+X_n \sim \mathscr{B}(n,p)\,.$$

Démonstration. Pour tout $k \in [\![1;n]\!]$, on pose $S_k = \sum_{i=1}^{k} X_i$. Montrons par récurrence que pour tout $k \in [\![1;n]\!]$ la proposition suivante est vraie :

$$\mathscr{P}(k) \qquad : \qquad \text{« } S_k \sim \mathscr{B}(k,p). \text{ »}$$

Initialisation. Si $k=1$ alors $S_1 = X_1 \sim \mathscr{B}(p) = \mathscr{B}(1,p)$. Donc $\mathscr{P}(1)$ est vrai.
Hérédité. Soit $k \in [\![1;n-1]\!]$. Supposons que $\mathscr{P}(k)$ est vrai i.e. $S_k \sim \mathscr{B}(k,p)$. La variable aléatoire S_k est une fonction des variables aléatoires $X_1,\dots,X_k$ qui sont indépendantes de la variable X_{k+1}. Donc par la proposition I.16, S_k et X_{k+1} sont indépendantes. Or $S_k \sim \mathscr{B}(k,p)$ et $X_{k+1} \sim \mathscr{B}(1,p)$. Donc par la proposition précédente,

$$S_{k+1} = S_k + X_{k+1} \sim \mathscr{B}(k+1,p)\,.$$

Donc $\mathscr{P}(k+1)$ est vraie.
Conclusion. Pour tout $k \in [\![1;n]\!]$, $\mathscr{P}(k)$ est vraie et notamment $\mathscr{P}(n)$, donc $S_n \sim \mathscr{B}(n,p)$.

□

Proposition III.10

Soient $n \in \mathbb{N}^*$, $p \in [0;1]$ et X une variable aléatoire sur $(\Omega,\mathbb{P})$ de loi binomiale de paramètre n et p, $X \sim \mathscr{B}(n,p)$. Alors

$$\mathbb{E}(X) = np, \qquad \mathbb{V}(X) = np(1-p) \qquad \text{et} \qquad \forall t \in \mathbb{R},\ G_X(t) = (1-p+pt)^n\,.$$

Démonstration. Version 1. Soient $X_1, \ldots, X_n$, n variables aléatoires indépendantes de loi de Bernoulli de paramètre p. Alors $X_1 + \cdots + X_n$ a même loi que X. Donc

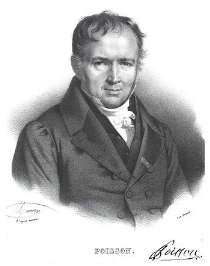
$$\begin{aligned}\mathbb{E}(X) = \mathbb{E}(X_1 + \cdots + X_n) &= \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) && \textit{par linéarité de l'espérance}\\ &= p + \cdots + p && \textit{par la proposition III.6}\\ &= np.\end{aligned}$$

De même,

$$\begin{aligned}\mathbb{V}(X) = \mathbb{V}(X_1 + \cdots + X_n) &= \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n) && \textit{par indépendance des } X_i \textit{ et non la linéarité !}\\ &= p(1-p) + \cdots + p(1-p) && \textit{par la proposition III.6}\\ &= np(1-p).\end{aligned}$$

Pour la fonction génératrice, voir la démonstration précédente.
Version 2. On a déjà vu que
$$\forall t \in \mathbb{R}, \qquad G_X(t) = (1 - p + pt)^n.$$
Donc G_X est deux fois dérivables sur $\mathbb{R}$ et pour $n \geqslant 2$,
$$G_X'(t) = np(1-p+pt)^{n-1} \qquad \text{et} \qquad G_X''(1) = n(n-1)p^2(1-p+pt)^{n-2}$$
Donc par la question II.13 on a
$$\mathbb{E}(X) = G_X'(1) = np(1-p+p)^{n-1} = np$$
et
$$\mathbb{V}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2 = n(n-1)p^2 + np - (np)^2 = np - np^2 = np(1-p).$$

□



Denis POISSON *(Pithiviers (Loiret) 1781 - Sceaux 1840) est un physicien et mathématicien français. A la mort de son père, un soldat ayant pris sa retraite dans l'administration, sa mère sans ressource envoie alors Poisson âgé de 15 ans chez un oncle chirurgien étudier la médecine. Après avoir causé la mort de quelques patients, il se tourne vers les mathématiques dans lesquelles il excelle. A dix-huit ans il est reçu premier à l'Ecole Polytechnique et obtient très rapidement une très grande reconnaissance. Il est très rapidement soutenu par Laplace et Lagrange et recevra de nombreuses récompenses au cours de sa carrière. Il devient membre de l'Académie des sciences en 1812, est fait baron par Louis XVIII et obtient un siège à la chambre des pairs en 1832. Travaillant d'abord sur la géométrie et la mécanique, son oeuvre concerne surtout la physique mathématique. Il pose les lois de l'électrosatique et travaille à décrire l'électricité. En mathématiques, on lui doit des résultats sur l'intégration et les séries de Fourier. En 1837 il publie un ouvrage de probabilités et décrit pour la première fois la loi qui porte son nom s'interprétant comme une loi limite d'une loi binomiale* $\mathscr{B}(n,p)$ *où* np *est constant. Il est l'un des premiers à maitriser toutes les profondes nouvelles notions d'analyse et à comprendre leur grande portée dans des domaines aussi variés qu'en mécanique, électricité ou probabilité.*

Alors âgé de 17 ans, Poisson se trompa dans l'horaire de cours d'histoire naturelle et déboucha dans un cours de mathématiques. Le professeur ayant tellement peu d'auditeur, qu'il interdit à Poisson de ressortir. Immédiatement passionné par le cours, Poisson continua par la suite à suivre ce cours, abandonna la médecine et se lança dans des études de mathématiques !

La nourrice du jeune Poisson accrochait le bébé en hauteur par les bretelles à un crochet du mur lorsqu'elle s'absentait : « C'est pour lui éviter la saleté du sol » expliqua-t-elle au père arrivé de manière impromptu. Poisson prétendit plus tard que ce fut ses efforts pour s'en délivrer qui l'avaient conduit à l'étude du mouvement du pendule...

Loi de Poisson

Soient $\lambda \in \mathbb{R}_+^*$ et X une variable aléatoire sur $(\Omega, \mathbb{P})$. On dit que X suit une **loi de Poisson** de paramètre λ si et seulement si

$$X(\Omega) = \mathbb{N} \qquad \text{et} \qquad \forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

On note alors $X \sim \mathscr{P}(\lambda)$.

« La vie n'est bonne qu'à deux choses : à faire des mathématiques et à les professer ».

Denis Poisson

Un étudiant n'ayant pas bien appris son cours se retrouve embêté lorsque le colleur lui demande :
« - Les variables X *et* Y *sont-elles indépendantes ?*
Audacieux, il tente malgré tout une réponse :
- X *oui, clairement ! Mais* Y *je pense que non. »*

N'ayant été pour autant pas détourné de sa noble vocation, il devient quelques années plus tard mathématicien. Lors d'un séminaire, il propose un nouveau théorème qu'il présente et explique consciencieusement à son auditoire. Il détaille pendant une heure sa preuve au tableau. Alors qu'il est pratiquement arrivé à la fin, son ancien professeur, présent au fond de la salle se lève et affirme :
« - Ce théorème est faux ! J'en ai un contre-exemple.
Notre héros, ne se démontant pas pour autant, efface tout son tableau et affirme avec enthousiasme :
- Aucune importance, j'ai une autre démonstration de ce résultat ! »