Variables aléatoires sur un univers fini

Plan du chapitre

1 Variables aléatoires sur un univers fini

1.1 Définition d'une variable aléatoire

DÉFINITION 1. Soient Ω un univers fini.

Une **variable aléatoire** X sur Ω est une application de Ω dans un certain ensemble non vide E.
Si E est une partie de $\mathbb{R}$, la variable aléatoire est dite **réelle**.

⇨ **Commentaire** .

◇ *On notera qu'une « variable aléatoire » n'est ni une variable (puisque c'est une application), ni aléatoire (aucune probabilité n'apparaît dans la définition d'une variable aléatoire).*

◇ *Le programme officiel prévoit qu'une variable aléatoire peut être à valeurs dans un ensemble quelconque. Mais, dans la pratique de maths sup et de maths spé, une variable aléatoire X sera quasi systématiquement à valeurs dans $\mathbb{R}$, ce qui permettra par exemple d'en calculer l'espérance et la variance.*

Exemples.

• On lance deux dés à six faces successivement. Un univers associé à cette expérience est $\Omega = [\![1,6]\!]^2$. L'application

$$X : \begin{array}{ccc} [\![1,6]\!]^2 & \rightarrow & [\![2,12]\!] \\ \omega = (x,y) & \mapsto & X(\omega) = x+y \end{array}$$

est une variable aléatoire réelle sur Ω : c'est l'application qui, à chaque lancer des deux dés, associe la somme des points obtenus.

• On choisit un élève d'une classe. Un univers associé à cette expérience est l'ensemble Ω des élèves de la classe. L'application qui à chaque élève associe son sexe, M ou F, est une variable aléatoire définie sur Ω à valeurs dans $\{M, F\}$. Ce n'est pas une variable aléatoire réelle.

• Soient Ω un univers fini puis A un événement donné (c'est-à-dire un élément de $\mathcal{P}(\Omega)$). La **variable aléatoire indicatrice** 1_A de l'événement A est l'application de Ω dans $\{0,1\}$ définie par :

$$\forall \omega \in \Omega,\ 1_A(\omega) = \left\{\begin{array}{l} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{array}\right. .$$

On a alors $A = \{\omega \in \Omega /\ 1_A(\omega) = 1\} = 1_A^{-1}(\{1\})$ et $\overline{A} = \{\omega \in \Omega /\ 1_A(\omega) = 0\} = 1_A^{-1}(\{0\})$.

❏

On doit maintenant définir certaines notations et se les approprier. Soient Ω un univers fini puis X une variable aléatoire sur Ω à valeurs dans un ensemble E. Soit A une partie de E (A n'est donc pas un événement mais est une partie de l'ensemble d'arrivée de X). On sait que la notation $X^{-1}(A)$ désigne l'image réciproque de la partie A par l'application X ou encore l'ensemble des antécédents des éléments de A par X :

$$X^{-1}(A) = \{\omega \in \Omega /\ X(\omega) \in A\}.$$

$X^{-1}(A)$ est une partie de Ω ou encore un élément de $\mathcal{P}(\Omega)$. $X^{-1}(A)$ **est donc un événement**. Il se note $\{X \in A\}$ ou aussi $(X \in A)$.

$$X^{-1}(A) = \{X \in A\} = \{\omega \in \Omega /\ X(\omega) \in A\}.$$

On suppose de plus que X est à valeurs dans $\mathbb{R}$. Si A est un singleton $\{x\}$ où x est un réel, l'événement $X^{-1}(A)$ se note $\{X = x\}$ ou aussi $(X = x)$:

$$\begin{array}{l}
\forall x \in \mathbb{R},\ X^{-1}(\{x\}) = \{X = x\} = \{\omega \in \Omega /\ X(\omega) = x\}. \\
\forall x \in \mathbb{R},\ X^{-1}(]-\infty, x]) = \{X \leqslant x\} = \{\omega \in \Omega /\ X(\omega) \leqslant x\}. \\
\forall x \in \mathbb{R},\ X^{-1}(]-\infty, x[) = \{X < x\} = \{\omega \in \Omega /\ X(\omega) < x\}. \\
\forall x \in \mathbb{R},\ X^{-1}([x, +\infty[) = \{X \geqslant x\} = \{\omega \in \Omega /\ X(\omega) \geqslant x\}. \\
\forall x \in \mathbb{R},\ X^{-1}(]x, +\infty[) = \{X > x\} = \{\omega \in \Omega /\ X(\omega) > x\}.
\end{array}$$

Si A est l'un des ensembles $]-\infty, x]$ ou $]-\infty, x[$ ou $[x, +\infty[$ ou $]x, +\infty[$, l'événement $X^{-1}(A)$ se note $\{X \leqslant x\}$ ou $\{X < x\}$ ou $\{X \geqslant x\}$ ou $\{X > x\}$:

$$\forall x \in \mathbb{R},\ X^{-1}(]-\infty, x]) = \{X \leqslant x\} = \{\omega \in \Omega /\ X(\omega) \leqslant x\}.$$

Reprenons l'exemple d'un lancer de deux dés, avec $\Omega = [\![1,6]\!]^2$, puis prenons pour X la variable aléatoire qui à chaque lancer (modélisé par un couple $(x,y) \in [\![1,6]\!]^2$) associe la somme $x+y$ des points obtenus. Alors

$$\{X = 2\} = \{(1,1)\},\ \{X = 4\} = \{(1,3),(2,2),(3,1)\},\ \{X \leqslant 3\} = \{(1,1),(1,2),(2,1)\},\ \{X \geqslant 2\} = [\![1,6]\!]^2 \text{ et } \{X > 12\} = \varnothing.$$

1.2 Loi de probabilité d'une variable aléatoire

On munit maintenant l'univers Ω d'une probabilité P : soit (Ω, P) un espace probabilisé fini. On considère une variable aléatoire X sur Ω à valeurs dans un certain ensemble E.

Définition 2. La **loi de probabilité** P_X de la variable aléatoire X est l'application

$$P_X : \begin{array}{ccc} \mathcal{P}(X(\Omega)) & \to & [0,1] \\ A & \mapsto & P(\{X \in A\}) \end{array}.$$

Ainsi,

$$\forall A \in \mathcal{P}(X(\Omega)),\ P_X(A) = P(\{X \in A\}) = P\left(X^{-1}(A)\right) = P\left(\{\omega \in \Omega / X(\omega) \in A\}\right).$$

Théorème 1. Soit (Ω, P) un espace probabilisé fini. Soit X une variable aléatoire sur Ω à valeurs dans un certain ensemble E.

$$\forall A \in \mathcal{P}(X(\Omega)),\ P_X(A) = \sum_{x \in A} P(\{X = x\}),$$

(avec la convention usuelle qu'une somme vide est nulle dans le cas où A est vide).

Démonstration. Si A est vide, $P_X(A) = P\left(X^{-1}(\varnothing)\right) = P(\varnothing) = 0 = \sum_{x \in A} P(\{X = x\})$.

Si A n'est pas vide, les événements $\{X = x\}$, $x \in A$, constituent une partition de l'événement $\{X \in A\}$ et donc,

$$P_X(A) = P(\{X \in A\}) = P\left(\bigcup_{x \in A} \{X = x\}\right) = \sum_{x \in A} P(\{X = x\}).$$

❑

Ainsi, les probabilités $P_X(A) = P(\{X \in A\})$ où A est une partie de $X(\Omega)$, sont entièrement déterminés par les nombres $P(\{X = x\})$ où $x \in X(\Omega)$. Pour cette raison, dans la pratique, quand on demande la loi de probabilité d'une variable aléatoire X, on se contente de calculer les $P(\{X = x\})$ où $x \in X(\Omega)$.

Exemples.

• Si on reprend un lancer de deux dés avec $\Omega = [\![1, 6]\!]^2$, que l'on munit Ω de la probabilité uniforme (cas de dés non truqués) et que X désigne la somme des points obtenus (de sorte que $X(\Omega) = [\![2, 12]\!]$), la loi de probabilité de X est donnée par :

x	2	3	4	5	6	7	8	9	10	11	12
$P(\{X = x\})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

• Soit (Ω, P) un espace probabilisé fini. Soit A un élément de $\mathcal{P}(\Omega)$ c'est-à-dire un événement. Considérons $X = 1_A$ la variable indicatrice de l'événement A. Alors,

$$P(A) = P_{1_A}(\{1\}) = P\left(\{1_A = 1\}\right) \quad \text{et} \quad P\left(\overline{A}\right) = P_{1_A}(\{0\}) = P\left(\{1_A = 0\}\right).$$

1.3 Fonction de répartition d'une variable aléatoire réelle

Définition 3. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

La **fonction de répartition** associée à la variable X est la fonction F définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R},\ F(x) = P(X \leqslant x).$$

Exemple. On reprend l'exemple du lancer successif de deux dés équilibrés, X désignant la somme des points obtenus. La fonction de répartition F associée à cette variable est définie pour tout réel x par :

- si $x < 2$, $F(x) = P(X \leqslant x) = 0$.
- si $2 \leqslant x < 3$, $F(x) = P(X \leqslant x) = P(X = 2) = \dfrac{1}{36}$.
- si $3 \leqslant x < 4$, $F(x) = P(X \leqslant x) = P(X = 2) + P(X = 3) = \dfrac{1}{36} + \dfrac{2}{36} = \dfrac{3}{36} = \dfrac{1}{12}$.

- si $4 \leqslant x < 5$, $F(x) = P(X \leqslant x) = P(X = 2) + P(X = 3) + P(X = 4) = \dfrac{3}{36} + \dfrac{3}{36} = \dfrac{6}{36} = \dfrac{1}{6}$.
- si $5 \leqslant x < 6$, $F(x) = P(X \leqslant x) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) = \dfrac{6}{36} + \dfrac{4}{36} = \dfrac{10}{36} = \dfrac{5}{18}$.
- si $6 \leqslant x < 7$, $F(x) = \dfrac{10}{36} + \dfrac{5}{36} = \dfrac{15}{36} = \dfrac{5}{12}$.
- si $7 \leqslant x < 8$, $F(x) = \dfrac{15}{36} + \dfrac{6}{36} = \dfrac{21}{36} = \dfrac{7}{12}$.
- si $8 \leqslant x < 9$, $F(x) = \dfrac{21}{36} + \dfrac{5}{36} = \dfrac{26}{36} = \dfrac{13}{18}$.
- si $9 \leqslant x < 10$, $F(x) = \dfrac{26}{36} + \dfrac{4}{36} = \dfrac{30}{36} = \dfrac{5}{6}$.
- si $10 \leqslant x < 11$, $F(x) = \dfrac{30}{36} + \dfrac{3}{36} = \dfrac{33}{36} = \dfrac{11}{12}$.
- si $11 \leqslant x < 12$, $F(x) = \dfrac{33}{36} + \dfrac{2}{36} = \dfrac{35}{36}$.
- si $x \geqslant 12$, $F(X) = 1$.

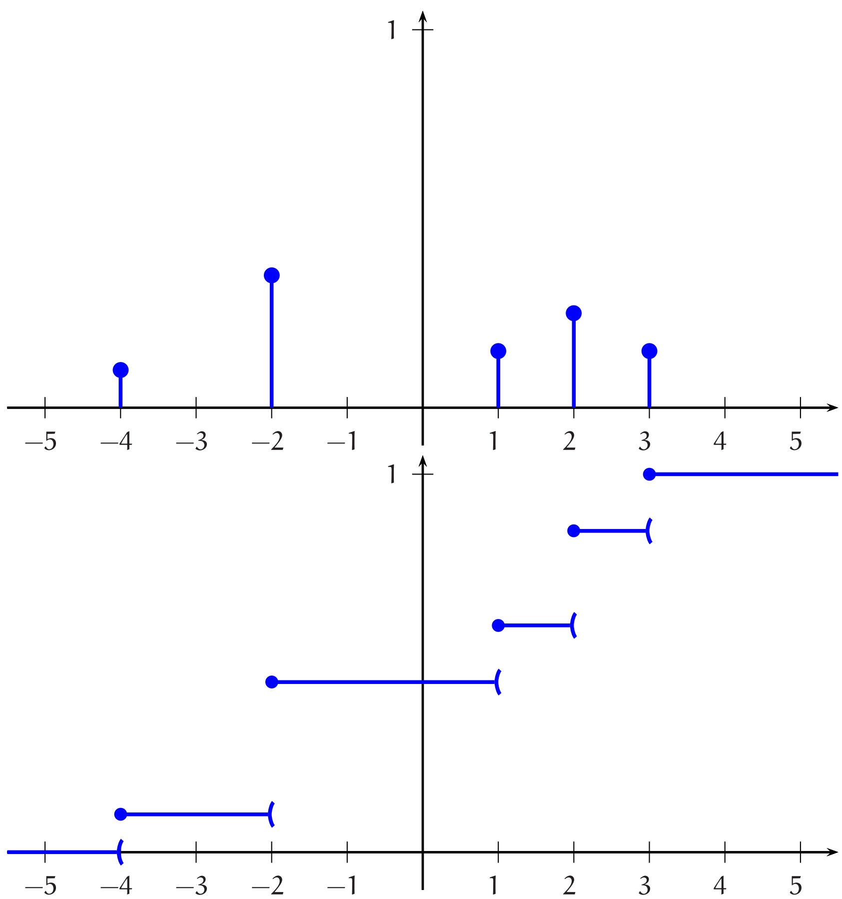
Exercice 1. La loi de probabilité d'une variable aléatoire X est donnée par le tableau suivant :

x_i	-4	-2	1	2	3
$P(X = x_i)$	$0,10$	$0,35$	0.15	$0,25$	$0,15$

1) Tracer l'histogramme de X ainsi que le graphe de la fonction de répartition de X. (L'histogramme de X représente les probabilités des x_i en fonction des x_i dans un repère du plan. Pour plus de lisibilité, on dessine des traits verticaux de l'axe des abscisses à chacun des 5 points).

2) Calculer $P(X < 0)$, $P(-3.5 < X \leqslant -2)$, $P(-3.5 < X < -2)$.

Solution 1. 1)



Pour la fonction de répartition F, on a pour $x \in \mathbb{R}$,
- si $x < -4$, $F(X) = 0$,
- si $-4 \leqslant x < -2$, $F(X) = P(X = -4) = 0,1$,
- si $-2 \leqslant x < 1$, $F(x) = P(X = -4) + P(X = -2) = 0,1 + 0,35 = 0,45$,

- si $1 \leqslant x < 2$, $F(x) = P(X = -4) + P(X = -2) + P(X = 1) = 0,1 + 0,35 + 0,15 = 0,6$,
- si $2 \leqslant x < 3$, $F(x) = P(X = -4) + P(X = -2) + P(X = 1) + P(X = 2) = 0,1 + 0,35 + 0,15 + 0,25 = 0,85$,
- si $x \leqslant 3$, $F(x) = P(X = -4) + P(X = -2) + P(X = 1) + P(X = 2) + P(X = 3) = 1$.

2) $P(X < 0) = P(X = -4) + P(X = -2) = 0,45$ ou aussi $P(X < 0) = P(X \leqslant -2) = F(-2) = 0,45$.
$P(-3,5 < X \leqslant -2) = P(X = -2) = 0,35$ ou aussi $P(-3,5 < X \leqslant -2) = P(X \leqslant -2) - P(X \leqslant -3,5) = F(-2) - F(-3,5) = 0,45 - 0,1 = 0,35$.
$P(-3,5 < X < -2) = 0$.

1.4 Image d'une variable aléatoire par une application

Soient X une variable aléatoire sur un univers fini Ω à valeurs dans un ensemble E puis f une application de E vers un certain ensemble F. On peut alors définir $Y = f \circ X$. Y est une application de Ω vers F ou encore Y **est une variable aléatoire sur Ω à valeurs dans** F. Elle se note abusivement

$$Y = f(X).$$

Ainsi, si X est une variable aléatoire à valeurs réelles, on peut par exemple définir les variables aléatoires réelles X^2, $-X$, $2X + 3$, $|X|$ ou e^X.

Exemple. On lance une pièce de monnaie. On associe à cette expérience l'univers $\Omega = \{P, F\}$. On considère la variable aléatoire définie sur Ω à valeurs dans $\{-1, 1\}$ définie par $X(P) = 1$ et $X(F) = -1$.
Les variables aléatoires X^2 et $|X|$ sont les variables aléatoires définies par : $X^2(P) = |X|(P)| = 1$ et $X^2(F) = |X|(F) = 1$. X^2 et $|X|$ sont des variables aléatoires constantes : $X^2 = |X| = 1$ où 1 est la variable aléatoire constante sur Ω : $\omega \mapsto 1$. L'événement $\left\{X^2 = 1\right\}$ est

$$\left\{X^2 = 1\right\} = \{X = 1\} \cup \{X = -1\} = \{P\} \cup \{F\} = \{P, F\} = \Omega$$

ou directement

$$\left\{X^2 = 1\right\} = \left(X^2\right)^{-1}(\{1\}) = \{\omega \in \Omega / \ X^2(\omega) = 1\} = \{P, F\} = \Omega.$$

❑

Il y a bien sûr un lien entre les lois de probabilités de X et $Y = f(X)$:

Théorème 2. Soit (Ω, P) un espace probabilisé fini. Soit X une variable aléatoire sur Ω à valeurs dans un certain ensemble non vide E. Soient f une application de E vers F puis $Y = f(X)$. Alors,

$$\forall y \in (f(X))(\Omega),\ P(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} P(X = x) = \sum_{x \in f^{-1}(\{y\})} P_X(\{x\}).$$

Démonstration . Soit $y \in (f(X))(\Omega)$. Soit $A = f^{-1}(y) = \{x \in E / \ f(x) = y\}$. Les événements $X^{-1}(x) = \{\omega \in \Omega / \ X(\omega) = x\}$ où $x \in A = f^{-1}(\{y\})$, sont deux à deux disjoints, de réunion l'événement $\{\omega \in \Omega / \ f(X(\omega)) = y\}$. Donc,

$$P(f(X) = y) = P\left(\bigcup_{x \in f^{-1}(y)} \{X = x\}\right) = \sum_{x \in f^{-1}(\{y\})} P(X = x).$$

❑

Ainsi, par exemple, si X est une variable aléatoire réelle, $P\left(X^2 = 1\right) = P(X = -1) + P(X = 1)$. Ici, f est la fonction $t \mapsto t^2$, $y = 1$ et $f^{-1}(\{y\}) = \{-1, 1\}$.

Exercice 2. On reprend la variable X de l'exercice n° 1. Déterminer les lois de probabilité des variables : $|X|$, $X^2 + X - 2$, $\min(X, 1)$, $\max\left(X, -X^2\right)$.

Solution 2. On pose $Y = |X|$, $Z = X^2 + X - 2$, $T = \min(X, 1)$, $U = \max\left(X, -X^2\right)$.

- $Y(\Omega) = \{1, 2, 3, 4\}$.

  - $Y = 1 \Leftrightarrow X = 1$. Donc, $P(Y = 1) = P(X = 1) = 0,15$.
  - $Y = 2 \Leftrightarrow (X = 2$ ou $X = -2)$. Donc, $P(Y = 2) = P(X = -2) + P(X = 2) = 0,35 + 0,25 = 0,6$.
  - $Y = 3 \Leftrightarrow X = 3$. Donc, $P(Y = 3) = P(X = 3) = 0,15$.
  - $Y = 4 \Leftrightarrow X = -4$. Donc, $P(Y = 4) = P(X = -4) = 0,1$.

La loi de probabilité de la variable Y est donnée par :

y_i	1	2	3	4
$P(Y=y_i)$	$0,15$	$0,6$	0.15	$0,1$

- $Z(\Omega)=\{0,4,10\}$.

  - $Y=0 \Leftrightarrow (X=-2$ ou $X=1)$. Donc, $P(Z=0)=P(X=-2)+P(X=1)=0,35+0,15=0,5$.
  - $Z=4 \Leftrightarrow X=2$. Donc, $P(Z=4)=P(X=2)=0,25$.
  - $Z=10 \Leftrightarrow (X=-4$ ou $X=3)$. Donc, $P(Z=10)=P(X=-4)+P(X=3)=0,1+0,15=0,25$.

La loi de probabilité de la variable Z est donnée par :

z_i	0	4	10
$P(Z=z_i)$	$0,5$	$0,25$	0.25

- $T(\Omega)=\{-4,-2,1\}$.

  - $T=-4 \Leftrightarrow X=-4$. Donc, $P(T=-4)=P(X=-4)=0,1$.
  - $T=-2 \Leftrightarrow X=-2$. Donc, $P(T=-2)=P(X=-2)=0,35$.
  - $T=1 \Leftrightarrow (X=1$ ou $X=2$ ou $X=3)$. Donc, $P(T=10)=P(X=1)+P(X=2)+P(X=3)=0,15+0,25+0,15=0,55$.

La loi de probabilité de la variable T est donnée par :

t_i	-4	-2	1
$P(T=t_i)$	$0,1$	$0,35$	0.55

- Il est clair que $U=X$.

---

2 Couples de variables aléatoires. n-uplets de variables aléatoires

2.1 Couples de variables aléatoires

2.1.1 Définition d'un couple de variables aléatoires

Soit Ω un univers fini. Soient X et Y deux variables aléatoires à valeurs dans des ensembles E et F respectivement. Le **couple de variables aléatoires** (X,Y) est l'application :

$$\begin{array}{ccc}\Omega & \to & E\times F \\ \omega & \mapsto & (X(\omega),Y(\omega))\end{array}.$$

Un couple de variables aléatoires sur Ω est donc une variable aléatoire sur Ω. Si les ensembles d'arrivées de X et Y sont E et F respectivement, alors le couple (X,Y) est une variable aléatoire dont l'ensemble d'arrivée est $E\times F$.

Exemple. On lance successivement deux dés à six faces. On associe à cette expérience l'univers $\Omega=[\![1,6]\!]^2$. On note X la somme des points obtenus et Y la différence entre le nombre obtenu sur le premier dé et le nombre obtenu sur le second dé. Par suite, on peut prendre $E=[\![2,12]\!]$ et $F=[\![-5,5]\!]$ (ou aussi $E=F=\mathbb{R}$).
Le couple (X,Y) est alors une application de $\Omega=[\![1,6]\!]^2$ vers $E\times F=[\![2,12]\!]\times[\![-5,5]\!]$ et si par exemple $\omega=(3,4)$, alors

$$(X,Y)(\omega)=(X(\omega),Y(\omega))=(X((3,4)),Y((3,4))=(7,-1).$$

2.1.2 Loi d'un couple de variables aléatoires

Définition 3. Soit (Ω,P) un espace probabilisé fini. Soient X et Y deux variables aléatoires sur Ω à valeurs dans des ensembles E et F respectivement. La **loi conjointe** des variables X et Y est la loi de probabilité du couple (X,Y).

Donc, par définition, la loi de probabilité du couple (X,Y) est l'application de $\mathcal{P}(E)\times\mathcal{P}(F)$ dans $[0,1]$ suivante :

$$P_{(X,Y)}\ :\ \begin{array}{ccc}\mathcal{P}(E)\times\mathcal{P}(F) & \to & [0,1] \\ (A,B) & \mapsto & P(\{(X,Y)\in A\times B\})\end{array}.$$

On note que l'événement $\{(X,Y) \in A \times B\}$ est encore l'événement $\{X \in A\} \cap \{Y \in B\}$. Ensuite, si A et B sont des parties non vides de E et F respectivement, alors $(\{X = x\} \cap \{Y = y\})_{(x,y)\in A\times B}$ est une partition de l'événement $\{(X,Y) \in A \times B\} = \{X \in A\} \cap \{Y \in B\}$. Donc,

$$P_{(X,Y)}(A,B) = P(\{(X,Y) \in A \times B\})P(\{X \in A\} \cap \{Y \in B\}) = \sum_{(x,y)\in A\times B} P(\{X = x\} \cap \{Y = y\}).$$

Dans la pratique, quand on demande la loi de probabilité du couple (X,Y), on se contente de donner les probabilités

$$P(\{X = x\} \cap \{Y = y\}) \text{ où } (x,y) \in (X,Y)(\Omega).$$

La famille d'événements $(\{X = x\} \cap \{Y = y\})_{(x,y)\in(X,Y)(\Omega)}$ est un système complet d'événements et donc

$$\sum_{(x,y)\in(X,Y)(\Omega)} P(\{X = x\} \cap \{Y = y\}) = 1.$$

Au passage, on fera attention au fait que $(X,Y)(\Omega)$ n'est pas $X(\Omega)\times Y(\Omega)$. $(X,Y)(\Omega)$ n'est pas l'ensemble des couples de la forme $(X(\omega), Y(\omega))$ où $\omega \in \Omega$ alors que $(X,Y)(\Omega)$ est l'ensemble des couples de la forme $(X(\omega), Y(\omega'))$ où $(\omega, \omega') \in \Omega^2$.

Exemple. On lance simultanément deux dés à six faces bien équilibrés. On note X le plus petit des nombres obtenus et Y le plus grand des nombres obtenus. On peut prendre $\Omega = [\![1,6]\!]^2$ muni de la probabilité uniforme. On a $(X,Y)(\Omega) = \{(x,y) \in [\![1,6]\!]^2/\ x \leqslant y\}$ puis $\text{card}((X,Y)(\Omega)) = 21$ (alors que $X(\Omega) \times Y(\Omega) = [\![1,6]\!]^2$ et que $\text{card}(X(\Omega) \times Y(\Omega)) = 36$).

Si (x,y) est un élément de $(X,Y)(\Omega)$ tel que $x = y$, alors l'événement $\{(X,Y) = (x,y)\}$ est l'événement $\{(x,x)\}$. Sa probabilité est $\dfrac{1}{36}$. Si (x,y) est un élément de $(X,Y)(\Omega)$ tel que $x < y$, alors l'événement $\{(X,Y) = (x,y)\}$ est l'événement $\{(x,y),(y,x)\}$. Sa probabilité est $\dfrac{2}{36} = \dfrac{1}{18}$. La loi de probabilité du couple (X,Y) est donc :

$$\forall (x,y) \in (X,Y)(\Omega),\ P(\{(X,Y) = (x,y)\}) = \begin{cases} \dfrac{1}{36} \text{ si } x = y \\ \dfrac{1}{18} \text{ si } x < y \end{cases}.$$

❑

2.1.3 Lois marginales

Commençons par un exemple. On se donne deux variables aléatoires X et Y, toutes deux à valeurs dans $\{0,1\}$. On suppose que la loi du couple (X,Y) est fournie par le tableau suivant :

X \ Y	0	1
0	$\frac{1}{6}$	$\frac{1}{4}$
1	$\frac{1}{3}$	$\frac{1}{4}$

Ainsi par exemple, $P(\{X = 0\} \cap \{Y = 1\}) = \dfrac{1}{4}$. On note que $\dfrac{1}{6} + \dfrac{1}{3} + \dfrac{1}{4} + \dfrac{1}{4} = 1$.
Le tableau ci-dessus permet de reconstituer les lois de X et Y. Par exemple,

$$P(X = 0) = P((X = 0) \cap (Y = 0)) + P((X = 0) \cap (Y = 1)) = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}.$$

La loi de X s'obtient en calculant la somme des probabilités de chaque ligne et la loi de Y s'obtient en calculant la somme des probabilités de chaque colonne. Ces totaux partiels sont alors reportés **en marge** de ce tableau :

X \ Y	0	1	X
0	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{5}{12}$
1	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{7}{12}$
Y	$\frac{1}{2}$	$\frac{1}{2}$	1

Pour cette raison, les lois de X et de Y sont appelées les lois marginales du couple (X, Y). Plus généralement,

DÉFINITION 4. Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé fini. La **première loi marginale** du couple (X, Y) est la loi de X et la **deuxième loi marginale** du couple (X, Y) est la loi de Y.

Puisque les familles $(\{X = x\})_{x \in X(\Omega)}$ et $(\{Y = y\})_{y \in Y(\Omega)}$ sont des systèmes complets d'événements, on a immédiatement :

Théorème 3. Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé fini (Ω, P). Alors,

$$\forall x \in X(\Omega),\ P(X = x) = \sum_{y \in Y(\Omega)} P((X, Y) = (x, y)) = \sum_{y \in Y(\Omega)} P(\{X = x\} \cap \{Y = y\})$$

et

$$\forall y \in Y(\Omega),\ P(Y = y) = \sum_{x \in X(\Omega)} P((X, Y) = (x, y)) = \sum_{x \in X(\Omega)} P(\{X = x\} \cap \{Y = y\}).$$

Ainsi, la connaissance de la loi conjointe du couple (X, Y) permet de déterminer les lois marginales de ce couple, c'est-à-dire les lois de X et de Y.

Exemple. Reprenons l'exemple du lancer de deux dés où on appelle X le plus petit des numéros obtenus et Y le plus grand. On a vu qu'on peut prendre $\Omega = [\![1, 6]\!]^2$ muni de la probabilité uniforme puis $(X, Y)(\Omega) = \left\{(x, y) \in \Omega^2 / x \leqslant y\right\}$.

Déterminons les lois marginales du couple (X, Y). On a $X(\Omega) = [\![1, 6]\!]$. Soit $i \in [\![1, 6]\!]$. Si $i < 6$,

$$P(X = i) = \sum_{j=i}^{6} P(\{X = i\} \cap \{Y = j\}) = \frac{1}{36} + \sum_{j=i+1}^{6} \frac{1}{18} = \frac{1}{36} + \frac{6 - i}{18} = \frac{13 - 2i}{36}$$

ce qui reste vrai si $i = 6$. De même, si $j > 1$,

$$P(Y = j) = \sum_{i=1}^{j} P(\{X = i\} \cap \{Y = j\}) = \sum_{i=1}^{j-1} \frac{1}{18} + \frac{1}{36} = \frac{j - 1}{18} + \frac{1}{36} = \frac{2j - 1}{36},$$

ce qui reste clairement vrai si $j = 1$. Donnons les lois de probabilité de X et Y dans un tableau :

k	1	2	3	4	5	6
$P(X = k)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$
$P(Y = k)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Exercice 3. Soient X et Y deux variables aléatoires réelles prenant leurs valeurs dans $\{0, 1, 2\}$. On suppose que la loi conjointe est donnée par le tableau suivant :

Y \ X	0	1	2
0	p	$p/2$	$p/4$
1	$2p$	p	$p/2$
2	$4p$	$2p$	p

1) Pour quelle valeur de p ce tableau représente-t-il effectivement la loi d'un couple ?

2) Déterminer les lois marginales.

Solution 3.

1) Le tableau représente effectivement une loi de couple si et seulement si les nombres écrits dans ce tableau sont positifs, de somme égale à 1. Or,

$$p + \frac{p}{2} + \frac{p}{4} + 2p + p + \frac{p}{2} + 4p + 2p + p = 12p + \frac{p}{4} = \frac{49p}{4},$$

et donc, le tableau représente effectivement une loi de couple si et seulement si $p = \frac{4}{49}$. Dorénavant, $p = \frac{4}{49}$.

2) La première loi marginale du couple (X, Y) est la loi de X :

- $P(X=0) = p + 2p + 4p = 7p = \frac{4}{7}$.
- $P(X=1) = \frac{p}{2} + p + 2p = \frac{1}{2}P(X=0) = \frac{2}{7}$.
- $P(X=2) = \frac{1}{2}P(X=1) = \frac{1}{7}$.

La deuxième loi marginale du couple (X, Y) est la loi de Y :

- $P(Y=0) = p + \frac{p}{2} + \frac{p}{4} = \frac{7p}{4} = \frac{1}{7}$.
- $P(Y=1) = 2P(Y=0) = \frac{2}{7}$.
- $P(Y=2) = 2P(Y=1) = \frac{4}{7}$.

2.1.4 Lois conditionnelles

Définition 5. Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé fini.
Pour $y \in Y(\Omega)$, tel que $P(Y=y) \neq 0$, la **loi conditionnelle de la variable X sachant l'événement** $\{Y=y\}$ est la loi de la variable X dans l'espace probabilisé $\left(\Omega, P_{\{Y=y\}}\right)$. Donner cette loi conditionnelle, c'est fournir les probabilités

$$P_{\{Y=y\}}(X=x),\ x \in X(\Omega).$$

Les formules usuelles sur les probabilités conditionnelles s'appliquent ici : définition, formule des probabiltés composées, formule de Bayes ... Par exemple, si $(x, y) \in (X, Y)(\Omega)$ est tel que $P(Y=y) \neq 0$, alors

$$P_{\{Y=y\}}(X=x) = \frac{P(\{X=x\} \cap \{Y=y\})}{P(Y=y)} = \frac{P((X, Y) = \{x, y\})}{P(Y=y)}$$

et aussi

$$P((X, Y) = \{x, y\}) = P(\{X=x\} \cap \{Y=y\}) = P(Y=y) \times P_{\{Y=y\}}(X=x).$$

2.2 n-uplets de variables aléatoires

Ce qui précède ce généralise aisément à n variables aléatoires X_1, ..., X_n sur un même espace probabilisé :

Définition 6. Soient $n \geqslant 2$ puis X_1, ..., X_n des variables aléatoires sur un univers fini Ω à valeurs dans des ensembles E_1, ..., E_n respectivement. L'application

$$\begin{array}{ccc} \Omega & \to & E_1 \times \ldots \times E_n \\ \omega & \mapsto & (X_1(\omega), \ldots, X_n(\omega)) \end{array}.$$

C'est une application de Ω dans $E_1 \times \ldots \times E_n$.

Si de plus, X_1, ..., X_n des variables aléatoires sur un espace probabilisé fini (Ω, P). La loi du n-uplet $(X_1, \ldots, X_n)$ est la **loi conjointe** des variables X_1, ..., X_n.

Les **lois marginales** du n-uplet $(X_1, \ldots, X_n)$ sont les lois des variables X_1, ..., X_n.

Pour $(x_1, \ldots, x_n) \in (X_1, \ldots, X_n)(\Omega) \subset X_1(\Omega) \times \ldots \times X_n(\Omega)$, on a

$$P\left((X_1, \ldots, X_n) = (x_1, \ldots, x_n)\right) = P\left(\{X_1 = x_1\} \cap \ldots \cap \{X_n = x_n\}\right).$$

La probabilité $P\left((X_1, \ldots, X_n) = (x_1, \ldots, x_n)\right)$ peut se noter plus simplement $P\left(X_1 = x_1, \ldots, X_n = x_n\right)$.

A partir de la formule des probabilités totales, les lois marginales s'obtiennent ainsi :

$$\forall (x_1, \ldots, x_n) \in (X_1, \ldots, X_n)(\Omega),\ \forall i \in [\![1, n]\!],$$
$$P(X_i = x_i) = \sum_{(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \prod_{j \neq i} X_j(\Omega)} P(X_1 = x_1, \ldots, X_n = x_n).$$

3 Variables aléatoires indépendantes

3.1 Couples de variables aléatoires indépendantes

Définition 7. Soient X et Y deux variables aléatoires sur un espace probabilisé fini (Ω, P) à valeurs dans des ensembles E et F respectivement.

Les variables X et Y sont **indépendantes** si et seulement si

$$\forall (A, B) \in \mathcal{P}(E) \times \mathcal{P}(F),\ P(\{X \in A\} \cap \{Y \in B\}) = P(X \in A) \times P(Y \in B).$$

ce qui équivaut à dire que pour tout $(A, B) \in \mathcal{P}(E) \times \mathcal{P}(F)$, les événements $\{X \in A\}$ et $\{Y = B\}$ sont indépendants.

Dans ce cas, on dit aussi que le couple (X, Y) est un couple de variables aléatoires indépendantes.

Si l'on utilise cette définition pour prouver que deux variables aléatoires sont indépendantes, la liste des vérifications à effectuer est très longue. Le théorème suivant diminue largement cette liste de vérifications :

Théorème 4. Soient X et Y deux variables aléatoires sur un espace probabilisé fini (Ω, P).

Les variables X et Y sont **indépendantes** si et seulement si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega),\ P(\{X = x\} \cap \{Y = y\}) = P(X = x) \times P(Y = y).$$

Démonstration . Supposons les variables X et Y indépendantes. Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. En appliquant la définition aux parties $A = \{x\}$ et $B = \{y\}$, on obtient $P(\{X = x\} \cap \{Y = y\}) = P(X = x) \times P(Y = y) \quad (*)$.

Réciproquement, supposons $(*)$. Soient A et B deux parties de E et F respectivement. Si l'un des événements $\{X \in A\}$ ou $\{Y \in B\}$ est vide, alors $\{X \in A\} \cap \{Y \in B\}$ est vide et on a donc $P(\{X \in A\} \cap \{Y \in B\}) = 0 = P(X \in A) \times P(Y \in B)$.
On suppose maintenant que les deux événements $\{X \in A\}$ et $\{Y \in B\}$ sont non vides. D'après la formule des probabilités totales,

$$\begin{aligned} P(\{X \in A\} \cap \{Y \in B\}) &= \sum_{(x,y) \in A \times B} P((X, Y) = (x, y)) \\ &= \sum_{(x,y) \in A \times B} P(X = x) \times P(Y = y) = \left(\sum_{x \in A} P(X = x)\right) \times \left(\sum_{y \in B} P(Y = y)\right) \\ &= P(X \in A) \times P(Y \in B). \end{aligned}$$

Les variables X et Y sont donc indépendantes.

❏

Théorème 5. Soient X et Y deux variables aléatoires indépendantes sur un espace probabilisé fini (Ω, P) à valeurs dans des ensembles E et F respectivement. Soient f et g des applications de E vers un ensemble E' et F vers un ensemble F' respectivement.

Alors, les variables $f(X)$ et $g(Y)$ sont indépendantes.

Démonstration . Soit $(A', B') \in \mathcal{P}(E') \times \mathcal{P}(F')$. Soient $A = f^{-1}(A')$ et $B = g^{-1}(B')$ (images réciproques des parties A' et B' par f et g respectivement). Puisque $f(X) \in A' \Leftrightarrow X \in f^{-1}(A')$, on a $\{f(X) \in A'\} = \{X \in A\}$ et de même $\{f(Y) \in B'\} = \{Y \in B\}$ et $\{f(X) \in A'\} \cap \{g(Y) \in B'\} = \{X \in A\} \cap \{Y \in B\}$ puis

$$P(\{f(X) \in A'\} \cap \{g(Y) \in B'\}) = \sum P(\{X \in A\} \cap \{Y \in B\}) = P(X \in A) \times P(Y \in B) = P(f(X) \in A') \times P(g(Y) \in B').$$

Les variables $f(x)$ et $g(Y)$ sont donc indépendantes.

❏

Ainsi, si X et Y sont indépendantes, X^2 et $|Y|$ sont indépendantes, X^p et Y^q sont indépendantes pour $(p, q) \in \mathbb{N}^2$, e^X et Y sont indépendantes, e^X et e^Y sont indépendantes ...

3.2 n-uplets de variables aléatoires indépendantes

Définition 8. Soient $n \geqslant 2$ puis X_1, ..., X_n des variables aléatoires sur un espace probabilisé fini (Ω, P) à valeurs dans des ensembles E_1, ..., E_n respectivement.

Les variables X_1, ..., X_n sont **mutuellement indépendantes** (ou plus simplement indépendantes) si et seulement si pour tout $(A_1, \ldots, A_n) \in \mathcal{P}(E_1) \times \ldots \mathcal{P}(E_n)$, les événements $\{X_1 \in A_1\}$, ..., $X_n \in \{A_n\}$ sont indépendants. Ceci revient à dire que pour tout $(A_1, \ldots, A_n) \in \mathcal{P}(E_1) \times \ldots \mathcal{P}(E_n)$ et toute partie non vide I de $[\![1, n]\!]$,

$$P\left(\bigcap_{i \in I} \{X_i \in A_i\}\right) = \prod_{i \in I} P(X_i \in A_i).$$

Par récurrence sur n à partir du théorème 4, on obtient

Théorème 6. Les variables X_1, ..., X_n sont indépendantes si et seulement si

$$\text{pour tout } (x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega), \text{ pour toute partie non vide } I \text{ de } [\![1, n]\!],$$
$$P\left(\bigcap_{i \in I} \{X_i = x_i\}\right) = \prod_{i \in I} P(X_i = x_i).$$

Définition 9. Soient $n \geqslant 2$ puis X_1, ..., X_n des variables aléatoires sur un espace probabilisé fini (Ω, P) à valeurs dans des ensembles E_1, ..., E_n respectivement.

Les variables X_1, ..., X_n sont **deux à deux indépendantes** si et seulement si, pour tout $(i, j) \in [\![1, n]\!]^2$ tel que $i \neq j$, les variables X_i et X_j sont indépendantes.

et on a bien sûr

Théorème 7. Si les variables X_1, ..., X_n sont indépendantes alors elles sont deux à deux indépendantes.

La réciproque est fausse.

Le théorème 5 peut se généraliser : soient X_1, ..., X_n, sont n variables aléatoires indépendantes. Soit $(I_1, \ldots, I_p)$ une partition de $[\![1, n]\!]$. Pour chaque $k \in [\![1, p]\!]$, on note Y_k une fonction des variables X_j, $j \in I_k$ (par exemple, $Y_1 = e^{X_1 X_3} - X_7^2$). Alors, les variables Y_1, ..., Y_p sont indépendantes. Ce résultat est connu sous le nom de : « lemme des coalitions ».

4 Espérance d'une variable aléatoire réelle

4.1 Définition de l'espérance

Définition 10. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

L'**espérance** de la variable aléatoire X est :

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

Si on note x_1, ..., x_n, les valeurs réelles deux à deux distinctes prises par la variable X, cette espérance peut aussi s'écrire

$$E(X) = \sum_{i=1}^{n} x_i \times P(X = x_i).$$

⇨ **Commentaire** .

◇ *Il faut bien faire la différence entre l'espérance d'une variable aléatoire et la moyenne d'une série statistique comme par exemple, la moyenne des notes d'une classe.*
La moyenne d'une série statistique donne un résultat moyen qui s'est produit dans le passé. L'espérance comme son nom l'indique nous donne un résultat moyen que l'on peut espérer dans le futur.

◇ *Supposons que la variable X prenne les valeurs x_1, ..., x_n et que les événements $\{X = x_i\}$, $i \in [\![1, n]\!]$, soient équiprobables. Donc, pour tout $i \in [\![1, n]\!]$, $P(X = x_i) = \dfrac{1}{n}$. (On dit alors que la variable X suit la **loi uniforme**. Ceci sera étudié plus loin.) L'espérance de X est donnée par :*

$$E(X) = \frac{x_1 + \ldots + x_n}{n}.$$

Dans ce cas, l'espérance de X *est tout simplement la moyenne arithmétique des valeurs prises par* X*. Dans le cas général, l'espérance de* X *doit être comprise comme la* ***moyenne pondérée*** *des valeurs prises par* X*, les coefficients de pondération étant les probabilités de chaque événement* $\{X = x_i\}$*.*

Les événements $\{\omega\}$ où $\omega \in \Omega$ n'apparaissent pas dans la définition ci-dessus. Il peut être utile en certaines circonstances d'avoir une expression de l'espérance utilisant ces événements élémentaires (on notera comme on l'a déjà fait p_ω la probabilité $P(\{\omega\})$) :

Théorème 8. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Alors,

$$E(X) = \sum_{\omega \in \Omega} X(\omega) p_\omega.$$

Démonstration .

$$\begin{aligned}
E(X) &= \sum_{x \in X(\Omega)} x \times P(X = x) = \sum_{x \in X(\Omega)} x \left(\sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P(\{\omega\}) \right) = \sum_{x \in X(\Omega)} \left(\sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} x p_\omega \right) \\
&= \sum_{x \in X(\Omega)} \left(\sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} X(\omega) p_\omega \right) \\
&= \sum_{\omega \in \Omega} X(\omega) p_\omega \text{ (car } (\{X = x\})_{x \in X(\Omega)} \text{ est un système complet d'événements).}
\end{aligned}$$

❑

4.2 Propriétés de l'espérance

Théorème 9. (linéarité de l'espérance)

Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P) et λ et μ deux réels. Alors,

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Démonstration .

$$\begin{aligned}
E(\lambda X + \mu Y) &= \sum_{\omega \in \Omega} (\lambda X(\omega) + \mu Y(\omega)) p_\omega = \lambda \sum_{\omega \in \Omega} X(\omega) p_\omega + \mu \sum_{\omega \in \Omega} Y(\omega) p_\omega \\
&= \lambda E(X) + \mu E(Y).
\end{aligned}$$

❑

Ensuite, on a immédiatement :

Théorème 10. (positivité de l'espérance)

Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Si X est à valeurs positives, alors $E(X) \geqslant 0$.

De la positivité et la linéarité de l'espérance, on déduit la :

Théorème 11. (croissance de l'espérance)

Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P). Si $X \leqslant Y$, alors $E(X) \leqslant E(Y)$.

Démonstration . Supposons $X \leqslant Y$ (c'est-à-dire $\forall \omega \in \Omega,\ X(\omega) \leqslant Y(\omega)$). Alors, $Y - X \geqslant 0$ puis

$$E(Y) - E(X) = E(Y - X) \geqslant 0$$

et donc $E(X) \leqslant E(Y)$.

❑

Théorème 12. Soit (Ω, P) un espace probabilisé fini. Soit $A \in \mathcal{P}(\Omega)$ (ou encore soit A un événement).
L'espérance de 1_A la variable indicatrice de A est $P(A)$.

Démonstration.

$$\begin{aligned} E(1_A) &= \sum_{\omega\in\Omega} 1_A(\omega)P(\{\omega\}) = \sum_{\omega\in A} 1\times P(\{\omega\}) + \sum_{\omega\notin A} 0\times P(\{\omega\}) \\ &= \sum_{\omega\in A} P(\{\omega\}) = P(A). \end{aligned}$$

❑

Théorème 13. Soit (Ω, P) un espace probabilisé fini.

1) Soit a un réel. L'espérance de la variable aléatoire constante égale à a est a.

2) Soient a et b deux réels. Soit X une variable aléatoire réelle sur (Ω, P). $E(aX+b) = aE(X)+b$.

Démonstration.

1) Notons X la variable constante égale à a.

$$E(X) = \sum_{\omega\in\Omega} X(\omega)P(\{\omega\}) = a\sum_{\omega\in\Omega} P(\{\omega\}) = a.$$

2) Par linéarité de l'espérance (en notant 1 la variable aléatoire constante égale à 1),

$$E(aX+b) = aE(X)+bE(1) = aE(X)+b.$$

❑

Définition 11. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).
X est **centrée** si et seulement si $E(X) = 0$.

Théorème 14. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Alors, la variable $X - E(X)$ est centrée.

La variable $X - E(X)$ est appelée la **variable centrée associée à** X.

L'espérance est souvent utilisée pour créer les règles d'un jeu de sorte que le jeu soit équitable. On considère la variable aléatoire « gain algébrique » d'un joueur (le gain algébrique est négatif quand le joueur perd de l'argent). Le jeu est équitable quand, en moyenne le joueur ne gagne rien ... et ne perd rien. Dit autrement, le jeu est équitable si et seulement si l'espérance est nulle (ou encore la variable X est centrée) :

Exercice 4. Deux joueurs A et B lancent deux pièces de monnaie.
Si les deux pièces tombent sur pile, A gagne. Sinon, B gagne 2 €. Combien doit gagner A pour que le jeu soit équitable ?

Solution 4. A partir de l'énoncé, on doit comprendre que quand B gagne, A donne 2 euros à B est perd donc deux euros et quand A gagne, B donne x euro à A, où x est à déterminer, et perd donc x euros.

Notons X et Y les variables aléatoires respectivement égales au gain algébrique de A et de B.

$X(\Omega) = \{x, -2\}$ puis $P(X = x) = \dfrac{1}{4}$ et $P(X = -2) = \dfrac{3}{4}$ puis $E(X) = \dfrac{x}{4} - \dfrac{3}{4}\times 2 = \dfrac{x-6}{4}$.
$Y = -X$ et donc $E(Y) = -E(X) = \dfrac{6-x}{4}$.

Le jeu est équitable si et seulement si $E(X) = E(Y) = 0$ ce qui équivaut à $x = 6$.

Exercice 5. Soient $n \in \mathbb{N}^*$ et $\lambda \in \mathbb{R}^+$. On considère une variable aléatoire réelle X prenant ses valeurs dans l'ensemble $[\![1, n]\!]$ et dont la loi est de la forme : $\forall k \in [\![1, n]\!], \quad P(X = k) = \lambda k$.

1) Déterminer λ.

2) Calculer $E(X)$.

Solution 5.

1) $1 = \sum_{k=1}^{n} P(X = k) = \lambda \sum_{k=1}^{n} k = \dfrac{\lambda n(n+1)}{2}$ et donc $\lambda = \dfrac{2}{n(n+1)}$ puis

$$\forall k \in [\![1, n]\!],\ P(X = k) = \frac{2k}{n(n+1)}.$$

2)

$$E(X) = \sum_{k=1}^{n} kP(X = k) = \frac{2}{n(n+1)} \sum_{k=1}^{n} k^2 = \frac{2}{n(n+1)} \times \frac{n(n+1)(2n+1)}{6} = \frac{2n+1}{3}.$$

4.3 Le théorème de transfert

Théorème 15. (le théorème de transfert)

Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Soit f une application de $X(\Omega)$ dans $\mathbb{R}$. Alors,

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x) \times P(X = x).$$

Démonstration. On sait que la loi de la variable $f(X)$ (théorème 2) est donnée par

$$\forall y \in f(X)(\Omega),\ P(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} P(X = x).$$

Donc

$$\begin{aligned} E(f(X)) &= \sum_{y \in f(X)(\Omega)} yP(f(X) = y) = \sum_{y \in f(X)(\Omega)} \left(y \sum_{x \in f^{-1}(\{y\})} P(X = x) \right) \\ &= \sum_{y \in f(X)(\Omega)} \left(\sum_{x \in f^{-1}(\{y\})} f(x)P(X = x) \right) \end{aligned}$$

Comme $X(\Omega)$ est la réunion disjointe des $f^{-1}(\{y\})$ où y décrit $f(X)(\Omega)$, on obtient

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

❏

Ainsi, par exemple, $E\left(X^2\right) = \sum_{x \in X(\Omega)} x^2 P(X = x)$.

4.4 L'inégalité de Markov

Théorème 16. (l'inégalité de Markov)

Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Si X est positive, alors

$$\forall a > 0,\ P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

Démonstration. Soit X une variable aléatoire réelle positive. Soit $a > 0$.

Démonstration 1.

$$\begin{aligned} E(X) &= \sum_{x \in X(\Omega)} xP(X = x) \\ &\geqslant \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} xP(X = x) \text{ (car X est à valeurs positives)} \\ &\geqslant a \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} P(X = x) = aP(X \geqslant a). \end{aligned}$$

Puisque $a > 0$, on en déduit que $P(X \geqslant a) \leqslant \dfrac{E(X)}{a}$.

Démonstration 2. Notons A l'événement $\{X \geqslant a\}$. On a vu que $P(X \geqslant a) = P(A) = E(1_A)$. On note que $X \geqslant a \Leftrightarrow \dfrac{1}{a}X \geqslant 1$.
Soit $\omega \in \Omega$. Si $X(\omega) \in A$, alors $1_A(\omega) = 1 \leqslant \dfrac{1}{a}X(\omega)$ et si $X(\omega) \notin A$, alors $1_A(\omega) = 0 \leqslant \dfrac{1}{a}X(\omega)$ (car $\dfrac{1}{a}X$ est une variable positive).

En résumé, $1_A \leqslant \dfrac{1}{a}X$. Par croissance et linéarité de l'espérance, on en déduit

$$P(X \geqslant a) = E(1_A) \leqslant E\left(\frac{1}{a}X\right) = \frac{E(X)}{a}.$$

❏

4.5 Espérance d'un produit de variables indépendantes

Théorème 17.

Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P). Si X et Y sont indépendantes, alors $E(XY) = E(X)E(Y)$.

Démonstration .

$$\begin{aligned}
E(XY) &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP((X,Y) = (x,y)) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(\{X = x\} \cap \{Y = y\}) \\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x) \times P(Y = y) \text{ (car } X \text{ et } Y \text{ sont indépendantes)} \\
&= \left(\sum_{x\in X(\Omega)} xP(X = x)\right)\left(\sum_{y\in Y(\Omega)} yP(Y = y)\right) = E(X)E(Y).
\end{aligned}$$

❏

Nous terminons cette séquence sur les variables aléatoires et leur espérance par deux exercices de la banque d'exercices des oraux des CCP.

Exercice 6. On dispose de n boules numérotées de 1 à n et d'une boîte formée de trois compartiments identiques également numérotés de 1 à 3.
On lance simultanément les n boules.
Elles viennent se ranger aléatoirement dans les 3 compartiments.
Chaque compartiment peut éventuellement contenir les n boules.
On note X la variable aléatoire qui à chaque expérience aléatoire fait correspondre le nombre de compartiments restés vides.

1) Préciser les valeurs prises par X.

2) a) Déterminer la probabilité $p(X = 2)$.

b) Finir de déterminer la loi de probabilité de X.

3) a) Calculer $E(X)$.

b) Déterminer $\lim\limits_{n\to+\infty} E(X)$. Interpréter ce résultat.

Solution 6.

1) X prend les valeurs 0, 1 ou 2.

2) a) $X = 2$ est l'événement « toutes les boules vont dans le même compartiment ». Il y a 3^n répartitions possibles des n boules dans les 3 compartiments (pour chacune des n boules, il y a 3 possibilités de compartiment). Parmi ces répartitions, il y en a une et une seule pour laquelle toutes les boules sont dans le compartiment n° 1, une et une seule pour laquelle toutes les boules sont dans le compartiment n° 2 et une et une seule pour laquelle toutes les boules sont dans le compartiment n° 3. Donc

$$p(X = 2) = \frac{3}{3^n} = \frac{1}{3^{n-1}}.$$

b) Soit E l'événement : « le troisième compartiment est vide et les deux premiers ne le sont pas ». On a alors

$$p(X=1) = 3 \times p(E).$$

Soit $k \in [\![1, n-1]\!]$. Soit E_k l'événement « k boules sont dans le compartiment n° 1 et $n-k$ sont dans le compartiment n° 2 ». $E = \bigcup_{1 \leqslant k \leqslant n-1} E_k$ et les E_k, $1 \leqslant k \leqslant n-1$, sont deux à deux disjoints. Donc,

$$p(X=1) = 3p(E) = 3\sum_{k=1}^{n-1} p\left(E_k\right).$$

Soit $k \in [\![1, n-1]\!]$. Le nombre de répartitions des n boules telles que k d'entre elles soient dans le compartiment n° 1 et $n-k$ soient dans le compartiment n° 2 est encore le nombre de tirages simultanés de k boules parmi les n à savoir $\binom{n}{k}$.

Donc $p\left(E_k\right) = \dfrac{\binom{n}{k}}{3^n}$. Par suite,

$$p\left(E\right) = 3\frac{\sum_{k=1}^{n-1}\binom{n}{k}}{3^n} = \frac{2^n-2}{3^{n-1}}.$$

Enfin,

$$p(X=0) = 1 - p(X=1) - p(X=2) = 1 - \frac{2^n-2}{3^{n-1}} - \frac{1}{3^{n-1}} = \frac{3^{n-1}-2^n+1}{3^{n-1}}.$$

$$\boxed{p(X=2) = \frac{1}{3^{n-1}},\ p(X=1) = \frac{2^n-2}{3^{n-1}} \text{ et } p(X=0) = \frac{3^{n-1}-2^n+1}{3^{n-1}}.}$$

3) a) $E(X) = 0 \times \dfrac{3^{n-1}-2^n+1}{3^{n-1}} + 1 \times \dfrac{2^n-2}{3^{n-1}} + 2 \times \dfrac{1}{3^{n-1}} = \dfrac{2^n}{3^{n-1}} = 3\left(\dfrac{2}{3}\right)^n.$

b) $\lim\limits_{n\to+\infty} E(X) = \lim\limits_{n\to+\infty} 3\left(\dfrac{2}{3}\right)^n = 0$. Ainsi, s'il y a un grand nombre de boules, il y a peu de chances qu'un compartiment reste vide.

Exercice 7. Soit $n \in \mathbb{N}^*$. Une urne contient n boules blanches numérotées de 1 à n et deux boules noires numérotées 1 et 2.
On effectue le tirage une à une, sans remise, de toutes les boules de l'urne.
On note X la variable aléatoire égale au rang d'apparition de la première boule blanche.
On note Y la variable aléatoire égale au rang d'apparition de la première boule numérotée 1.

1) Déterminer la loi de X.

2) Déterminer la loi de Y.

Solution 7. Ω est l'ensemble des tirages successifs sans remise des $n+2$ boules ou encore l'ensemble des permutations des $n+2$ boules. Le nombre de tirages successifs et sans remise des $n+2$ boules est $(n+2)!$ ou encore $\text{card}\,(\Omega) = (n+2)!$.

1) L'urne contient $n+2$ boules. La première boule blanche peut apparaître au premier, deuxième ou troisième tirage ou encore $X(\Omega) = [\![1, 3]\!]$.

• $X = 1$ est l'événement : « la première boule tirée est blanche ». On a n possibilités de tirer la première boule parmi les n blanches puis pour chacune de ces n possibilités, on a $(n+1)!$ possibilités de tirer les $n+1$ boules restantes. Donc

$$p(X=1) = \frac{n \times (n+1)!}{(n+2)!} = \frac{n}{n+2}.$$

• $X = 3$ est l'événement : « les deux premières boules tirées sont noires ». On a $2! = 2$ possibilités de tirer les deux premières boules puis pour chacune de ces deux possibilités, on a $n!$ possibilités de tirer les n boules restantes. Donc,

$$p(X=3) = \frac{2 \times n!}{(n+2)!} = \frac{2}{(n+1)(n+2)}.$$

• Enfin

$$p(X=2) = 1 - p(X=1) - p(X=3) = 1 - \frac{n}{n+2} - \frac{2}{(n+1)(n+2)} = \frac{(n+1)(n+2) - n(n+1) - 2}{(n+1)(n+2)}$$
$$= \frac{2n}{(n+1)(n+2)}.$$

$$\boxed{X(\Omega) = [\![1,3]\!] \text{ et } p(X=1) = \frac{n}{n+2},\ p(X=2) = \frac{2n}{(n+1)(n+2)} \text{ et } p(X=3) = \frac{2}{(n+1)(n+2)}.}$$

2) La première boule numérotée 1 peut sortir au premier, deuxième, ..., $(n+1)$-ème tirage ou encore $Y(\Omega) = [\![1, n+1]\!]$.
Soit $k \in [\![2, n+1]\!]$. L'événement $Y = k$ est l'événement « les $k-1$ premières boules ne portent pas le numéro 1 et la k-ème porte le numéro 1 ». Pour les $k-1$ premières boules, on a $n(n-1) \times \ldots \times (n-k+2) = \frac{n!}{(n-k+1)!}$ tirages possibles puis pour chacun des ces tirages on a 2 possibilités pour la k-ème boule et donc $2 \times \frac{n!}{(n-k+1)!}$ tirages possibles pour les k premières boules. Pour chacun de ces tirages, on a $(n+2-k)!$ tirages possibles des $n+2-k$ boules restantes. Finalement,

$$p(Y=k) = \frac{\frac{n!}{(n-k+1)!} \times 2 \times (n+2-k)!}{(n+2)!} = \frac{2(n+2-k)}{(n+1)(n+2)}.$$

L'événement $Y = 1$ est l'événement « la première boule porte le numéro 1 ». Il y a 2 tirages possibles pour la première boule puis pour chacun de ces deux tirages, il y a $(n+1)!$ tirages possibles des $n+1$ boules restantes. Donc

$$p(Y=1) = \frac{2 \times (n+1)!}{(n+2)!} = \frac{2(n+1)}{(n+1)(n+2)} = \frac{2(n+2-1)}{(n+1)(n+2)}.$$

Finalement

$$\boxed{Y(\Omega) = [\![1, n+1]\!] \text{ et } \forall k \in [\![1, n+1]\!],\ p(Y=k) = \frac{2(n+2-k)}{(n+1)(n+2)}.}$$

---

5 Variance et écart-type. Covariance

5.1 Moments d'une variable aléatoire

Définition 12. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Soit k un entier naturel.
Le **moment d'ordre** k de la variable X est $E\left(X^k\right)$.

Avec le théorème de transfert, on a immédiatement

Théorème 18. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Soit k un entier naturel.

$$E\left(X^k\right) = \sum_{x \in X(\Omega)} x^k P(X = x).$$

5.2 Variance et écart-type d'une variable aléatoire

Définition 13. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).
La **variance** de la variable X est : $V(X) = E\left((X - E(X))^2\right)$.

La variable aléatoire $(X - E(X))^2$ est positive. Par positivité de l'espérance, on en déduit que la variance d'une variable aléatoire réelle est toujours un nombre positif. On peut donc poser :

Définition 14. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).
L'écart-type de la variable X est : $\sigma(X) = \sqrt{V(X)}$.

Le théorème de transfert fournit :

Théorème 19. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

$$V(X) = \sum_{x \in X(\Omega)} (x - E(X))^2 P(X = x).$$

Théorème 20. (formule de Koenig-Huygens)
Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

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

Démonstration . Par linéarité de l'espérance,

$$V(X) = E\left(X^2 - 2XE(X) + (E(X))^2\right) = E\left(X^2\right) - 2E(X)E(X) + (E(X))^2 = E\left(X^2\right) - (E(X))^2.$$

❏

5.3 Propriétés de la variance et de l'écart-type

Théorème 21. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P). Soient a et b deux réels.

$$V(aX + b) = a^2 V(X) \quad \text{et} \quad \sigma(aX + b) = |a|\sigma(X).$$

Démonstration .

$$V(aX+b) = E\left((aX + b - E(aX+b))^2\right) = E\left((aX - aE(X))^2\right) = a^2 E\left(X - E(X))^2\right) = a^2 V(X),$$

puis, en prenant la racine carrée, $\sigma(aX+b) = \sqrt{a^2V(X)} = |a|\sqrt{V(X)} = |a|\sigma(X)$.

❏

Voici encore un exercice issu de la banque d'oraux des CCP.

Exercice 8. Une urne contient deux boules blanches et huit boules noires.

1) Un joueur tire successivement, avec remise, cinq boules dans cette urne.
Pour chaque boule blanche tirée, il gagne 2 points et pour chaque boule noire tirée, il perd 3 points.
On note X la variable aléatoire représentant le nombre de boules blanches tirées.
On note Y le nombre de points obtenus par les joueur sur une partie.

a) Déterminer la loi de X, son espérance et sa variance.

b) Déterminer la loi de Y, son espérance et sa variance.

2) Dans cette question, on suppose que les cinq tirages successifs se font sans remise.

a) Déterminer la loi de X.

b) Déterminer la loi de Y.

Solution 8.

1) a) La variable aléatoire X suit une loi binomiale (enseignée en classe de première et rappelée en terminale). En effet,

- 5 expériences identiques et indépendantes (car les tirages se font avec remise) sont effectuées.
- chaque expérience a deux issues : « la boule tirée est blanche » avec une probabilité $p = \frac{2}{10} = \frac{1}{5}$ et « la boule tirée n'est pas blanche » avec une probabilité $1 - p = \frac{4}{5}$.

La variable aléatoire X est régie par une loi binomiale de paramètres $n = 5$ et $p = \frac{1}{5}$. On sait alors que

$$X(\Omega) = [\![0, 5]\!] \text{ et } \forall k \in [\![0, 5]\!],\ p(X = k) = \binom{5}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{5-k}.$$

Plus explicitement,

- $p(X = 0) = \left(\frac{4}{5}\right)^5 = \frac{1024}{3125} = 0,32768.$

- $p(X=1) = 5 \times \dfrac{1}{5} \times \left(\dfrac{4}{5}\right)^4 = \dfrac{256}{625} = 0,4096$.
- $p(X=2) = 10 \times \left(\dfrac{1}{5}\right)^2 \times \left(\dfrac{4}{5}\right)^3 = \dfrac{128}{625} = 0,2048$.
- $p(X=3) = 10 \times \left(\dfrac{1}{5}\right)^3 \times \left(\dfrac{4}{5}\right)^2 = \dfrac{32}{625} = 0,0512$.
- $p(X=4) = 5 \times \left(\dfrac{1}{5}\right)^4 \times \dfrac{4}{5} = \dfrac{4}{625} = 0,0064$.
- $p(X=5) = \left(\dfrac{1}{5}\right)^5 = \dfrac{1}{3125} = 0,00032$.

L'espérance de X est $E(X) = np = 5 \times \dfrac{1}{5} = 1$ et la variance de X est $V(X) = np(1-p) = 5 \times \dfrac{1}{5} \times \dfrac{4}{5} = \dfrac{4}{5} = 0,8$.

b) $Y = 2X - 3(5-X) = 5X - 15$. Par suite, $Y(\Omega) = \{5k-15,\ k \in [\![0,5]\!]\} = \{-15,-10,-5,0,5,10\}$. Ensuite,

$$\forall k \in [\![0,5]\!],\ p(Y = 5k-15) = p(X=k) = \binom{5}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{5-k}.$$

Enfin, $E(Y) = 5E(X) - 15 = 5 \times 1 - 15 = -10$ et $V(Y) = V(5X-15) = 5^2V(X) = 25 \times \dfrac{4}{5} = 20$.

2) a) $X(\Omega) = \{0,1,2\}$. La loi de probabilité ne change pas si on suppose les tirages simultanés.
Le nombre de tirages simultanés de 5 boules parmi 10 est $\dbinom{10}{5}$.
Soit $k \in [\![0,2]\!]$. Au cours d'un tirage de 5 boules, on obtient k boules blanches si et seulement si on tire k boules parmi les 2 blanches et $5-k$ boules parmi les 8 noires. Il y a donc $\dbinom{2}{k} \times \dbinom{8}{5-k}$ tirages où on obtient k boules blanches. Donc,

$$\forall k \in [\![0,2]\!],\ p(X=k) = \frac{\dbinom{2}{k} \times \dbinom{8}{5-k}}{\dbinom{10}{5}}.$$

Plus explicitement,

- $p(X=0) = \dfrac{\dfrac{8\times7\times6\times5\times4}{5\times4\times3\times2}}{\dfrac{10\times9\times8\times7\times6}{5\times4\times3\times2}} = \dfrac{5\times4}{10\times9} = \dfrac{2}{9}$.
- $p(X=1) = \dfrac{2\times\dfrac{8\times7\times6\times5}{4\times3\times2}}{\dfrac{10\times9\times8\times7\times6}{5\times4\times3\times2}} = \dfrac{2\times5\times5}{10\times9} = \dfrac{5}{9}$.
- $p(X=2) = \dfrac{\dfrac{8\times7\times6}{3\times2}}{\dfrac{10\times9\times8\times7\times6}{5\times4\times3\times2}} = \dfrac{5\times4}{10\times9} = \dfrac{2}{9}$.

L'espérance de X est $E(X) = 0 \times \dfrac{2}{9} + 1 \times \dfrac{5}{9} + 2 \times \dfrac{2}{9} = 1$ et la variance de X est

$$V(X) = E\left(X^2\right) - (E(X))^2 = 0 \times \frac{2}{9} + 1 \times \frac{5}{9} + 4 \times \frac{2}{9} - 1^2 = \frac{4}{9}.$$

b) Comme à la question 1), $Y(\Omega) = \{5k-15,\ k \in [\![0,2]\!]\} = \{-15,-10,-5\}$ et

$$\forall k \in [\![0,2]\!],\ p(Y = 5k-15) = p(X=k) = \frac{\dbinom{2}{k} \times \dbinom{8}{5-k}}{\dbinom{10}{5}}.$$

Ensuite, $E(Y) = 5E(X) - 15 = -10$ et $V(Y) = 5^2E(X) = \dfrac{100}{9}$.

Définition 15. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

La variable X est **réduite** si et seulement si $\sigma(X) = 1$.

Théorème 22. Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

Si $\sigma(X) > 0$, la variable $\dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Démonstration. $E\left(\dfrac{X - E(X)}{\sigma(X)}\right) = \dfrac{1}{\sigma(X)}E(X - E(X)) = 0$ et $V\left(\dfrac{X - E(X)}{\sigma(X)}\right) = \dfrac{1}{(\sigma(X))^2}V(X) = 1$ puis $\sigma(X) = 1$.

❏

La variable $\dfrac{X - E(X)}{\sigma(X)}$ s'appelle la **variable centrée réduite associée à** X.

5.4 L'inégalité de Bienaymé-Tchebychev

Théorème 23. (inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle sur un espace probabilisé fini (Ω, P).

Pour tout $\varepsilon > 0$, $P(|X - E(X)| \geqslant \varepsilon) \leqslant \dfrac{V(X)}{\varepsilon^2}$.

Démonstration. Soit $\varepsilon > 0$. On note que $|X - E(X)| \geqslant \varepsilon \Leftrightarrow (X - E(X))^2 \geqslant \varepsilon^2$. D'après l'inégalité de Markov appliquée à la variable positive $(X - E(X))^2$ et au réel strictement positif ε^2,

$$P(|X - E(X)| \geqslant \varepsilon) = P\left((X - E(X)^2 \geqslant \varepsilon^2\right) \leqslant \frac{1}{\varepsilon^2}E\left((X - E(X))^2\right) = \frac{V(X)}{\varepsilon^2}.$$

❏

Exercice 9. On lance n fois un dé à 6 faces. Comment choisir n pour que la probabilité d'obtenir un nombre de 6 compris entre 0 et $n/3$ soit supérieure à $1/2$?

Solution 9. Soit X la variable aléatoire égale au nombre de 6 obtenus en n lancers. X suit une loi binomiale de paramètres n et $p = \dfrac{1}{6}$. Donc, $E(X) = \dfrac{n}{6} = \mu$ et $V(X) = n \times \dfrac{1}{6} \times \dfrac{5}{6} = \dfrac{5n}{36}$.

$$P\left(0 \leqslant X \leqslant \frac{n}{3}\right) = P\left(\left|X - \frac{n}{6}\right| \leqslant \frac{n}{6}\right) \geqslant P\left(\left|X - \frac{n}{6}\right| < \frac{n}{6}\right) = 1 - P\left(|X - \mu| \geqslant \frac{n}{6}\right)$$

D'après l'inégalité de Bienaymé-Tchebychev,

$$P\left(|X - \mu| \geqslant \frac{n}{6}\right) \leqslant \frac{6^2 V(X)}{n^2} = \frac{36 \times \dfrac{5n}{36}}{n^2} = \frac{5}{n}$$

et donc

$$P\left(0 \leqslant X \leqslant \frac{n}{3}\right) \geqslant 1 - P\left(\left|X - \frac{n}{6}\right| \geqslant \frac{n}{6}\right) \geqslant 1 - \frac{5}{n},$$

puis

$$P\left(0 \leqslant X \leqslant \frac{n}{3}\right) \geqslant \frac{1}{2} \Leftarrow 1 - \frac{5}{n} \geqslant \frac{1}{2} \Leftarrow n \geqslant 10.$$

Pour $n \geqslant 10$, $P\left(0 \leqslant X \leqslant \dfrac{n}{3}\right) \geqslant \dfrac{1}{2}$.

On a obtenu des valeurs de n pour lesquelles on est sûr que $P\left(0 \leqslant X \leqslant \dfrac{n}{3}\right) \geqslant \dfrac{1}{2}$. Si l'on veut toutes les valeurs de n telles que $P\left(0 \leqslant X \leqslant \dfrac{n}{3}\right) \geqslant \dfrac{1}{2}$, il reste tester à la main les valeurs de $n \in [\![1, 9]\!]$, ce que nous ne ferons pas ici.

5.5 Covariance d'un couple de variables aléatoires

DÉFINITION 16. Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P).

La **covariance** du couple (X, Y) est : $\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$.

⇨ **Commentaire** . *On doit remarquer que* $\text{cov}(X, X) = V(X)$.

Immédiatement,

Théorème 24. Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P).

$$\text{cov}(X, Y) = \sum_{(x,y)\in(X,Y)(\Omega)} (x - E(X))(y - E(Y))P((X, Y) = (x, y)).$$

Théorème 25. Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P).

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y).$$

DÉMONSTRATION . Par linéarité de l'espérance,

$$\begin{aligned}\text{cov}(X, Y) &= E((X - E(X))(Y - E(Y))) = E(XY - XE(Y) - YE(X) + E(X)E(Y)) = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)\\ &= E(XY) - E(X)E(Y).\end{aligned}$$

❏

On a vu que l'espérance du produit de deux variables indépendantes X et Y est : $E(XY) = E(X)E(Y)$. Donc,

Théorème 26. Soient X et Y deux variables aléatoires réelles sur un espace probabilisé fini (Ω, P).

Si X et Y sont indépendantes, alors $\text{cov}(X, Y) = 0$.

Par contraposition, $\text{cov}(X, Y) \neq 0 \Rightarrow X$ et Y ne sont pas indépendantes. Ainsi, la covariance est un outil pour se convaincre que deux variables ne sont pas indépendantes. Mais attention, la covariance n'est pas un outil pour savoir si deux variables sont indépendantes :

($\text{cov}(X, Y) = 0 \not\Rightarrow X$ et Y indépendantes) ou aussi (X et Y non indépendantes $\not\Rightarrow \text{cov}(X, Y) \neq 0$).

Construisons explicitement un exemple de deux variables non indépendantes dont la covariance est nulle. Commençons par prendre une variable X d'espérance nulle : $X(\Omega) = \{-1, 0, 1\}$ avec $P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}$. Pour cette variable X, on a $E(X) = 0$ et donc $E(X)E(Y) = 0$. Choisissons alors pour Y une variable telle que $XY = 0$ et non indépendante de X. On peut prendre pour Y la variable qui prend la valeur 1 quand $X = 0$ et 0 sinon, c'est-à-dire la variable indicatrice de l'événement $\{X = 0\}$. Pour ce choix de Y, on a $XY = 0$ puis $\text{cov}(X, Y) = E(XY) - E(X)E(Y) = 0$.
Vérifions enfin que X et Y ne sont pas indépendantes. L'événement $\{X = 1\}\cap\{Y = 1\}$ est vide et donc $P(\{X = 1\} \cap \{Y = 1\}) = 0$ mais

$$P(\{X = 1\}) \times P(\{Y = 1\}) = P(\{X = 1\}) \times P(\{X = 0\}) = \frac{1}{9} \neq 0.$$

Les événements $\{X = 1\}$ et $\{Y = 1\}$ ne sont pas indépendants et donc les variables X et Y ne sont pas indépendantes, bien que de covariance nulle.

Théorème 27. La covariance est une forme bilinéaire, symétrique, positive sur l'espace vectoriel des variables aléatoires réelles sur l'espace probabilisé fini (Ω, P).

DÉMONSTRATION . On sait que l'ensemble des variables aléatoires réelles, c'est-à-dire l'ensemble des applications de Ω vers $\mathbb{R}$, est un $\mathbb{R}$-espace vectoriel.

• Soient X et Y deux variables aléatoires réelles. $\text{cov}(X, Y) = E(XY) - E(X)E(Y) = E(YX) - E(Y)E(X) = \text{cov}(Y, X)$. Donc, la covariance est symétrique.

• Soient X_1, X_2 et Y trois variables aléatoires réelles et λ et μ deux réels. Par linéarité de l'espérance,

$$\begin{aligned}\text{cov}(\lambda X_1 + \mu X_2, Y) &= E((\lambda X_1 + \mu X_2) Y) - E(\lambda X_1 + \mu X_2) E(Y) = \lambda E(X_1 Y) + \mu E(X_2 Y) - \lambda E(X_1) E(Y) + \mu E(X_2) E(Y)\\ &= \lambda (E(X_1 Y) - E(X_1) E(Y)) + \mu (E(X_2 Y) - E(X_1) E(Y)) = \lambda\text{cov}(X_1, Y) + \mu\text{cov}(X_2, Y).\end{aligned}$$

Donc, la covariance est linéaire par rapport à sa première variable puis bilinéaire par symétrie.

• Soit X une variable aléatoire réelle. $\text{cov}(X, X) = V(X) \geqslant 0$. Donc, la covariance est positive.

❑

5.6 Variance d'une somme de variables aléatoires

Théorème 28. Soient $n \geqslant 2$ puis $X_1, \ldots, X_n$ n variables aléatoires réelles sur un espace probabilisé fini (Ω, P).

$$V(X_1 + \ldots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \text{cov}(X_i, X_j).$$

En particulier, pour deux variables aléatoires X et Y :

$$V(X+Y) = V(X) + 2\text{cov}(X, Y) + V(Y).$$

Démonstration . Par bilinéarité et symétrie de la covariance,

$$\begin{aligned} V(X_1 + \ldots + X_n) &= \text{cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{1 \leqslant i,j \leqslant n} \text{cov}(X_i, X_j) = \sum_{i=1}^{n} \text{cov}(X_i, X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \text{cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \text{cov}(X_i, X_j). \end{aligned}$$

❑

Le théorème 26 et le théorème précédent fournissent alors :

Théorème 29. Soient $n \geqslant 2$ puis $X_1, \ldots, X_n$ n variables aléatoires réelles sur un espace probabilisé fini (Ω, P), deux à deux indépendantes.

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i).$$

En particulier, pour deux variables aléatoires X et Y indépendantes :

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

6 Lois usuelles

Dans ce paragraphe, nous étudions trois lois usuelles : la loi uniforme, la loi de Bernoulli et la loi binomiale. Cette étude est complétée en math spé par deux autres lois : la loi géométrique et la loi de Poisson. Quand une variable aléatoire X suit une loi $\mathcal{L}$, on écrira $X \sim \mathcal{L}$.

6.1 Loi uniforme

6.1.1 Définition de la loi uniforme

On dit qu'une variable aléatoire Y prenant un nombre fini de valeurs $y_1, \ldots, y_n$, suit une loi uniforme si et seulement si les événements $\{Y = y_i\}$ sont équiprobables. Ceci modélise la situation qui consiste à choisir « au hasard » un objet un parmi n objets.

Si X est la variable égale au numéro de l'objet, on a bien sûr $P(Y = y_i) = P(X = i) = \dfrac{1}{n}$. Pour cette raison, les définitions et résultats de ce paragraphe se concentre sur une variable aléatoire prenant de manière équiprobable les valeurs $1, 2, \ldots, n$ ou plus généralement, une variable prenant des valeurs entières consécutives.

Définition 17. Soit X une variable aléatoire sur un espace probabilisé fini (Ω, P) telle que $X(\Omega) = [\![1, n]\!]$, $n \in \mathbb{N}^*$.

X suit la loi uniforme sur $[\![1, n]\!]$ si et seulement si $P(X = 1) = P(X = 2) = \ldots = P(X = n)$.

Plus généralement, si $X(\Omega) = [\![a, b]\!]$ où a et b sont deux entiers relatifs tels que $a < b$, X suit la loi uniforme sur $[\![a, b]\!]$ si et seulement si $P(X = a) = P(X = a + 1) = \ldots = P(X = b)$.

Notation. Quand une variable aléatoire X suit la loi uniforme sur $[\![1, n]\!]$ ou plus généralement sur $[\![a, b]\!]$, on écrit $X \sim \mathcal{U}([\![1, n]\!])$ ou plus généralement $X \sim \mathcal{U}([\![a, b]\!])$.

Par exemple, quand on jette une fois un dé à 6 faces non truqué, la variable X égale au numéro obtenu suit la loi uniforme sur $[\![1, 6]\!]$: $X \sim \mathcal{U}([\![1, 6]\!])$.

Immédiatement, on a :

Théorème 30. Soit $n \in \mathbb{N}^*$. $\left(X \sim \mathcal{U}([\![1, n]\!]) \Leftrightarrow \forall k \in [\![1, n]\!],\ P(X = k) = \dfrac{1}{n}\right)$.

Plus généralement, pour tout $(a, b) \in \mathbb{Z}^2$ tel que $a < b$, $\left(X \sim \mathcal{U}([\![a, b]\!]) \Leftrightarrow \forall k \in [\![a, b]\!],\ P(X = k) = \dfrac{1}{b - a + 1}\right)$.

6.1.2 Espérance, variance et écart-type de la loi uniforme

Théorème 31. Soit $n \in \mathbb{N}^*$. Soit X une variable aléatoire suivant la loi uniforme sur $[\![1, n]\!]$ ($X \sim \mathcal{U}([\![1, n]\!])$).

$$E(X) = \frac{n+1}{2}, \quad V(X) = \frac{n^2 - 1}{12}, \quad \sigma(X) = \sqrt{\frac{n^2 - 1}{12}}.$$

Démonstration . Soit $n \geqslant 1$. Soit X une variable aléatoire suivant la loi uniforme sur $[\![1, n]\!]$. Donc, $X(\Omega) = [\![1, n]\!]$ et $\forall k \in [\![1, n]\!]$, $P(X = k) = \dfrac{1}{n}$.

$$E(X) = \sum_{k=1}^{n} k \times P(X = k) = \frac{1}{n} \sum_{k=1}^{n} k = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Ensuite,

$$\begin{aligned}
V(X) = E\left(X^2\right) - (E(X))^2 &= \sum_{k=1}^{n} k^2 \times P(X = k) - \left(\frac{n+1}{2}\right)^2 = \frac{1}{n} \sum_{k=1}^{n} k^2 - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{1}{12}\left(2(n+1)(2n+1) - 3(n+1)^2\right) \\
&= \frac{(n+1)(n-1)}{12} = \frac{n^2 - 1}{12},
\end{aligned}$$

et enfin $\sigma(X) = \sqrt{\dfrac{n^2 - 1}{12}}$.

❑

Théorème 32. Soit $(a, b) \in \mathbb{Z}^2$ tel que $a \leqslant b$. Soit X une variable aléatoire suivant la loi uniforme sur $[\![a, b]\!]$ ($X \sim \mathcal{U}([\![a, b]\!])$).

$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(n^2 - 1)b - a)(b - a + 2)}{12}, \quad \sigma(X) = \sqrt{\frac{(b - a)(b - a + 2)}{12}}.$$

Démonstration . Soient a et b deux entiers relatifs tels que $a < b$. Soit X une variable aléatoire suivant la loi uniforme sur $[\![a, b]\!]$. Donc, $X(\Omega) = [\![a, b]\!]$ et $\forall k \in [\![a, b]\!]$, $P(X = k) = \dfrac{1}{b - a + 1}$.

Soient $n = b - a + 1$ et $Y = X - a + 1$ de sorte que $X = Y + a - 1$. Y suit la loi uniforme sur $[\![1, n]\!]$. Donc,

$$E(X) = E(Y) + a - 1 = \frac{(b - a + 1) + 1}{2} + a - 1 = \frac{a+b}{2},$$

Ensuite,

$$V(X) = V(Y) = \frac{n^2 - 1}{12} = \frac{(b - a + 1)^2 - 1}{12} = \frac{(b - a)(b - a + 2)}{12}$$

et enfin $\sigma(X) = \sqrt{\dfrac{(b - a)(b - a + 2)}{12}}$.

❑

6.2 Loi de Bernoulli

6.2.1 Définition de la loi de Bernoulli

Dans ce paragraphe, on s'intéresse à un événement A donné dans une certaine expérience aléatoire. Pour cet événement, on a que deux possibilités : il est réalisé ou pas. La réalisation de cet événement est appelée *succès* et la non réalisation de cet événement est appelée *échec*. On code le succès et l'échec par les nombres 1 et 0 ou encore on considère la variable X égale à 1 si A est réalisé et à 0 si A n'est pas réalisé. L'intérêt d'avoir choisi les nombres 0 et 1 pour coder réside dans le fait qu'en recommençant l'expérience la somme de ces 0 et 1 donnera le nombre de succès.

On est dans cette situation quand on lance une pièce de monnaie une fois et que l'on marque 1 si on a obtenu Pile et 0 si on a obtenu Face. On est aussi dans cette situation quand on lance un dé à 6 faces et que l'on marque 1 si on a obtenu un 6 et 0 si on n'a pas obtenu un 6.

Définition 18. Soit p un réel élément de $[0,1]$. Soit X une variable aléatoire réelle.

X suit la **loi de Bernoulli de paramètre** p si et seulement si $X(\Omega)=\{0,1\}$ et $P(X=1)=p$ (et $P(X=0)=1-p$).

Notation. Quand une variable aléatoire X suit une loi de Bernoulli de paramètre p, on écrit $X\sim\mathcal{B}(p)$.

La notion de variable aléatoire suivant une loi de Bernoulli modélise toute expérience aléatoire qui a exactement deux issues. Par exemple, la variable indicatrice d'un événement A donné suit une loi de Bernoulli.

Remarque. Les nombres 0 et 1 ont un autre avantage : si X et Y sont deux variables aléatoires suivant une loi de Bernoulli (de paramètres éventuellement différents), alors XY suit une loi de Bernoulli et X^2 ou plus généralement X^k, $k\in\mathbb{N}^*$, suivent une loi de Bernoulli.

6.2.2 Espérance, variance et écart-type de la loi de Bernoulli

Théorème 33. Soit $p\in[0,1]$. Soit X une variable aléatoire réelle suivant la loi de Bernoulli de paramètre p ($X\sim\mathcal{B}(p)$).

Alors, $E(X)=p$, $V(X)=p(1-p)$ et $\sigma(X)=\sqrt{p(1-p)}$.

Démonstration . $X(\Omega)=\{0,1\}$ puis $P(X=1)=p$ et $P(X=0)=1-p$.

$$E(X)=0\times P(X=0)+1\times P(X=1)=p.$$

Ensuite,

$$V(X)=E\left(X^2\right)-(E(X))^2=0^2\times P(X=0)+1^1\times P(X=1)-p^2=p-p^2=p(1-p),$$

et enfin

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

❏

Exercice 10. Quelle la valeur maximale de l'écart-type d'une variable aléatoire suivant une loi de Bernoulli ?

Solution 10. Soit X une variable aléatoire suivant une loi de Bernoulli de paramètre $p\in[0,1]$.

$$(\sigma(X))^2=V(X)=p(1-p)=\frac{1}{4}-\left(\frac{1}{2}-p\right)^2\leqslant\frac{1}{4}$$

avec égalité si et seulement si $p=\dfrac{1}{2}$. La valeur maximale de l'écart-type est donc $\sqrt{\dfrac{1}{4}}=\dfrac{1}{2}$, valeur obtenue si et seulement si $p=\dfrac{1}{2}$.

Exercice 11. Soient X et Y deux variables aléatoires suivant des lois de Bernoulli de paramètres respectifs p et p'.

Montrer que les variables X et Y sont indépendantes si et seulement si $\mathrm{cov}(X,Y)=0$.

Solution 11. On sait déjà que si X et Y sont indépendantes, alors $\mathrm{cov}(X,Y)=0$.

Réciproquement, supposons $\mathrm{cov}(X,Y)=0$. Ceci fournit $E(XY)-E(X)E(Y)=0$ puis $E(XY)=pp'$. Maintenant, $XY(\Omega)=\{0,1\}$ puis $XY=1\Leftrightarrow X=1$ et $Y=1$. Par suite,

$$P(\{X=1\}\cap\{Y=1\}) = P(XY=1) = 0\times P(XY=0)+1\times P(XY=1) = E(XY) = pp' = P(X=1)\times P(Y=1).$$

Les événements $\{X=1\}$ et $\{Y=1\}$ sont indépendants. On sait qu'il en est de même des événements $\{X=0\}=\overline{\{X=1\}}$ et $\{Y=1\}$, des événements $\{X=1\}$ et $\{Y=0\}$ et des événements $\{X=0\}$ et $\{Y=0\}$. On a montré que $(x,y)\in X(\Omega)\times Y(\Omega) = \{0,1\}^2$, $P(\{X=x\}\cap\{Y=y\}) = P(X=x)\times P(Y=y)$ et finalement, les variables X et Y sont indépendantes.

6.3 Loi binomiale

6.3.1 Modélisation de n expériences aléatoires indépendantes par un n-uplet de variables aléatoires indépendantes

On peut se passer de la lecture de ce paragraphe si on veut « aller au fait ». Le but de ce paragraphe est de convaincre qu'il existe une modélisation, ou plutôt un espace probabilisé, représentant la répétition d'une même épreuve de Bernoulli, de « manière indépendante ». Nous sommes dans cette situation quand par exemple, nous lançons 10 fois un dé avec pour possibilités à chaque lancer, obtenir ou ne pas obtenir le 6, la probabilité d'obtenir ou de ne pas obtenir le 6 à un lancer ne dépendant pas de ce qui se passe aux autres lancers.

• Analysons d'abord une situation générale. On veut modéliser une succession d'expériences aléatoires où la probabilité d'obtenir tel ou tel résultat au cours de l'une des expériences ne dépend pas de ce qui se passe au cours des autres expériences ou encore on veut que la notion intuitive d'indépendance de différentes expériences se traduise par l'indépendance au sens des probabilités.

On considère un certain nombre d'expériences aléatoires $\mathcal{E}_1$, $\mathcal{E}_2$, ..., $\mathcal{E}_n$. A chaque expérience $\mathcal{E}_i$, $1\leqslant i\leqslant n$, est associé un espace probabilisé, fini en maths sup, (Ω_i, P_i).

L'univers à considérer, associé à la succession d'expériences $\mathcal{E}_1$, $\mathcal{E}_2$, ..., $\mathcal{E}_n$, est clair : il s'agit de $\Omega=\Omega_1\times\ldots\times\Omega_n$. C'est la probabilité P sur Ω qui ne l'est pas encore.

On définit la probabilité P par :

$$\forall A=(A_1,\ldots,A_n)\in\Omega=\Omega_1\times\ldots\times\Omega_n,\ P(A)=P_1(A_1)\times\ldots\times P_n(A_n)\quad(*),$$

(définition de la « probabilité produit »). En particulier,

$$\forall i\in[\![1,n]\!],\ \forall A=(A_1,\ldots,A_n)\in\Omega,\ P_i(A_i)=P(\Omega_1\times\ldots\times\Omega_{i-1}\times A_i\times\Omega_{i+1}\times\ldots\times\Omega_n)\quad(**).$$

On peut démontrer facilement que P ainsi définie est une probabilité sur l'univers Ω. Les égalités $(*)$ et $(**)$ traduisent mathématiquement la notion intuitive d'« indépendance » d'une expérience $\mathcal{E}_i$ avec les autres expériences.

• On se place maintenant dans la situation où les expériences $\mathcal{E}_i$ sont identiques et toujours indépendantes. Les modélisations, à savoir les espaces probabilisés (Ω_i,P_i), sont un seul et même espace probabilisé (Ω,P) (Ω ne désigne donc plus le produit des Ω_i mais désigne chaque Ω_i et P n'est plus la probabilité produit mais P est chaque P_i).

L'espace probabilisé modélisant la répétition de manière indépendante de l'expérience $\mathcal{E}$ est (Ω',P') où $\Omega'=\Omega^n$ et P' est la probabilité définie par :

$$\forall A=(A_1,\ldots,A_n)\in\Omega^n,\ P'(A)=\prod_{i=1}^{n}P(A_i).$$

Tout ceci montre l'existence d'un espace probabilisé modélisant la répétition d'une même expérience à deux issues, et ceci de manière indépendante.

6.3.2 Définition de la loi binomiale

On considère une même expérience $\mathcal{E}$ à deux issues : le succès de probabilité $p\in[0,1]$ et l'échec de probabilité $1-p$.
On effectue de manière indépendante cette même expérience n fois, $n\in\mathbb{N}^*$. On note X la variable aléatoire égale au nombre de succès. Donc, $X(\Omega)=[\![0,n]\!]$.

Soit $k\in[\![0,n]\!]$. Puisque les expériences sont indépendantes, la probabilité d'obtenir k succès au cours des k premières expériences puis $n-k$ échecs au cours des $n-k$ dernières expériences est

$$\underbrace{p\times\ldots\times p}_{\text{p facteurs}}\times\underbrace{(1-p)\times\ldots\times(1-p)}_{n-\text{p facteurs}}=p^k(1-p)^{n-k}.$$

Maintenant, il y a $\binom{k}{n}$ successions de n résultats comportant k succès et $n-k$ échecs (nombre de choix de k emplacements parmi n), tous de probabilité $p^k(1-p)^{n-k}$. Donc,

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

Ceci nous conduit à la définition suivante :

Définition 19. Soient $n \in \mathbb{N}^*$ et $p \in [0,1]$. Soit X une variable aléatoire réelle.

X suit la loi binomiale de paramètres n et p si et seulement si

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

Notation. Quand une variable aléatoire suit une loi binomiale de paramètres n et p, on écrit $X \sim \mathcal{B}(n,p)$.

Remarque. $\displaystyle\sum_{k=0}^{n} P(X=k)=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=(p+1-p)^n=1$ (heureusement). Ceci, entre autres, met en évidence le lien existant entre la loi binomiale et la formule du binôme de Newton et explique le nom donnée à cette loi.

6.3.3 Espérance, variance et écart-type de la loi binomiale (premier calcul)

Théorème 34. Soient $n \in \mathbb{N}^*$ et $p \in [0,1]$. Soit X une variable aléatoire réelle suivant la loi binomiale de paramètres n et p ($X \sim \mathcal{B}(n,p)$).

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

Démonstration. Soit $n \geqslant 2$ (le calcul ayant déjà été fait quand $n=1$). $X(\Omega)=[\![0,n]\!]$ puis, $\forall k \in [\![0,n]\!]$, $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$. On sait que pour tout $k \in [\![1,n]\!]$, $k\binom{n}{k}=n\binom{n-1}{k-1}$ et donc

$$\begin{aligned}
E(X)&=\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=1}^{n}k\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=1}^{n}n\binom{n-1}{k-1}p^k(1-p)^{n-k}\\
&=\sum_{k'=0}^{n}n\binom{n-1}{k'}p^{k'+1}(1-p)^{n-(k'+1)} \text{ (en posant } k'=k-1 \text{ ou encore } k=k'+1)\\
&=np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-1-k}=np(p+1-p)^{n-1}=np.
\end{aligned}$$

Ensuite, pour $n \geqslant 2$ puis $1 \leqslant k \leqslant n-1$, $k(n-k)\binom{n}{k}=k(n-k)\dfrac{n!}{k!(n-k)!}=n(n-1)\dfrac{(n-2)!}{(k-1)!(n-1-k)!}=n(n-1)\binom{n-2}{k-1}$ et donc

$$\begin{aligned}
V(X)=E\left(X^2\right)-(E(X))^2&=\sum_{k=1}^{n}k^2\binom{n}{k}p^k(1-p)^{n-k}-n^2p^2=\sum_{k=1}^{n}\left(k^2-kn+kn\right)\binom{n}{k}p^k(1-p)^{n-k}-n^2p^2\\
&=\sum_{k=1}^{n}k(k-n)\binom{n}{k}p^k(1-p)^{n-k}+n\sum_{k=1}^{n}k\binom{n}{k}p^k(1-p)^{n-k}-n^2p^2\\
&=-n(n-1)\sum_{k=1}^{n-1}\binom{n-2}{k-1}p^k(1-p)^{n-k}+n\times np-n^2p^2\\
&=-n(n-1)\sum_{k=0}^{n-2}\binom{n-2}{k}p^{k+1}(1-p)^{n-(k+1)}+n^2p(1-p)\\
&=-n(n-1)p(1-p)\sum_{k=0}^{n-2}\binom{n-2}{k}p^k(1-p)^{n-2-k}+n^2p(1-p)\\
&=-n(n-1)p(1-p)(p+1-p)^{n-2}+n^2p(1-p)=-n(n-1)p(1-p)+n^2p(1-p)\\
&=np(1-p),
\end{aligned}$$

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

❏

6.3.4 Somme de n variables indépendantes suivant une même loi de Bernoulli

Théorème 35. Soit $(n, m) \in (\mathbb{N}^*)$ et $p \in [0, 1]$. Soient X et Y deux variable aléatoires réelles indépendantes telles que $X \sim \mathcal{B}(n, p)$ et $Y \sim \mathcal{B}(m, p)$.

Alors, $X + Y$ suit la loi binomiale $\mathcal{B}(n + m, p)$.

Démonstration . Tout d'abord, $X(\Omega) = [\![0, n]\!]$ et $Y(\Omega) = [\![0, m]\!]$. Donc, $(X+Y)(\Omega) = [\![0, n+m]\!]$. Ensuite, pour $k \in [\![0, n+m]\!]$,

$$\begin{aligned}
P(X+Y=k) &= \sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} P(\{X=i\}\cap\{Y=j\}) \\
&= \sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} P(\{X=i\})\times P(\{Y=j\}) \text{ (car les variables } X \text{ et } Y \text{ sont indépendantes)} \\
&= \sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{j} p^j (1-p)^{m-j} \\
&= \sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} \binom{n}{i}\binom{m}{j} p^{i+j}(1-p)^{n+m-(i+j)} = p^k(1-p)^{n+m-k} \sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} \binom{n}{i}\binom{m}{j}
\end{aligned}$$

Ensuite, $(1+\alpha)^n = \displaystyle\sum_{i=0}^{n} \binom{n}{i}\alpha^i$ et $(1+\alpha)^m = \displaystyle\sum_{j=0}^{m} \binom{m}{j}\alpha^j$ puis

$$(1+\alpha)^{n+m} = (1+\alpha)^n(1+\alpha)^m = \left(\sum_{i=0}^{n}\binom{n}{i}\alpha^i\right)\left(\sum_{j=0}^{m}\binom{m}{j}\alpha^j\right).$$

Pour $0 \leqslant k \leqslant n+m$, $\displaystyle\sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} \binom{n}{i}\binom{m}{j}$ est le coefficient de α^k dans le développement du produit précédent. Ce coefficient est aussi le coefficient de α^k dans le développement de $(1+\alpha)^{n+m}$ par la formule du binôme du Newton. Ceci montre que pour $0 \leqslant k \leqslant n+m$,

$$\sum_{\substack{(i,j)\in[\![0,n]\!]\times[\![0,m]\!]\\ i+j=k}} \binom{n}{i}\binom{m}{j} = \binom{n+m}{k} \text{ (formule de Vandermonde)}$$

et donc

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

La variable $X + Y$ suit donc la loi binomiale $\mathcal{B}(n + m, p)$.

❏

Par récurrence, on obtient

Théorème 36. Soit $p \in [0, 1]$. Soient $n \geqslant 2$ puis $X_1, \ldots, X_n$, n variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre p. Alors

$$X_1 + \ldots + X_n \sim \mathcal{B}(n, p).$$

Démonstration .

- Si X_1 et X_2 suivent $\mathcal{B}(p) = \mathcal{B}(1, p)$ et sont indépendantes, alors $X_1 + X_2$ suit la loi $\mathcal{B}(1 + 1, p) = \mathcal{B}(2, p)$ d'après le théorème précédent.

- Soit $n \geqslant 2$. Supposons le résultat pour n variables. Soient $X_1, \ldots, X_n, X_{n+1}$, $n + 1$ variables indépendantes suivant la loi de Bernoulli de paramètre p. Par hypothèse de récurrence, $X_1 + \ldots + X_n$ suit $\mathcal{B}(n, p)$. Ensuite, d'après les lemme des coalitions, les variables $X_1 + \ldots + X_n$ et X_{n+1} sont indépendantes.
  Toujours d'après le théorème précédent, $X_1 + \ldots + X_n + X_{n+1} = (X_1 + \ldots + X_n) + X_{n+1}$ suit la loi binomiale $\mathcal{B}(n + 1, p)$.

Le résultat est démontré par récurrence.

❏

Retrouvons alors l'espérance et la variance d'une variable aléatoire X suivant la loi binomiale $\mathcal{B}(n,p)$ (2ème calcul). D'après le théorème 36, X peut s'écrire sous la forme $X = X_1 + \ldots + X_n$ où les variables X_i sont des variables mutuellement indépendantes suivant une même loi de Bernoulli de paramètre p. Pour tout $i \in [\![1,n]\!]$, $E(X_i) = p$ et $V(X_i) = p(1-p)$.

Par linéarité de l'espérance,

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np.$$

D'autre part, les variables X_i sont indépendantes et en particulier, les variables X_i sont deux à deux indépendantes. On en déduit que

$$V(X) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p).$$