23 Variables aléatoires sur un univers fini

23.1 Généralités

23.1.1 Variable aléatoire

Définition 1: Variable aléatoire sur un univers fini

Soit Ω un univers fini.
Une variable aléatoire sur Ω est une application X définie sur l'univers Ω et à valeurs dans un ensemble E.

Remarque 1. Si Ω est fini, $X(\Omega)$ est nécessairement fini également.

Exemple 1. • Soit Ω un univers fini. Soit $A \subset \Omega$.

La fonction indicatrice $\begin{array}{rcl} \mathbb{1}_A : \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \mathbb{1}_A(\omega) = \left\{ \begin{array}{ll} 1 & \text{si}\, \omega \in A \\ 0 & \text{si}, \omega \notin A \end{array} \right. \end{array}$ est une variable aléatoire.

Rappelons ses propriétés vues dans le TD « Applications » :

$$\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A; \quad \mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B; \quad \mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B.$$

• On lance deux dés à 6 faces. On considère l'univers $\Omega = [\![1, 6]\!]^2$.

Soit $X : \begin{array}{rcl} \Omega & \longrightarrow & \mathbb{R} \\ (a, b) & \longmapsto & a + b. \end{array}$

L'application X est une variable aléatoire qui renvoie la somme des deux dés. On a alors $X(\Omega) = [\![2, 12]\!]$.

Définition 2

Soit X une variable aléatoire définie sur un univers fini Ω et à valeurs dans E.
Soit $A \subset E$.
On note $(X \in A)$ l'événement

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

En particulier, si $E = \mathbb{R}$, pour tout réel x, on note
• $(X = x)$ l'événement $(X \in \{x\}) = \{\omega \in \Omega, X(\omega) \in \{x\}\}$.
• $(X \leqslant x)$ l'événement $(X \in]-\infty, x]) = \{\omega \in \Omega, X(\omega) \in]-\infty, x]\}$.
• $(X < x)$ l'événement $(X \in]-\infty, x[) = \{\omega \in \Omega, X(\omega) \in]-\infty, x[\}$.
• $(X \geqslant x)$ l'événement $(X \in [x + \infty[) = \{\omega \in \Omega, X(\omega) \in [x + \infty[\}$.
• $(X > x)$ l'événement $(X \in]x + \infty[) = \{\omega \in \Omega, X(\omega) \in]x + \infty[\}$.
et pour tout couple (x, y) de réels avec $x \leqslant y$, on note

$$(x \leqslant X \leqslant y) = \{\omega \in \Omega, X(\omega) \in [x, y]\}.$$

Exemple 2. Reprenons l'exemple précédent. On a

$$(X = 7) = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\},$$

$$(X < 13) = \Omega$$

et

$$(10 \leqslant X \leqslant 12) = \{(4,6), (5,5), (6,4), (5,6), (6,5), (6,6)\}.$$

Proposition 1: Système complet d'événements associé à une variable aléatoire

Soit X une variable aléatoire définie sur un univers fini Ω. Notons $X(\Omega) = \{x_1, \ldots, x_n\}$ (où $x_i \neq x_j$ si $i \neq j$) les valeurs prises par la variable aléatoire X (qui sont nécessairement en nombre fini).
Alors les événements $(X = x_i)_{1 \leqslant i \leqslant n}$ forment un système complet d'événements pour Ω.

Démonstration. • Soient $(i, j) \in [\![1, n]\!]^2$ avec $i \neq j$.
Soit $\omega \in (X = x_i) \cap (X = x_j)$.
Alors $X(\omega) = x_i$ et $X(\omega) = x_j$. Or $x_i \neq x_j$.
Il ne peut donc pas exister d'élément $\omega \in (X = x_i) \cap (X = x_j)$, ce qui prouve que

$$(X = x_i) \cap (X = x_j) = \emptyset$$

et ce pour tout $(i, j) \in [\![1, n]\!]^2$ avec $i \neq j$.

• Montrons que $\Omega = \displaystyle\bigcup_{i=1}^{n} (X = x_i)$.

On a clairement $\displaystyle\bigcup_{i=1}^{n} (X = x_i) \subset \Omega$. Montrons l'inclusion réciproque.

Soit $\omega \in \Omega$. Alors $X(\omega) \in X(\Omega) = \{x_1, \ldots, x_n\}$ donc il existe $i \in [\![1, n]\!]$ tel que $X(\omega) = x_i$, i.e. $\omega \in (X = x_i)$ donc $\omega \in \displaystyle\bigcup_{i=1}^{n} (X = x_i)$.

On a donc bien prouvé l'inclusion $\Omega \subset \displaystyle\bigcup_{i=1}^{n} (X = x_i)$ d'où finalement l'égalité

$$\Omega = \bigcup_{i=1}^{n} (X = x_i).$$

Les deux points ci-dessus montrent que les événements $(X = x_i)_{1\leqslant i\leqslant n}$ forment un système complet d'événements pour Ω. ■

Remarque 2. En particulier, ceci implique que $\displaystyle\sum_{i=1}^{n} \mathbb{P}(X = x_i) = 1$.

Exemple 3. Dans l'exemple précédent, les événements $(X = k)_{2\leqslant k\leqslant 12}$ forment un système complet d'événements pour Ω.

23.1.2 Loi d'une variable aléatoire

Définition 3: Loi d'une variable aléatoire

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit X une variable aléatoire sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs dans un ensemble E.
On appelle loi (de probabilité) de la variable aléatoire X l'application

$$\begin{array}{rcl} \mathbb{P}_X : \mathcal{P}(E) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbb{P}(X \in A). \end{array}$$

Cette application $\mathbb{P}_X$ est une probabilité sur $(E, \mathcal{P}(E))$.
Si Y est une autre variable aléatoire définie sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ telle que $\mathbb{P}_X = \mathbb{P}_Y$, on dit que X et Y ont même loi et on note $X \sim Y$.

Démonstration. Montrons que $\mathbb{P}_X$ est bien une probabilité sur $(E, \mathcal{P}(E))$.
• On a $\mathbb{P}_X(E) = \mathbb{P}(X \in E) = \mathbb{P}(\Omega) = 1$.
• Soient $(A_1, \dots, A_n)$ des événements de E deux à deux incompatibles. Par propriété de la probabilité $\mathbb{P}$, on a alors

$$\mathbb{P}_X\left(\bigsqcup_{k=1}^{n} A_k\right) = \mathbb{P}\left(X \in \bigsqcup_{k=1}^{n} A_k\right) = \mathbb{P}\left(\bigsqcup_{k=1}^{n} (X \in A_k)\right) = \sum_{k=1}^{n} \mathbb{P}(X \in A_k) = \sum_{k=1}^{n} \mathbb{P}_X(A_k)$$

donc $\mathbb{P}_X$ est bien une probabilité sur $(E, \mathcal{P}(E))$. ■

Remarque 3. La probabilité $\mathbb{P}_X$ est déterminée par la distribution de probabilités $(\mathbb{P}(X = x))_{x\in E}$.

Exemple 4. On lance un dé à 6 faces équilibré. Si on obtient 1 ou 2, on perd 1 point ; si on obtient $3, 4$ ou 5, il ne se passe rien. Si on obtient 6, on gagne 3 points.

Pour modéliser cette expérience, on pose $\Omega = [\![1, 6]\!]$.

Puisqu'on est face à une situation d'équiprobabilité, on définit la probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ par

$$\forall k \in [\![1, 6]\!], \mathbb{P}(\{k\}) = \frac{1}{6}.$$

Par ailleurs, on définit la variable aléatoire $X : [\![1, 6]\!] \to \{-1, 0, 3\}$ définie par

$$X(\omega) = \begin{cases} -1 & \text{si } \omega \in \{1, 2\} \\ 0 & \text{si } \omega \in \{3, 4, 5\} \\ 3 & \text{si } \omega = 6. \end{cases}$$

On a alors

$$\mathbb{P}_X(\{-1\}) = \mathbb{P}(X = -1) = \mathbb{P}(\{1, 2\}) = \frac{1}{3};$$

$$\mathbb{P}_X(\{0\}) = \mathbb{P}(X = 0) = \mathbb{P}(\{3, 4, 5\}) = \frac{1}{2};$$

et
$$\mathbb{P}_X(\{3\}) = \mathbb{P}(X=3) = \mathbb{P}(\{6\}) = \frac{1}{6}.$$

D'autre part, on a
$$\mathbb{P}(X \leqslant 0) = \mathbb{P}(X=-1) + \mathbb{P}(X=0) = \frac{5}{6}.$$
La probabilité d'avoir un gain négatif est donc grande, mais pourtant, on a tout intérêt à jouer à ce jeu (calculer l'espérance...).

Proposition 2

Soit $n \in \mathbb{N}^*$. On considère $(x_i)_{1\leqslant i\leqslant n}$ une famille de réels distincts et $(p_i)_{1\leqslant i\leqslant n}$ une famille de réels positifs vérifiant $\sum\limits_{i=1}^{n} p_i = 1$.
Alors il existe une variable aléatoire X sur un espace probabilisé fini vérifiant
$$\forall i \in [\![1,n]\!], \mathbb{P}(X = x_i) = p_i.$$

Démonstration. Soit $\Omega = [\![1,n]\!]$. On définit une application
$$\begin{array}{rcl} X : \Omega = [\![1,n]\!] & \longrightarrow & \mathbb{R} \\ i & \longmapsto & x_i \end{array}$$
et une application
$$\begin{array}{rcl} \mathbb{P} : \mathcal{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \sum\limits_{\substack{i\in[\![1,n]\!] \\ i\in A}} p_i \end{array}.$$
On a vu dans le chapitre précédent que $\mathbb{P}$ définit bien une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.

L'application X est alors une variable aléatoire sur l'espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ qui vérifie les conditions demandées puisque pour tout $i \in [\![1,n]\!]$, on a
$$\mathbb{P}(X = x_i) = \sum_{\substack{k\in[\![1,n]\!] \\ k\in(X=x_i)}} p_k = p_i.$$
$\blacksquare$

Remarque 4. Ce théorème nous permettra de définir les lois classiques : en effet, une loi sera donnée par l'ensemble des valeurs atteintes $\{x_i | 1 \leqslant i \leqslant n\}$ et les probabilités $\mathbb{P}(X = x_i)$.

Proposition 3

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit X une variable aléatoire sur Ω et à valeurs dans un ensemble E.
Soit f une fonction définie sur E à valeurs dans un ensemble F.
Alors $f \circ X$ est une variable aléatoire définie sur Ω à valeurs dans F, notée $f(X)$.
Si $Y : \Omega \to E$ est une autre variable aléatoire telle que $X \sim Y$, alors $f(X) \sim f(Y)$.

Démonstration. Pour tout $A \in \mathcal{P}(F)$, puisque $\mathbb{P}_X = \mathbb{P}_Y$, on a
$$\mathbb{P}(f(X) \in A) = \mathbb{P}(X \in f^{-1}(A)) = \mathbb{P}(Y \in f^{-1}(A)) = \mathbb{P}(f(Y) \in A)$$
donc $f(X) \sim f(Y)$. $\blacksquare$

Définition 4: Loi conditionnelle d'une variable aléatoire sachant un événement

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé.
Soit X une variable aléatoire sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs dans un ensemble E.
Soit A un événement inclus dans Ω tel que $\mathbb{P}(A) \neq 0$.
On appelle loi conditionnelle de la variable aléatoire X sachant l'événement A la loi

$$\begin{array}{rcl} \mathcal{P}(E) & \longrightarrow & [0,1] \\ B & \longmapsto & \mathbb{P}_A(X \in B) = \dfrac{\mathbb{P}(A \cap (X \in B))}{\mathbb{P}(A)} \end{array}.$$

Remarque 5. Cette application est une probabilité sur $(E, \mathcal{P}(E))$.

23.1.3 Couple de variables aléatoires

Définition 5: Couple de variables aléatoires

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soient X et Y deux variables aléatoires définies sur Ω à valeurs dans E et F respectivement.
On appelle couple de variables aléatoires défini par X et Y la variable aléatoire

$$\begin{array}{rcl} Z = (X,Y) : \Omega & \longrightarrow & E \times F \\ \omega & \longmapsto & (X(\omega), Y(\omega)). \end{array}$$

On appelle loi conjointe de X et Y la loi du couple (X,Y) qui est donnée par l'application

$$\begin{array}{rcl} \mathbb{P}_{(X,Y)} : E \times F & \longrightarrow & [0,1] \\ (x,y) & \longmapsto & \mathbb{P}((X=x) \cap (Y=y)) = \mathbb{P}(X=x, Y=y). \end{array}$$

Remarque 6. • L'événement $(X = x) \cap (Y = y)$ se note $(X = x, Y = y)$. Plus généralement, pour tout $(A, B) \in \mathcal{P}(E) \times \mathcal{P}(F)$, l'événement $(X \in A) \cap (Y \in B)$ se note $(X \in A, Y \in B)$.

• La connaissance de la distribution de probabilités $(\mathbb{P}(X = x, Y = y))_{(x,y)\in E\times F}$ entraîne la connaissance de la distribution de probabilités $(\mathbb{P}(X \in A, Y \in B))_{(A,B)\in\mathcal{P}(E)\times\mathcal{P}(F)}$ car pour tout $(A, B) \in \mathcal{P}(E) \times \mathcal{P}(F)$, on a

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}\left(\bigsqcup_{x\in A}(X = x), \bigsqcup_{y\in B}(Y = y)\right) = \sum_{x\in A}\sum_{y\in B}\mathbb{P}(X = x, Y = y).$$

• On peut généraliser cette définition à n variables aléatoires $(X_1, \ldots, X_n)$.

Définition 6: Lois marginales

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soient X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{P}(\Omega))$ à valeurs dans E et F respectivement.
On appelle première loi marginale du couple (X, Y) la loi de la variable aléatoire X dans le sens suivant :

$$\begin{array}{rcl} \mathbb{P}_X : \mathcal{P}(E) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(X \in A, Y \in F). \end{array}$$

On appelle deuxième loi marginale du couple (X, Y) la loi de la variable aléatoire Y dans le sens suivant :

$$\begin{array}{rcl} \mathbb{P}_Y : \mathcal{P}(F) & \longrightarrow & [0,1] \\ B & \longmapsto & \mathbb{P}_Y(B) = \mathbb{P}(Y \in B) = \mathbb{P}(X \in E, Y \in B). \end{array}$$

Remarque 7. On voit sur la définition que la connaissance de la loi conjointe du couple implique la connaissance des lois marginales.

En revanche, la connaissance des lois marginales n'implique pas la connaissance de la loi du couple. Pour s'en convaincre, regardons l'exemple suivant.

Alice et Bob passent un examen et répondent à une question où deux réponses sont proposées. On note X la variable aléatoire qui à la réponse d'Alice associe 0 si elle est fausse, 1 si elle est correcte. On considère Y la variable aléatoire similaire pour Bob.

Intéressons-nous aux trois situations suivantes :

1. Alice et Bob répondent au hasard et indépendamment.
2. Alice répond au hasard et Bob copie sur Alice.
3. Alice répond au hasard et Bob décide systématiquement de prendre la réponse opposée.

   Dans les trois cas, les lois marginales sont les mêmes.

   En effet, on a

   $$\mathbb{P}_X(\{0\}) = \mathbb{P}_X(\{1\}) = \mathbb{P}_Y(\{0\}) = \mathbb{P}_Y(\{1\}) = \frac{1}{2}.$$

   En revanche, les lois du couple (X, Y) sont différentes dans ces trois situations.

   Dans la première situation, on a

   $$\mathbb{P}_{(X,Y)}(\{(0,0)\}) = \frac{1}{4};$$

   dans la deuxième,

   $$\mathbb{P}_{(X,Y)}(\{(0,0)\}) = \frac{1}{2};$$

   et dans la troisième,

   $$\mathbb{P}_{(X,Y)}(\{(0,0)\}) = 0.$$

23.2 Espérance et variance

23.2.1 Espérance

Définition 7: Espérance d'une variable aléatoire

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit X une variable aléatoire définie sur $(\Omega, \mathcal{P}(\Omega))$.
On note $X(\Omega) = \{x_1, \dots, x_n\}$, avec $x_i \neq x_j$ si $i \neq j$.
On appelle espérance de X, noté $\mathbb{E}(X)$, le réel

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

On dit que la variable aléatoire X est centrée si $\mathbb{E}(X) = 0$.

Remarque 8. • On note également $\mathbb{E}(X) = \displaystyle\sum_{x \in X(\Omega)} x\mathbb{P}(X = x)$.

• L'espérance d'une variable aléatoire représente la moyenne des valeurs prises par cette variable aléatoire pondérées par leurs probabilités. C'est un indicateur dit de position.

Exemple 5. Si la variable aléatoire X est constante, i.e. il existe $a \in \mathbb{R}$ tel que $\mathbb{P}(X = a) = 1$, alors $\mathbb{E}(X) = a$.

Exemple 6. Reprenons l'exemple 4.
On a $\mathbb{E}(X) = -1 \times \mathbb{P}(X = -1) + 0 \times \mathbb{P}(X = 0) + 3\mathbb{P}(X = 3) = -\dfrac{1}{3} + 3 \times \dfrac{1}{6} = \dfrac{1}{6}$.
L'espérance est positive donc le jeu est favorable au joueur.

Proposition 4

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit X une variable aléatoire définie sur $(\Omega, \mathcal{P}(\Omega))$.
Alors
$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}).$$

Démonstration. Notons $X(\Omega) = \{x_1, \dots, x_n\}$ avec $x_i \neq x_j$ si $i \neq j$.

Pour tout $k \in [\![1, n]\!]$, notons $A_k = (X = x_k) = \{\omega \in \Omega, X(\omega) = x_k\}$. Notons que $\Omega = \displaystyle\bigsqcup_{k=1}^{n} A_k$.

On a alors

$$\sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) = \sum_{k=1}^{n} \sum_{\omega \in A_k} \underbrace{X(\omega)}_{=x_k} \mathbb{P}(\{\omega\}) = \sum_{k=1}^{n} x_k \sum_{\omega \in A_k} \mathbb{P}(\{\omega\}) = \sum_{k=1}^{n} x_k \mathbb{P}\left(\underbrace{\bigsqcup_{\omega \in A_k} \{\omega\}}_{=(X=x_k)} \right) = \sum_{k=1}^{n} x_k \mathbb{P}(X = x_k)$$

d'où $\mathbb{E}(X) = \displaystyle\sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$. ∎

Exemple 7. Reprenons l'exemple 4. On a

$$\begin{aligned}
\mathbb{E}(X) &= X(1)\mathbb{P}(\{1\}) + X(2)\mathbb{P}(\{2\}) + X(3)\mathbb{P}(\{3\}) + X(4)\mathbb{P}(\{4\}) + X(5)\mathbb{P}(\{5\}) + X(6)\mathbb{P}(\{6\}) \\
&= \frac{1}{6}(-1 + (-1) + 0 + 0 + 0 + 3) \\
&= \frac{1}{6}.
\end{aligned}$$

Proposition 5: Espérance d'une indicatrice

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit $A \in \mathcal{P}(\Omega)$.
Alors

$$\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A).$$

Démonstration. On a $\mathbb{1}_A(\Omega) = \{0, 1\}$ et par définition de $\mathbb{1}_A$, on a les événements

$$(\mathbb{1}_A = 0) = \overline{A} \quad \text{et} \quad (\mathbb{1}_A = 1) = A$$

donc

$$\mathbb{E}(\mathbb{1}_A) = 0 \times \mathbb{P}(\mathbb{1}_A = 0) + 1 \times \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A).$$

■

Proposition 6: Positivité de l'espérance

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit X une variable aléatoire définie sur $(\Omega, \mathcal{P}(\Omega))$ à valeurs positives.
Alors $\mathbb{E}(X) \geqslant 0$.

Démonstration. Notons $X(\Omega) = \{x_1, \dots, x_n\}$. Par hypothèse, pour tout $k \in [\![1, n]\!], x_k \geqslant 0$ donc

$$\mathbb{E}(X) = \sum_{k=1}^{n} x_k \mathbb{P}(X = x_k) \geqslant 0.$$

■

Théorème 1: Formule de transfert

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit X une variable aléatoire définie sur $(\Omega, \mathcal{P}(\Omega))$ à valeurs dans un ensemble E. On note $X(\Omega) = \{x_1, \dots, x_n\}$ avec $x_i \neq x_j$ si $i \neq j$.
Soit $f : E \longrightarrow \mathbb{R}$ une application.
Alors la variable aléatoire $f \circ X$, notée $f(X)$, a pour espérance

$$\mathbb{E}(f(X)) = \sum_{k=1}^{n} f(x_k)\mathbb{P}(X = x_k).$$

Démonstration. Puisque $\Omega = \bigsqcup_{k=1}^{n} (X = x_k)$, on a

$$\begin{aligned} \mathbb{E}(f(X)) &= \sum_{\omega \in \Omega} f(X(\omega))\mathbb{P}(\{\omega\}) \\ &= \sum_{k=1}^{n} \sum_{\omega \in (X = x_k)} f(X(\omega))\mathbb{P}(\{\omega\}) \\ &= \sum_{k=1}^{n} f(x_k) \sum_{\omega \in (X = x_k)} \mathbb{P}(\{\omega\}) \\ &= \sum_{k=1}^{n} f(x_k)\mathbb{P}(X = x_k). \end{aligned}$$

■

Remarque 9. On peut également appliquer la formule de transfert aux couples ou aux n-uplets de variables aléatoires.

Exemple 8. Reprenons l'exemple des dés.

On a

$$\mathbb{E}(X^2) = \mathbb{P}(X = -1) + 9\mathbb{P}(X = 3) = \frac{1}{3} + \frac{9}{6} = \frac{11}{6}.$$

Proposition 7: Linéarité de l'espérance

Soient X et Y deux variables aléatoires définies sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs réelles.

Alors pour tout $(\lambda, \mu) \in \mathbb{R}^2$, l'espérance de la variable aléatoire $\lambda X + \mu Y$ vérifie

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

Démonstration. Par linéarité des sommes finies, on a

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

■

Remarque 10. Soit X une variable aléatoire définie sur un espace probabilisé fini. On considère la variable aléatoire $Y = X - \mathbb{E}(X)$.

Par linéarité de l'espérance, on a

$$\mathbb{E}(Y) = \mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0,$$

où on a utilisé que $\mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X)$ car $\mathbb{E}(X)$ est une variable aléatoire constante.

Ainsi, pour toute variable aléatoire X, la variable aléatoire $X - \mathbb{E}(X)$ est centrée.

Corollaire 1: Croissance de l'espérance

Soient X et Y deux variables aléatoires définies sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs réelles.
On suppose que $X \leqslant Y$, i.e. pour tout $\omega \in \Omega, X(\omega) \leqslant Y(\omega)$.
Alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

Démonstration. Par hypothèse, on a $Y - X \geqslant 0$. Par positivité de l'espérance, on a

$$\mathbb{E}(Y - X) \geqslant 0.$$

Or, par linéarité de l'espérance, on a $\mathbb{E}(Y - X) = \mathbb{E}(Y) - \mathbb{E}(X)$, donc $\mathbb{E}(Y) - \mathbb{E}(X) \geqslant 0$ d'où $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$. ■

Corollaire 2: Inégalités triangulaires

Soient X et Y deux variables aléatoires définies sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs réelles.
Alors

1. $$|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|).$$
2. $$\mathbb{E}(|X + Y|) \leqslant \mathbb{E}(|X|) + \mathbb{E}(|Y|).$$

Démonstration.

1. D'après l'inégalité triangulaire réelle, on a

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

2. D'après l'inégalité triangulaire réelle, on a $|X + Y| \leqslant |X| + |Y|$.
Par croissance et linéarité de l'espérance, on en déduit que

$$\mathbb{E}(|X + Y|) \leqslant \mathbb{E}(|X| + |Y|) = \mathbb{E}(|X|) + \mathbb{E}(|Y|).$$

■

23.2.2 Variance

Définition 8: Moments d'une variable aléatoire

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Notons $X(\Omega) = \{x_1, \ldots, x_n\}$, où $x_i \neq x_j$ si $i \neq j$.
Pour tout $r \in \mathbb{N}^*$, on appelle moment d'ordre r de la variable aléatoire X le nombre réel $\mathbb{E}(X^r)$, i.e. d'après le théorème de transfert :

$$\mathbb{E}(X^r) = \sum_{k=1}^{n} x_k^r \mathbb{P}(X = x_k).$$

Remarque 11. • Le moment d'ordre 1 de la variable aléatoire X est $\mathbb{E}(X)$.

• On appelle moment centré d'ordre r de X le moment d'ordre r de la variable aléatoire centrée $X - \mathbb{E}(X)$, i.e.

$$\mathbb{E}((X - \mathbb{E}(X))^r) = \sum_{k=1}^{n} (x_k - \mathbb{E}(X))^r \mathbb{P}(X = x_k).$$

Définition 9: Variance d'une variable aléatoire

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
On appelle variance de X et on note $V(X)$ le réel

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

Remarque 12. • La variance est le moment d'ordre 2 de la variable aléatoire centrée $X - \mathbb{E}(X)$.

• Puisque $(X - \mathbb{E}(X))^2 \geqslant 0$, on en déduit que $V(X) \geqslant 0$ par positivité de l'espérance.

• La variance mesure la moyenne des carrés des écarts entre les valeurs prises par la variable aléatoire X et sa moyenne (son espérance). C'est un indicateur dit de dispersion.

En pratique, on calcule la variance d'une variable aléatoire grâce à la formule suivante :

Proposition 8: Formule de König-Huygens

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
Alors

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

Démonstration. On a bien d'après la linéarité de l'espérance

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

(On a utilisé que si $X = a$ est une variable aléatoire constante, alors $\mathbb{E}(X) = a$). ■

Remarque 13. Puisque $V(X) \geqslant 0$, on en déduit que $\mathbb{E}(X)^2 \leqslant \mathbb{E}(X^2)$, i.e. $|\mathbb{E}(X)| \leqslant \sqrt{\mathbb{E}(X^2)}$.

Exemple 9. Reprenons l'exemple 4. On a calculé $\mathbb{E}(X) = \dfrac{1}{6}$.

En appliquant la formule de transfert, on a

$$\mathbb{E}(X^2) = (-1)^2 \times \mathbb{P}(X = -1) + 0^2 \times \mathbb{P}(X = 0) + 3^2\mathbb{P}(X = 3) = \frac{1}{3} + 9 \times \frac{1}{6} = \frac{11}{6}.$$

Ainsi, $V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \dfrac{11}{6} - \dfrac{1}{36} = \dfrac{65}{36}$.

Définition 10: Ecart-type

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
On appelle écart-type de X, et on note $\sigma(X)$, le réel

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

Remarque 14. • Cette définition est légitime puisque $V(X) \geqslant 0$.

• L'écart-type mesure l'écart moyen par rapport à l'espérance. C'est encore un indicateur de dispersion.

• Si la variable aléatoire X s'exprime dans une unité, $V(X)$ s'exprime dans le carré de cette unité donc $\sigma(X)$ s'exprime dans la même unité que X.

Exemple 10. Dans l'exemple précédent, $V(X) = \frac{\sqrt{65}}{6}$.

Proposition 9: Caractérisation des variables aléatoires de variance nulle

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
Alors
$$V(X) = 0 \Leftrightarrow \mathbb{P}(X = \mathbb{E}(X)) = 1.$$
Autrement dit, une variable aléatoire est de variance nulle si et seulement si elle est presque sûrement égale à son espérance.

Remarque 15. On dit qu'une égalité est vraie presque sûrement si elle est vraie sur un ensemble de probabilité 1.

Démonstration. Par définition de la variance, on a
$$V(X) = 0 \Leftrightarrow \mathbb{E}((X - \mathbb{E}(X))^2) = 0.$$
Notons $X(\Omega) = \{x_1, \dots, x_n\}$, où $x_i \neq x_j$ si $i \neq j$. On a alors d'après le théorème de transfert :
$$V(X) = 0 \Leftrightarrow \sum_{k=1}^{n} (x_k - \mathbb{E}(X))^2 \mathbb{P}(X = x_k) = 0.$$
Or, une somme de termes positifs est nulle si et seulement si tous les termes sont nuls donc on a
$$V(X) = 0 \Leftrightarrow \forall k \in [\![1, n]\!], (x_k - \mathbb{E}(X))^2 \mathbb{P}(X = x_k) = 0 \Leftrightarrow \forall k \in [\![1, n]\!], x_k = \mathbb{E}(X) \,\text{ou}\, \mathbb{P}(X = x_k) = 0.$$
Or, les événements $(X = x_k)_{1 \leqslant n \leqslant n}$ forment un système complet d'événements donc
$$\sum_{k=1}^{n} \mathbb{P}(X = x_k) = 1.$$
Nécessairement, il existe $i \in [\![1, n]\!]$ tel que $\mathbb{P}(X = x_i) \neq 0$ et dans ce cas $x_i = \mathbb{E}(X)$. On a alors pour tout $k \neq i, x_k \neq \mathbb{E}(X)$ donc $\mathbb{P}(X = x_k) = 0$.

Puisque $\sum_{k=1}^{n} \mathbb{P}(X = x_k) = 1$, nécessairement $\mathbb{P}(X = x_i) = 1$. On a donc bien l'équivalence
$$V(X) = 0 \Leftrightarrow \mathbb{P}(X = \mathbb{E}(X)) = 1.$$
■

Exemple 11. Soit X une variable aléatoire telle que $\mathbb{P}(X = -1) = \mathbb{P}(X = 1) = \frac{1}{2}$ et soit Y la variable aléatoire nulle.

Ces deux variables aléatoires sont d'espérance nulle mais leurs variances diffèrent.

En effet, on a $V(X) = \mathbb{E}(X^2) = 1 \times \mathbb{P}(X = -1) + 1 \times \mathbb{P}(X = 1) = 1$ et $V(Y) = 0$.

La variance de X ne peut pas être nulle puisque X n'est pas presque sûrement égale à son espérance, contrairement à Y.

Proposition 10

Soit X une variable aléatoire réelle définie sur un espace probabilisé fini.
Alors pour tout $(a, b) \in \mathbb{R}^2$, on a
$$V(aX + b) = a^2 V(X).$$

Démonstration. C'est un simple calcul. On utilise la linéarité de l'espérance :

$$\begin{aligned}
V(aX+b) &= \mathbb{E}((aX+b)^2) - (\mathbb{E}(aX+b))^2 \\
&= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - (a\mathbb{E}(X)+b)^2 \\
&= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - a^2\mathbb{E}(X)^2 - 2ab\mathbb{E}(X) - b^2 \\
&= a^2(\mathbb{E}(X^2) - \mathbb{E}(X)^2) \\
&= a^2V(X).
\end{aligned}$$

■

Remarque 16.
1. En particulier, on a $\sigma(aX+b) = |a|\sigma(X)$.
   Une translation par b ne change pas l'écart par rapport à la moyenne (c'est à dire l'espérance), alors que multiplier la variable aléatoire par a multiplie par $|a|$ également cet écart.
   Ainsi, si $\sigma(X) > 0$, la variable aléatoire $Y = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ vérifie $\mathbb{E}(Y) = 0$ et $V(Y) = 1$: on dit qu'elle est centrée réduite.
2. On en déduit que la variance, contrairement à l'espérance, n'est pas linéaire.
   En effet, $V(2X) = 4V(X) \neq 2V(X)$ si $V(X) \neq 0$.
   Ainsi, on n'a pas forcément $V(X+Y) = V(X)+V(Y)$ (en l'occurence, c'est faux si $Y = X$ et $V(X) \neq 0$).

23.2.3 Covariance de deux variables aléatoires

Définition 11: Covariance

Soient X et Y deux variables aléatoires réelles définies sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ fini.
On appelle covariance de X et Y le réel

$$\mathrm{Cov(X,Y)} = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Si $\mathrm{Cov(X,Y)} = 0$, on dit que les variables aléatoires X et Y sont décorrélées.

Remarque 17.
1. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
2. $\mathrm{Cov}(X,X) = V(X)$.

Proposition 11: Variance d'une somme de variables aléatoires

Soient $X_1, \dots, X_n$ des variables aléatoires sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ fini.
Alors

$$V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k) + 2\sum_{1 \leq i < j \leq n} \mathrm{Cov}(X_i, X_j).$$

En particulier, si les variables aléatoires $(X_1, \dots, X_n)$ sont deux à deux décorrélées, on a

$$V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k).$$

Démonstration. On a
$$\left(\sum_{k=1}^{n} X_k\right)^2 = \sum_{k=1}^{n} X_k^2 + 2 \sum_{1\leq i<j\leq n} X_i X_j.$$
Par linéarité de l'espérance, on a alors
$$\mathbb{E}\left(\left(\sum_{k=1}^{n} X_k\right)^2\right) = \sum_{k=1}^{n} \mathbb{E}(X_k^2) + 2 \sum_{1\leq i<j\leq n} \mathbb{E}(X_i X_j).$$
D'autre part,
$$\mathbb{E}\left(\sum_{k=1}^{n} X_k\right)^2 = \left(\sum_{k=1}^{n} \mathbb{E}(X_k)\right)^2 = \sum_{k=1}^{n} \mathbb{E}(X_k)^2 + 2 \sum_{1\leq i<j\leq n} \mathbb{E}(X_i)\mathbb{E}(X_j).$$
Finalement, on a
$$\begin{aligned} V\left(\sum_{k=1}^{n} X_k\right) &= \mathbb{E}\left(\left(\sum_{k=1}^{n} X_k\right)^2\right) - \mathbb{E}\left(\sum_{k=1}^{n} X_k\right)^2 \\ &= \sum_{k=1}^{n} \mathbb{E}(X_k^2) - \mathbb{E}(X_k)^2 + 2 \sum_{1\leq i<j\leq n} \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) \\ &= \sum_{k=1}^{n} V(X_k) + 2 \sum_{1\leq i<j\leq n} \mathrm{Cov}(X_i, X_j). \end{aligned}$$
■

23.3 Indépendance de variables aléatoires

23.3.1 Variables aléatoires indépendantes

Définition 12: Variables aléatoires indépendantes

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
On dit que les variables aléatoires X et Y sont indépendantes, et on note $X \perp\!\!\!\perp Y$, si pour tout $A \in \mathcal{P}(X(\Omega))$ et pour tout $B \in \mathcal{P}(Y(\Omega))$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants, i.e.
$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

Proposition 12: Variables aléatoires indépendantes

Deux variables aléatoires X et Y définies sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ sont indépendantes si et seulement si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,
$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Démonstration. • Le sens direct est clair : il suffit d'appliquer la définition d'indépendace des variables aléatoires X et Y aux événements $(X = x)$ et $(Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

• Réciproquement, supposons que pour tout $(x, y) \in X(\Omega) \times Y(\Omega), \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$. Montrons que les vaiables aléatoires X et Y sont indépendantes.

Soient A et B des parties de $X(\Omega)$ et $Y(\Omega)$ respectivement. On a alors

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

ce qui prouve l'équivalence souhaitée. ∎

Remarque 18. • Les variables aléatoires X et Y sont indépendantes si et seulement si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants.

• Ceci signifie que les variables aléatoires X et Y sont indépendantes si et seulement si la loi conjointe du couple (X, Y) est donnée par la distribution de probabilités

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

Exemple 12. Alice et Bob passent un examen et répondent à une question où deux réponses sont proposées. On note X la variable aléatoire qui à la réponse d'Alice associe 0 si elle est fausse, 1 si elle est correcte. On considère Y la variable aléatoire similaire pour Bob.

Intéressons-nous aux trois situations suivantes :

1. Alice et Bob répondent au hasard et indépendamment.

2. Alice répond au hasard et Bob copie sur Alice.

3. Alice répond au hasard et Bob décide systématiquement de prendre la réponse opposée.

Dans la première des situations, les variables aléatoires X et Y sont indépendantes, mais pas dans les deux suivantes.

En effet, dans les trois cas, on a $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \dfrac{1}{2}$.

Dans la deuxième situation, $\mathbb{P}((X = 0) \cap (Y = 0)) = \mathbb{P}(X = 0) = \dfrac{1}{2} \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$.

Dans la troisième situation, $\mathbb{P}((X = 0) \cap (Y = 0)) = 0 \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$.

Définition 13: Variables aléatoires mutuellement indépendantes

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit $(X_k)_{1\leq k\leq n}$ une famille finie de variables aléatoires définies sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

1. On dit que les variables aléatoires $(X_k)_{1\leqslant k\leqslant n}$ sont deux à deux indépendantes si pour tout $(i,j) \in [\![1,n]\!]^2$ avec $i \neq j$, les variables aléatoires X_i et X_j sont indépendantes.
2. On dit que les variables aléatoires sont mutuellement indépendantes si pour tout $k \in [\![1,n]\!]$ et tout $(x_1, \ldots, x_k) \in \mathbb{R}^k$, on a

$$1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n \Longrightarrow \mathbb{P}\left(\bigcap_{j=1}^{k} (X_{i_j} = x_j)\right) = \prod_{j=1}^{k} \mathbb{P}(X_{i_j} = x_j).$$

Remarque 19. Il découle directement de la définition que si les variables aléatoires $(X_k)_{1\leqslant k\leqslant n}$ sont mutuellement indépendantes, alors toute sous-famille de $(X_k)_{1\leqslant k\leqslant n}$ est également constituée de variables aléatoires mutuellement indépendantes.

23.3.2 Propriétés de l'indépendance de variables aléatoires

Proposition 13

Si X et Y sont des variables aléatoires indépendantes sur un même espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, alors pour toutes fonctions f et g, les variables aléatoires $f(X)$ et $g(Y)$ sont également indépendantes.

Démonstration. Soient $A \in \mathcal{P}(f \circ X(\Omega))$ et $B \in \mathcal{P}(g \circ Y(\Omega))$. Par indépendance de X et Y, on a

$$\begin{aligned} \mathbb{P}(f(X) \in A, g(Y) \in B) &= \mathbb{P}(X \in f^{-1}(A), Y \in g^{-1}(B)) \\ &= \mathbb{P}(X \in f^{-1}(A))\mathbb{P}(Y \in g^{-1}(B)) \\ &= \mathbb{P}(f(X) \in A)\mathbb{P}(g(Y) \in B), \end{aligned}$$

ce qui prouve l'indépendance des variables aléatoires $f(X)$ et $g(Y)$. ∎

Remarque 20. En pratique, cela signifie que si les variables aléatoires X et Y sont indépendantes, alors les variables aléatoires X^2 et $-Y$ aussi, ou encore $\exp(X)$ et $4Y^3 + Y - 1$...

Lemme 1: Lemme des coalitions

Soient $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{n+p}$ des variables aléatoires mutuellement indépendantes. Alors pour tout couple de fonctions (u, v), les variables aléatoires $u(X_1, \ldots, X_n)$ et $v(X_{n+1}, \ldots, X_{n+p})$ sont indépendantes.

Démonstration. Démonstration hors-programme. ∎

Remarque 21. Le résultat s'étend naturellement au cas de plus de deux coalitions.

En particulier, si $u_1, \ldots, u_n$ sont n fonctions, alors les variables aléatoires $u_1(X_1), \ldots, u_n(X_n)$ sont mutuellement indépendantes.

Proposition 14: Espérance du produit de deux variables aléatoires indépendantes

Soient X et Y deux variables aléatoires indépendantes définies sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
Alors
$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Démonstration. Notons $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_p\}$ où les $(x_i)_{1 \leqslant i \leqslant n}$ sont deux à deux distincts, ainsi que les $(y_j)_{1 \leqslant j \leqslant p}$.

Alors $XY(\Omega) = \{x_i y_j, (i,j) \in [\![1,n]\!] \times [\![1,p]\!]\}$ où les $x_i y_j$ ne sont pas nécessairement deux à deux distincts.

Posons $XY(\Omega) = \{z_1, \dots, z_r\}$ avec pour tout $(i,j) \in [\![1,r]\!]^2$, avec $i \neq j, z_i \neq z_j$. On a par définition de l'espérance $\mathbb{E}(XY) = \displaystyle\sum_{k=1}^{r} z_k \mathbb{P}(XY = z_k)$.

Or, pour tout $k \in [\![1,r]\!]$, on a $(XY = z_k) = \displaystyle\bigsqcup_{\substack{(i,j) \in [\![1,n]\!] \times [\![1,p]\!] \\ x_i y_j = z_k}} (X = x_i) \cap (Y = y_j)$ donc

$$\mathbb{P}(XY = z_k) = \sum_{\substack{(i,j) \in [\![1,n]\!] \times [\![1,p]\!] \\ x_i y_j = z_k}} \mathbb{P}((X = x_i) \cap (Y = y_j))$$

d'où

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{k=1}^{r} z_k \sum_{\substack{(i,j) \in [\![1,n]\!] \times [\![1,p]\!] \\ x_i y_j = z_k}} \mathbb{P}((X = x_i) \cap (Y = y_j)) \\
&= \sum_{k=1}^{r} \sum_{\substack{(i,j) \in [\![1,n]\!] \times [\![1,p]\!] \\ x_i y_j = z_k}} x_i y_j \mathbb{P}((X = x_i) \cap (Y = y_j)) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{p} x_i y_j \mathbb{P}((X = x_i) \cap (Y = y_j)) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{p} x_i y_j \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j) \quad \text{par indépendance de } X \text{ et } Y \\
&= \left(\sum_{i=1}^{n} x_i \mathbb{P}(X = x_i) \right) \left(\sum_{j=1}^{p} y_j \mathbb{P}(y = y_j) \right) \\
&= \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$

■

Remarque 22. • Le résultat s'étend naturellement à n variables aléatoires indépendantes.

• La réciproque est fausse ! On peut avoir $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ sans que X et Y soient indépendantes.

Soit X une variable aléatoire suivant une loi uniforme sur $\{-1, 0, 1\}$, i.e.

$$\mathbb{P}(X = -1) = \mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{3}.$$

Soit $Y = X^2$.

Alors $\mathbb{E}(XY) = \mathbb{E}(X^3) = \mathbb{E}(X) = 0$

Mais X et Y ne sont pas indépendantes car

$$\mathbb{P}((X=0)\cap(Y=1)) = 0 \neq \mathbb{P}(X=0)\mathbb{P}(Y=1) = \frac{2}{9}.$$

Remarque 23. Soient $(A,B)\in\mathcal{P}(\Omega)^2$ telles que $X=\mathbb{1}_A$ et $Y=\mathbb{1}_B$.

Tout d'abord, vérifions que X et Y sont indépendantes si et seulement si les événements A et B sont indépendants.

• Si X et Y sont indépendantes, les événements $(X=1)$ et $(Y=1)$ sont indépendants, i.e. A et B sont indépendants.

• Si $A=(X=1)$ et $B=(Y=1)$ sont indépendants, on a vu dans le chapitre précédent que ceci impliquait l'indépendance des couples d'événements $(A,\overline{B}) = ((X=1),(Y=0))$, $(\overline{A},B) = ((X=0),(Y=1)$ et $(\overline{A},\overline{B}) = ((X=0),(Y=0)$, ce qui prouve que les variables aléatoires X et Y sont indépendantes.

D'après la proposition précédente, on a alors

$$\mathbb{P}(A\cap B) = \mathbb{E}(\mathbb{1}_{A\cap B}) = \mathbb{E}(\mathbb{1}_A\mathbb{1}_B) = \mathbb{E}(\mathbb{1}_A)\mathbb{E}(\mathbb{1}_B) = \mathbb{P}(A)\mathbb{P}(B),$$

ce qui n'est rien d'autre que la définition de l'indépendance des événements A et B !

Corollaire 3: Covariance de deux variables aléatoires indépendantes

Soient X et Y deux variables aléatoires indépendantes définies sur un espace probabilisé fini $(\Omega,\mathcal{P}(\Omega),\mathbb{P})$.
Alors $\mathrm{Cov}(X,Y)=0$.

Démonstration. En effet, d'après la proposition précédente, puisque X et Y sont indépendantes, on a

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0.$$

■

Remarque 24. Puisqu'on peut avoir $\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$ sans que les variables aléatoires X et Y sont indépendantes, on peut avoir $\mathrm{Cov}(X,Y)=0$ sans que X et Y soient indépendantes.

Proposition 15: Variance de la somme de deux variables aléatoires indépendantes

Soient X et Y deux variables aléatoires indépendantes définies sur un espace probabilisé fini $(\Omega,\mathcal{P}(\Omega),\mathbb{P})$.
Alors

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

Démonstration. En effet, on sait que $V(X+Y) = V(X)+V(Y)+2\mathrm{Cov}(X,Y)$. Or, puisque X et Y sont indépendantes, on a $\mathrm{Cov}(X,Y)=0$ d'où le résultat. ■

Remarque 25. Si $(X_1,\dots,X_n)$ sont des variables mutuellement indépendantes, alors

$$V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k).$$

23.4 Lois usuelles

Dans cette section, nous appliquerons implicitement la proposition 23.1.2 qui nous permet de définir des lois pourvu que toutes les probabilités soient des nombres positifs et que leur somme fasse 1.

23.4.1 Loi certaine

Définition 14: Loi certaine

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
On dit que X suit une loi certaine si X est constante, i.e.

$$\exists a \in \mathbb{R}, \forall \omega \in \Omega, X(\omega) = a.$$

Remarque 26. Cette loi modélise un phénomène dont l'issue est certaine.

On a déjà vu l'espérance et la variance d'une loi certaine : rappelons-les.

Proposition 16: Espérance et variance d'une loi certaine

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
On suppose que X est constante égale à $a \in \mathbb{R}$.
Alors
$$\mathbb{E}(X) = a \quad \text{et} \quad V(X) = 0.$$

23.4.2 Loi uniforme

Définition 15: Loi uniforme

Soit X une variable aléatoire définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Notons $X(\Omega) = \{x_1, \ldots, x_n\}$ avec $n \in \mathbb{N}^*$.
On dit que X suit une loi uniforme si

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

Remarque 27. On remarque qu'on a bien $\displaystyle\sum_{k=1}^{n} \mathbb{P}(X = x_k) = \frac{n}{n} = 1$.

Exemple 13. La loi uniforme modélise classiquement le résultat d'un lancer de dé équilibré.

Proposition 17: Espérance et variance d'une loi uniforme

Soit X une variable aléatoire suivant une loi uniforme sur $X(\Omega) = [\![1, n]\!]$.
On note $X \hookrightarrow \mathcal{U}(n)$. Alors

$$\mathbb{E}(X) = \frac{n+1}{2} \quad \text{et} \quad V(X) = \frac{n^2-1}{12}.$$

Démonstration.

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

2. On a $V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
   Or, d'après le théorème du transfert,
$$\mathbb{E}(X^2) = \sum_{k=1}^{n} k^2\mathbb{P}(X = k) = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6n} = \frac{(n+1)(2n+1)}{6}.$$

Ainsi,

$$V(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n+1)(4n+2-3n-3)}{12} = \frac{n^2-1}{12}.$$

■

23.4.3 Loi de Bernoulli

Définition 16: Loi de Bernoulli

Soit X une variable aléatoire à valeurs dans $\{0,1\}$. Soit $p \in [0,1]$.
On dit que la variable aléatoire X suit une loi de Bernoulli de paramètre p, et on note $X \hookrightarrow \mathcal{B}(p)$, si

$$\mathbb{P}(X=1) = p \quad \text{et} \quad \mathbb{P}(X=0) = 1-p.$$

On note souvent $q = 1-p$.

Remarque 28. • On remarque qu'on a bien $\mathbb{P}(X=0) + \mathbb{P}(X=1) = 1$.

• Une variable aléatoire X qui suit une loi de Bernoulli peut en fait être vue comme l'indicatrice de l'événement $(X=1)$.

Réciproquement, l'indicatrice d'un événement A est une variable aléatoire qui suit une loi de Bernoulli de paramètre $p = \mathbb{P}(A)$.

Exemple 14. La loi de Bernoulli modélise un lancer de pièce dont la probabilité d'obtenir pile serait égale à p.

Plus généralement, elle modélise une expérience dont le résultat est binaire (échec ou succès) et dont la probabilité du succès est p.

Proposition 18: Espérance et variance d'une loi de Bernoulli

Soit $X \hookrightarrow \mathcal{B}(p)$, avec $p \in [0,1]$. Alors

$$\mathbb{E}(X) = p \quad \text{et} \quad V(X) = p(1-p).$$

Démonstration.

1. On a
$$\mathbb{E}(X) = 0\mathbb{P}(X=0) + 1\mathbb{P}(X=1) = p.$$
2. On a d'après le théorème du transfert
$$\mathbb{E}(X^2) = 0^2\mathbb{P}(X=0) + 1^2\mathbb{P}(X=1) = p$$
donc
$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1-p).$$

■

23.4.4 Loi binomiale

Définition 17: Loi binomiale

Soit $n \in \mathbb{N}^*$, soit $p \in [0,1]$. Soit X une variable aléatoire à valeurs dans $[\![0,n]\!]$.
On dit que X suit une loi binomiale de paramètres n et p, et on note $X \hookrightarrow \mathcal{B}(n,p)$, si pour tout $k \in [\![0,n]\!]$,

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

Remarque 29. On a bien d'après la formule du binôme de Newton

$$\sum_{k=0}^{n} \mathbb{P}(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + 1 - p)^n = 1.$$

Exemple 15. La loi binomiale $\mathcal{B}(n,p)$ modélise le nombre de succès lors de la réalisation successive de n expériences de Bernoulli indépendantes de paramètre p (par exemple n lancers de pièce de monnaie).

En effet, l'événement $(X = k)$ signifie qu'on a obtenu k succès (par exemple k pile). Il y a $\binom{n}{k}$ façons de placer ces succès dans les n expériences. Pour calculer la probabilité de cet événement, il faut ensuite multiplier k fois par la probabilité d'obtenir un succès, ce qui donne p^k, et $n-k$ fois par la probabilité d'échouer, ce qui donne $(1-p)^{n-k}$.

On peut donner cette définition alternative, confirmée par l'interprétation de la loi binomiale, et qui est très utile en pratique.

Définition 18: Somme de variables aléatoires de Bernoulli indépendantes et de même paramètre

Soit $n \in \mathbb{N}^*$, soit $p \in [0,1]$.
On dit que la variable aléatoire X suit une loi binomiale de paramètre n et p si

$$X = \sum_{k=1}^{n} X_k$$

où les variables aléatoires X_k sont mutuellement indépendantes et suivent toutes une loi de Bernoulli de paramètre p.

Remarque 30. Si $X_1, \dots, X_m$ sont indépendantes et suivent toutes une loi binomiale de paramètres (n,p), alors $\displaystyle\sum_{k=1}^{m} X_k$ suit une loi binomiale de paramètres (nm, p).

Proposition 19: Espérance et variance d'une loi binomiale

Soit $X \hookrightarrow \mathcal{B}(n,p)$ avec $n \in \mathbb{N}^*$ et $p \in [0,1]$. Alors

$$\mathbb{E}(X) = np \quad \text{et} \quad V(X) = np(1-p).$$

Démonstration.

Dans toute cette preuve, nous écrirons $X = \displaystyle\sum_{k=1}^{n} X_k$ où les variables aléatoires X_k suivent toutes une loi de Bernoulli de paramètre p et sont mutuellement indépendantes.

1. Par linéarité de l'espérance, on a

$$\mathbb{E}(X) = \sum_{k=1}^{n} \mathbb{E}(X_k) = np.$$

2. Puisque les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes, on a

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

■

23.5 Inégalités probabilistes

23.5.1 Inégalités de Markov et de Bienaymé-Tchebychev

Proposition 20: Inégalité de Markov

Soit X une variable aléatoire à valeurs dans $\mathbb{R}_+$, définie sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ fini.
Alors pour tout $a > 0$,
$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}.$$

Démonstration. Soit $a > 0$.

Soit $\omega \in \Omega$.

- Si $X(\omega) \geqslant a$ alors $\omega \in (X \geqslant a)$ donc $a\mathbb{1}_{(X\geqslant a)}(\omega) = a \leqslant X(\omega)$.
- Si $X(\omega) < a$, alors $\omega \notin (X \geqslant a)$ donc $a\mathbb{1}_{(X\geqslant a)}(\omega) = 0 \leqslant X(\omega)$ par positivité de X.

Ainsi, pour tout $\omega \in \Omega, X(\omega) \geqslant a\mathbb{1}_{(X\geqslant a)}(\omega)$, i.e. $X \geqslant a\mathbb{1}_{(X\geqslant a)}$.

Par croissance de l'espérance, on en déduit que

$$\mathbb{E}(X) \geqslant \mathbb{E}(a\mathbb{1}_{(X\geqslant a)}) = a\mathbb{E}(\mathbb{1}_{(X\geqslant a)}) = a\mathbb{P}(X \geqslant a).$$

En divisant par $a > 0$, on obtient bien $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X)}{a}$. ■

Proposition 21: Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle définie sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.
Alors pour tout $a > 0$,
$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant a) \leqslant \frac{V(X)}{a^2}.$$

Démonstration. Soit $a > 0$.

Il s'agit d'appliquer l'inégalité de Markov à la variable aléatoire $(X - \mathbb{E}(X))^2$ et au réel strictement positif a^2. On a ainsi,

$$\mathbb{P}((X - \mathbb{E}(X))^2 \geqslant a^2) \leqslant \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{a^2} = \frac{V(X)}{a^2}.$$

On obtient le résultat en remarquant que $((X - \mathbb{E}(X))^2 \geq a^2) = (|X - \mathbb{E}(X)| \geq a)$ (car $a > 0$). ■

Remarque 31. Cette inégalité dit que plus la variance d'une variable aléatoire est petite, plus la probabilité que la variable aléatoire soit proche de son espérance est grande.

23.5.2 Loi faible des grands nombres

Théorème 2: Loi faible des grands nombres

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires deux à deux indépendantes et de même loi.

On note $m = \mathbb{E}(X_1), \sigma^2 = V(X_1)$ et pour tout $n \in \mathbb{N}, S_n = \displaystyle\sum_{k=1}^{n} X_k$.

Alors pour tout $\varepsilon > 0$, on a

$$\mathbb{P}\left(\left|\frac{1}{n}S_n - m\right| \geqslant \varepsilon\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

Démonstration. Puisque toutes les variables aléatoires ont la même loi, on a pour tout $n \in \mathbb{N}, \mathbb{E}(X_n) = m$ et $V(X_n) = \sigma^2$.

Soit $n \in \mathbb{N}^*$.

Par linéarité de l'espérance, on a $\mathbb{E}(\frac{S_n}{n}) = \dfrac{1}{n}\mathbb{E}\left(\displaystyle\sum_{k=1}^{n} X_k\right) = \dfrac{nm}{n} = m$.

Puisque les variables aléatoires $X_1, \ldots, X_n$ sont deux à deux indépendantes, on a

$$V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\left(\sum_{k=1}^{n} V(X_k)\right) = \frac{\sigma^2}{n}.$$

D'après l'inégalité de Bienaymé-Tchebychev appliquée à la variable aléatoire $\frac{S_n}{n}$, on a pour tout $\varepsilon > 0$

$$0 \leqslant \mathbb{P}\left(\left|\frac{1}{n}S_n - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2} \underset{n\to+\infty}{\longrightarrow} 0.$$

■

Remarque 32. On lance un nombre infini (ou en tout cas très grand) de fois une pièce équilibrée. Pour tout $n \in \mathbb{N}$, la probabilité d'obtenir pile est modélisée par la variable aléatoire $X_n \hookrightarrow \mathcal{B}(\frac{1}{2})$.

Avec les notations du théorème, la variable aléatoire $\dfrac{S_n}{n}$ représente la moyenne du nombre de piles obtenus au cours des n premiers lancers, calculée de façon empirique.

Soit $\varepsilon > 0$ fixé. Le théorème affirme que pour un nombre n de lancers très grand, la probabilité que l'écart entre la moyenne empirique et la moyenne théorique m, ici égale à $\frac{1}{2}$, soit supérieur à ε tend vers 0.

Concrètement, si on lance 100 fois une pièce de monnaie, le nombre de piles obtenus devrait s'approcher de 50.

D'après l'estimation obtenue dans le théorème, pour $\varepsilon = 0,1$, la probabilité que le nombre de piles soit en-dessous de 40 ou au-dessus de 60 est inférieure à $\dfrac{100}{4\times 100} = \dfrac{1}{4}$.

Si on lance la pièce 1000 fois, le nombre de piles devrait être encore plus proche de 500. Cette fois, pour $\varepsilon = 0,1$, la probabilité que le nombre de piles soit en-dessous de 400 ou au-dessus de 600 est inférieure à $\dfrac{100}{4000} = \dfrac{1}{40}$.