

V - Variables aléatoires discrètes finies

Révisions

Probabilités.

Dans tout le cours, Ω désigne un univers et \mathbf{P} est une probabilité.

I - Variables aléatoires réelles finies

I.1 - Définition

Définition 1 - Variable aléatoire réelle

On suppose que $\Omega = \{\omega_1, \dots, \omega_n\}$ est un ensemble fini. Une *variable aléatoire* est une application de Ω dans \mathbb{R} .

Exemple 1 - Somme de 2 lancers d'un dé équilibré

Un dé équilibré est lancé successivement 2 fois. On note les résultats obtenus à l'issue de chacun des lancers.

L'univers est $\Omega = \{(i, j), 1 \leq i, j \leq 6\} = \llbracket 1, 6 \rrbracket^2$.

La somme S des résultats obtenus à l'issue des 2 lancers est une variable aléatoire :

$$\begin{aligned} S : \quad \Omega &\rightarrow \mathbb{R} \\ (d_1, d_2) &\mapsto d_1 + d_2 \end{aligned}$$

Notations

Si $x \leq y$ sont des réels, on notera

$$[X = x] = \{\omega \in \Omega ; X(\omega) = x\}$$

$$[X \leq x] = \{\omega \in \Omega ; X(\omega) \leq x\}$$

$$[X \geq x] = \{\omega \in \Omega ; X(\omega) \geq x\}$$

$$[x \leq X \leq y] = \{\omega \in \Omega ; x \leq X(\omega) \leq y\}$$

Exemple 2 - Somme de 2 lancers d'un dé équilibré

On reprend les notations de l'exemple précédent.

$$[S = 3] = \{(d_1, d_2) \in \Omega ; d_1 + d_2 = 3\}$$

$$= \{(1, 2), (2, 1)\}.$$

$$[S \leq 4] = \{(d_1, d_2) \in \Omega ; d_1 + d_2 \leq 4\}$$

$$= \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\}.$$

Définition 2 - Système complet

Soit Ω un univers fini et X une variable aléatoire. Notons x_1, \dots, x_p les valeurs prises par X . Alors, $\{[X = x_i], i \in \llbracket 1, p \rrbracket\}$ est un système complet d'événements. C'est le *système complet* associé à la variable aléatoire X .

Exemple 3 - Somme de 2 lancers d'un dé équilibré

On reprend les notations de l'exemple précédent. Le système complet associé à S est :

$$\begin{aligned} &\{[S = 2], [S = 3], [S = 4], [S = 5], [S = 6], [S = 7], \\ &\quad [S = 8], [S = 9], [S = 10], [S = 11], [S = 12]\}. \end{aligned}$$

I.2 - Loi de probabilité

Définition 3 - Loi de probabilité

La loi de la variable aléatoire X est la donnée :

- des valeurs x_1, \dots, x_p prises par X ,
- de la famille de probabilités

$$(\mathbf{P}([X = x_1]), \dots, \mathbf{P}([X = x_p])).$$

La loi d'une variable aléatoire peut être représentée sous forme d'un tableau.

Lois usuelles - À connaître par cœur

- **Somme de deux dés équilibrés.** Nous obtenons successivement :

| k | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| $\mathbf{P}([S = k])$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

- **Loi certaine.** La variable aléatoire X ne prend qu'une seule valeur c et $\mathbf{P}([X = c]) = 1$. L'expérience renvoie toujours le même résultat.
- **Loi uniforme sur $\llbracket 1, n \rrbracket$.** Notée $\mathcal{U}(\llbracket 1, n \rrbracket)$. L'expérience renvoie les résultats $1, 2, \dots$, ou n et chacun de ces résultats a la même probabilité d'apparaître. Les valeurs prises par X sont $1, \dots, n$ et

$$\forall i \in \llbracket 1, n \rrbracket, \mathbf{P}([X = i]) = \frac{1}{n}.$$

- **Loi de Bernoulli de paramètre p .** Notée $\mathcal{B}(p)$. L'expérience (dite expérience de Bernoulli) renvoie le résultat 0 (on parle d'échec) ou 1 (on parle de succès). Le succès apparaît avec probabilité p . Soit $p \in [0, 1]$. Les valeurs prises par X sont 0 et 1 et

$$\mathbf{P}([X = 0]) = 1 - p \text{ et } \mathbf{P}([X = 1]) = p.$$

- **Loi binomiale de paramètres n et p .** Notée $\mathcal{B}(n, p)$.

La variable aléatoire compte le nombre de succès dans une succession de n expériences de Bernoulli indépendantes (appelé schéma de Bernoulli) dont la probabilité de succès vaut p .

Soit $p \in [0, 1]$. Les valeurs prises par X sont $0, \dots, n$ et

$$\forall k \in \llbracket 0, n \rrbracket, \mathbf{P}([X = k]) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Définition 4 - Fonction de répartition

Soit X une variable aléatoire. La fonction de répartition de X est la fonction F_X définie sur \mathbb{R} par $F_X(x) = \mathbf{P}([X \leq x])$.

Exemple 4 - Fonctions de répartition

- Si X suit une loi certaine de valeur $\frac{3}{2}$.

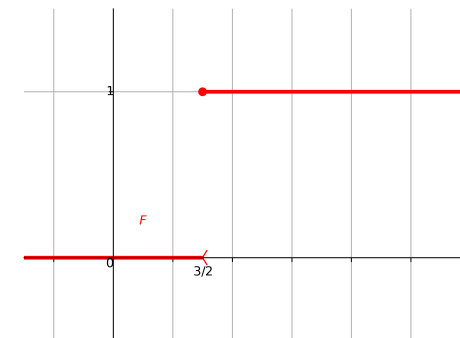
★ Si $x < \frac{3}{2}$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}(\emptyset) = 0.$$

★ Si $x \geq \frac{3}{2}$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}([X = 3/2]) = 1.$$

Ainsi, la fonction de répartition de X est :



- Si X suit une loi de Bernoulli de paramètre $\frac{3}{4}$.
 - ★ Si $x < 0$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}(\emptyset) = 0.$$

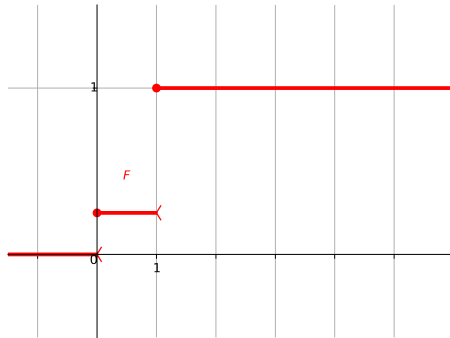
- ★ Si $0 \leq x < 1$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}([X = 0]) = \frac{1}{4}.$$

- ★ Si $x \geq 1$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}(\Omega) = 1.$$

Ainsi, la fonction de répartition de X est :



- Si X suit une loi uniforme sur $\llbracket 1, 4 \rrbracket$.
 - ★ Si $x < 1$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}(\emptyset) = 0.$$

- ★ Si $1 \leq x < 2$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}([X = 1]) = \frac{1}{4}.$$

- ★ Si $2 \leq x < 3$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}([X = 1] \sqcup [X = 2]) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

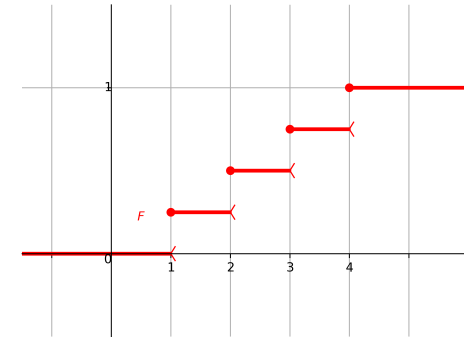
- ★ Si $3 \leq x < 4$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}([X = 1] \sqcup [X = 2] \sqcup [X = 3]) = \frac{3}{4}.$$

- ★ Si $x \geq 4$, alors

$$\mathbf{P}([X \leq x]) = \mathbf{P}(\Omega) = 1.$$

Ainsi, la fonction de répartition de X est :



- Si S est la somme des résultats de 2 lancers successifs d'un dé équilibré à 6 faces.
 - ★ Si $x < 2$. Alors,

$$F_S(x) = \mathbf{P}([S \leq x]) = \mathbf{P}(\emptyset) = 0.$$

- ★ Si $2 \leq x < 3$. Alors,

$$F_S(x) = \mathbf{P}([S \leq x]) = \mathbf{P}([S = 2]) = \frac{1}{36}.$$

- ★ Si $3 \leq x < 4$. Alors,

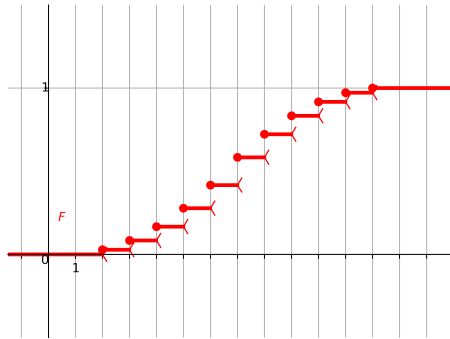
$$F_S(x) = \mathbf{P}([S \leq x]) = \mathbf{P}([S = 2] \sqcup [S = 3]) = \frac{1}{36} + \frac{2}{36}.$$

- ★ ...

- ★ Si $x \geq 12$. Alors,

$$F_S(x) = \mathbf{P}([S \leq x]) = \mathbf{P}(\Omega) = 1.$$

Ainsi, la fonction de répartition de S est :



Proposition 1

Les variables aléatoires X et Y sont de même loi si et seulement si elles ont mêmes fonctions de répartition.

II - Espérance & Variance

II.1 - Espérance

Définition 5 - Espérance

Soit X une variable aléatoire et x_1, \dots, x_p les valeurs prises par X . L'espérance de X , notée $\mathbf{E}[X]$, est le réel

$$\mathbf{E}[X] = x_1 \mathbf{P}([X = x_1]) + \dots + x_p \mathbf{P}([X = x_p]).$$

Exemple 5 - Lois usuelles - À connaître par cœur

- Loi certaine de valeur c .

$$\mathbf{E}[X] = c \mathbf{P}([X = c]) = c.$$

- Loi uniforme sur $\llbracket 1, n \rrbracket$.

$$\begin{aligned} \mathbf{E}[X] &= 1 \mathbf{P}([X = 1]) + 2 \mathbf{P}([X = 2]) + \dots + n \mathbf{P}([X = n]) \\ &= 1 \times \frac{1}{n} + 2 \times \frac{1}{n} + \dots + n \times \frac{1}{n} = \frac{1}{n} (1 + 2 + \dots + n) \\ &= \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2}. \end{aligned}$$

- Loi de Bernoulli de paramètre p .

$$\begin{aligned} \mathbf{E}[X] &= 0 \times \mathbf{P}([X = 0]) + 1 \times \mathbf{P}([X = 1]) \\ &= 0 \times (1 - p) + 1 \times p = p. \end{aligned}$$

- Loi binomiale de paramètres n et p .

$$\mathbf{E}[X] = \sum_{k=0}^n k \times \binom{n}{k} p^k (1-p)^{n-k} = n \times p.$$

Proposition 2 - Linéarité

Soit X, Y deux variables aléatoires et a, b deux réels. Alors, $\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$.

Exemple 6 - Somme de deux dés

Notons X_1 le résultat du premier lancer et X_2 le résultat du second. Alors, X_1 et X_2 suivent une loi uniforme sur $\llbracket 1, 6 \rrbracket$ et $S = X_1 + X_2$. Ainsi,

$$\mathbf{E}[S] = \mathbf{E}[X_1 + X_2] = \mathbf{E}[X_1] + \mathbf{E}[X_2] = \frac{6+1}{2} + \frac{6+1}{2} = 7.$$

Proposition 3 - Théorème de transfert

Soit X une variable aléatoire, x_1, \dots, x_p les valeurs prises par X et g une fonction à valeurs réelles. On note $Y = g(X)$ la variable aléatoire définie par

$$\forall \omega \in \Omega, Y(\omega) = g(X(\omega)).$$

Alors,

$$\mathbf{E}[Y] = \mathbf{E}[g(X)] = \sum_{k=1}^p g(x_k) \mathbf{P}([X = x_k]).$$

Exemple 7 - Carré de la somme

On reprend les notations de l'exemple précédent. Posons $Y = S^2$. Alors,

$$\begin{aligned} \mathbf{E}[Y] &= 2^2 \frac{1}{36} + 3^2 \frac{2}{36} + 4^2 \frac{3}{36} + 5^2 \frac{4}{36} + 6^2 \frac{5}{36} + 7^2 \frac{6}{36} + \dots \\ &\quad \dots + 8^2 \frac{5}{36} + 9^2 \frac{4}{36} + 10^2 \frac{3}{36} + 11^2 \frac{2}{36} + 12^2 \frac{1}{36} \end{aligned}$$

II.2 - Variance**Définition 6 - Variance, Écart-type**

Soit X une variable aléatoire. La *variance* de X , notée $\mathbf{V}(X)$, est le réel :

$$\mathbf{V}(X) = \mathbf{E}[(X - \mathbf{E}[X])^2].$$

L'écart-type est $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

Variance

Comme l'espérance s'interprète comme une moyenne, la variance est la moyenne des écarts quadratiques (c'est-à-dire au carré) à la moyenne.

Exemple 8 - Lois usuelles - À connaître par cœur

- **Loi certaine de valeur c .**

$$\mathbf{V}(X) = 0.$$

- **Loi uniforme sur $\llbracket 1, n \rrbracket$.**

$$\mathbf{V}(X) = \frac{n^2 - 1}{12}.$$

- **Loi de Bernoulli de paramètre p .**

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

- **Loi binomiale de paramètres n et p .**

$$\mathbf{V}(X) = np(1 - p).$$

Proposition 4

Soit X une variable aléatoire et a, b deux réels. Alors,

$$\mathbf{V}(aX + b) = a^2 \mathbf{V}(X).$$

Exemple 9

Soit Y une variable suivant une loi uniforme sur $\llbracket 2, 12 \rrbracket$. On pose $X = Y - 1$.

Déterminons la loi de X .

| k | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| $\mathbf{P}([Y = k])$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

Comme $X = Y - 1$,

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| $\mathbf{P}([X = j])$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

Ainsi, $X \sim \mathcal{U}(\llbracket 1, 11 \rrbracket)$.

Déterminons l'espérance et la variance de Y .

$$\mathbf{E}[Y] = \mathbf{E}[X + 1] = \mathbf{E}[X] + \mathbf{E}[1] = \frac{12}{2} + 1 = 6 + 1 = 7,$$

$$\mathbf{V}(Y) = \mathbf{V}(X + 1) = \mathbf{V}(X) = \frac{11^2 - 1}{2} = 60.$$

Théorème 1 - Formule de Kœnig-Huygens

Soit X une variable aléatoire. Alors,

$$\mathbf{V}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2.$$

Définition 7 - Variable centrée, réduite

Soit X une variable aléatoire.

- X est une variable centrée si $\mathbf{E}[X] = 0$.
- X est une variable réduite si $\mathbf{V}(X) = 1$.

Proposition 5

Soit X une variable aléatoire qui ne soit pas de loi certaine. La variable $X^* = \frac{X - \mathbf{E}[X]}{\sigma(X)}$ est une variable centrée réduite.

III - Couple de variables aléatoires

On considère une variable aléatoire X qui prend les valeurs x_1, \dots, x_p et une variable aléatoire Y qui prend les valeurs y_1, \dots, y_q .

III.1 - Loi du couple

Définition 8 - Loi du couple

La *loi du couple* (X, Y) est la donnée :

- des valeurs (x_i, y_j) prises par le couple (X, Y) ,
- des probabilités $\mathbf{P}([X = x_i] \cap [Y = y_j])$.

Exemple 10 - 2 feuilles et 2 tiroirs

On dispose d'un bureau à 2 tiroirs et de 2 feuilles de papier. On dispose aléatoirement chacune des feuilles de papier dans l'un des tiroirs.

On note X le nombre de feuilles dans le premier tiroir et Y le nombre de tiroirs vides.

Le nombre de feuilles de papier dans le premier tiroir peut être égal à 0, 1 ou 2.

Le nombre de tiroirs vides peut être égal à 0 (il y a 1 feuille dans chaque tiroir) ou 1 (les 2 feuilles sont dans le même tiroir).

De plus,

$$\mathbf{P}([X = 0] \cap [Y = 0]) = 0$$

$$\mathbf{P}([X = 0] \cap [Y = 1]) = \frac{1}{4}$$

$$\mathbf{P}([X = 1] \cap [Y = 0]) = \frac{1}{2}$$

$$\mathbf{P}([X = 1] \cap [Y = 1]) = 0$$

$$\mathbf{P}([X = 2] \cap [Y = 0]) = 0$$

$$\mathbf{P}([X = 2] \cap [Y = 1]) = \frac{1}{4}$$

On peut représenter ces résultats dans un tableau : à chaque ligne correspond une valeur x que peut prendre X ; à chaque colonne correspond une valeur y que peut prendre Y ; à l'intersection d'une ligne et d'une colonne se lit la probabilité $\mathbf{P}([X = x] \cap [Y = y])$:

| $x \backslash y$ | 0 | 1 |
|------------------|---------------|---------------|
| 0 | 0 | $\frac{1}{4}$ |
| 1 | $\frac{1}{2}$ | 0 |
| 2 | 0 | $\frac{1}{4}$ |

Définition 9 - Marginales

Les lois de X et de Y sont les *marginales* du couple (X, Y) . En utilisant le système complet associé à la variable aléatoire Y (resp. X), on obtient

$$\forall i \in \llbracket 1, p \rrbracket, \mathbf{P}([X = x_i]) = \sum_{j=1}^q \mathbf{P}([X = x_i] \cap [Y = y_j])$$

$$\forall j \in \llbracket 1, q \rrbracket, \mathbf{P}([Y = y_j]) = \sum_{i=1}^p \mathbf{P}([X = x_i] \cap [Y = y_j])$$

Exemple 11 - 2 feuilles et 2 tiroirs

En reprenant les notations de l'exemple précédent, les marginales s'obtiennent en sommant les lignes / les colonnes du tableau.

| $x \backslash y$ | 0 | 1 | $\mathbf{P}([X = x])$ |
|-----------------------|---------------|---------------|-----------------------|
| 0 | 0 | $\frac{1}{4}$ | $\frac{1}{4}$ |
| 1 | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |
| 2 | 0 | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $\mathbf{P}([Y = y])$ | $\frac{1}{2}$ | $\frac{1}{2}$ | |

Définition 10 - Loi conditionnelle

La *loi conditionnelle* de X sachant $[Y = y_j]$ est la donnée :

- des valeurs x_1, \dots, x_p prises par X ,
- des probabilités $\mathbf{P}_{[Y=y_j]}([X = x_1]), \dots, \mathbf{P}_{[Y=y_j]}([X = x_p])$.

Exemple 12 - 2 feuilles et 2 tiroirs

On reprend les notations de l'exercice précédent. La loi conditionnelle de X sachant $[Y = 0]$ est égale à

| i | 0 | 1 | 2 |
|-------------------------------|---|---|---|
| $\mathbf{P}_{[Y=0]}([X = i])$ | 0 | 1 | 0 |

C'est une loi presque certaine : *sachant qu'aucun des tiroirs n'est vide, on est certain qu'il y a une feuille dans le premier tiroir.*

La loi conditionnelle de X sachant $[Y = 1]$ est égale à

| i | 0 | 1 | 2 |
|-------------------------------|---------------|---|---------------|
| $\mathbf{P}_{[Y=1]}([X = i])$ | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |

Sachant qu'un des deux tiroirs est vide, il y a une chance sur deux que le premier tiroir contienne 0 feuille et une chance sur deux qu'il contienne les 2 feuilles.

III.2 - Indépendance**Définition 11 - Indépendance**

Les variables aléatoires X et Y sont *indépendantes* si pour tout $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$,

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

Exemple 13

- On considère le couple de variables aléatoires (X, Y) dont la loi est définie par

| $x \backslash y$ | 0 | 1 | $\mathbf{P}([X = x])$ |
|-----------------------|---------------|---------------|-----------------------|
| | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{2}$ |
| 0 | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{2}$ |
| 1 | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{2}$ |
| $\mathbf{P}([Y = y])$ | $\frac{1}{3}$ | $\frac{2}{3}$ | |

On étudie le comportement de **tous** les couples de valeurs possibles :

$$\mathbf{P}([X = 0] \cap [Y = 0]) = \frac{1}{6} = \frac{1}{2} \times \frac{1}{3} = \mathbf{P}([X = 0]) \times \mathbf{P}([Y = 0])$$

$$\mathbf{P}([X = 0] \cap [Y = 1]) = \frac{1}{3} = \frac{1}{2} \times \frac{2}{3} = \mathbf{P}([X = 0]) \times \mathbf{P}([Y = 1])$$

$$\mathbf{P}([X = 1] \cap [Y = 0]) = \frac{1}{6} = \frac{1}{2} \times \frac{1}{3} = \mathbf{P}([X = 1]) \times \mathbf{P}([Y = 0])$$

$$\mathbf{P}([X = 1] \cap [Y = 1]) = \frac{1}{3} = \frac{1}{2} \times \frac{2}{3} = \mathbf{P}([X = 1]) \times \mathbf{P}([Y = 1])$$

Ainsi, les variables aléatoires X et Y sont indépendantes.

- 2 feuilles et 2 tiroirs.** En reprenant la loi du couple,

$$\mathbf{P}([X = 0] \cap [Y = 0]) = 0 \neq \frac{1}{4} \times \frac{1}{2} = \mathbf{P}([X = 0]) \times \mathbf{P}([Y = 0]).$$

Les variables aléatoires X et Y **ne** sont **pas** indépendantes.

Indépendance en pratique

- Pour montrer que X et Y **ne** sont **pas** indépendantes, il suffit de trouver une case du tableau telle que

$$\mathbf{P}([X = x_k] \cap [Y = y_j]) \neq \mathbf{P}([X = x_k]) \times \mathbf{P}([Y = y_j]).$$

- Pour montrer que X et Y sont indépendantes, il faut montrer que cette égalité est vraie **pour toutes** les cases du tableau.

Proposition 6 - Loi certaine

Si Y est une variable aléatoire certaine, alors X et Y sont indépendantes.

III.3 - Covariance**Proposition 7 - Espérance d'un produit**

En utilisant la loi du couple,

$$\mathbf{E}[X \times Y] = \sum_{i=1}^p \sum_{j=1}^q x_i \times y_j \times \mathbf{P}([X = x_i] \cap [Y = y_j]).$$

Exemple 14 - 2 feuilles et 2 tiroirs

En reprenant l'exemple des feuilles et des tiroirs,

$$\begin{aligned} \mathbf{E}[XY] &= 0 \times 0 \times 0 + 0 \times 1 \times \frac{1}{4} + \dots \\ &\quad \dots + 1 \times 0 \times \frac{1}{2} + 1 \times 1 \times 0 + \dots \\ &\quad \dots + 2 \times 0 \times 0 + 2 \times 1 \times \frac{1}{4} \\ &= \frac{1}{2}. \end{aligned}$$

Théorème 2 - Espérance et Indépendance

Si X et Y sont indépendantes, alors

$$\mathbf{E}[X \times Y] = \mathbf{E}[X] \times \mathbf{E}[Y].$$

Définition 12 - Covariance

La *covariance* de X et Y , notée $\text{Cov}(X, Y)$, est le réel

$$\text{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}[X]) \times (Y - \mathbf{E}[Y])].$$

Proposition 8 - Propriétés de la covariance

Soit a, b, c trois réels et X, Y, Z trois variables aléatoires.

- $\text{Cov}(X, Y) = \mathbf{E}[X \times Y] - \mathbf{E}[X] \times \mathbf{E}[Y]$.
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- $\text{Cov}(X, X) = \mathbf{V}(X)$.
- $\text{Cov}(X, c) = 0$.
- $\text{Cov}(aX + bY, Z) = a\text{Cov}(X, Z) + b\text{Cov}(Y, Z)$.

Exemple 15 - 2 feuilles et 2 tiroirs

En reprenant les calculs précédents,

$$\mathbf{E}[X] = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1,$$

$$\mathbf{E}[Y] = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2},$$

$$\text{Cov}(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X] \mathbf{E}[Y] = \frac{1}{2} - 1 \times \frac{1}{2} = 0.$$

Proposition 9 - Covariance et Somme

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

Exemple 16 - Variance d'une somme

On considère le couple (X_1, X_2) dont la loi est définie par :

| $x_1 \backslash x_2$ | 0 | 1 | $\mathbf{P}([X_1 = x_1])$ |
|---------------------------|---------------|---------------|---------------------------|
| | 0 | $\frac{1}{4}$ | $\frac{1}{4}$ |
| 1 | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{3}{4}$ |
| $\mathbf{P}([X_2 = x_2])$ | $\frac{1}{4}$ | $\frac{3}{4}$ | |

X_1 et X_2 ont même loi, donc ils ont même espérance et même variance. De plus,

$$\mathbf{E}[X_1] = 0 \times \frac{1}{4} + 1 \times \frac{3}{4} = \frac{3}{4} = \mathbf{E}[X_1^2],$$

$$\mathbf{V}(X_1) = \mathbf{E}[X_1^2] - \mathbf{E}[X_1]^2 = \frac{3}{16},$$

$$\mathbf{E}[X_1 X_2] = 1 \times 1 \times \frac{1}{2} = \frac{1}{2}.$$

Ainsi,

$$\text{Cov}(X_1, X_2) = \mathbf{E}[X_1 X_2] - \mathbf{E}[X_1] \mathbf{E}[X_2] = \frac{1}{2} - \frac{9}{16} = -\frac{1}{16}.$$

Finalement,

$$\begin{aligned} \mathbf{V}(X_1 + X_2) &= \mathbf{V}(X_1) + \mathbf{V}(X_2) + 2\text{Cov}(X_1, X_2) \\ &= \frac{3}{16} + \frac{3}{16} - \frac{2}{16} = \frac{1}{4}. \end{aligned}$$

Proposition 10 - Covariance et Indépendance

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

Exemple 17 - 2 feuilles et 2 tiroirs

L'exemple des feuilles et des tiroirs montre que la covariance peut être nulle alors que les variables aléatoires ne sont pas indépendantes.

Définition 13 - Coefficient de corrélation linéaire

Le coefficient de corrélation linéaire de X et Y , noté $\rho(X, Y)$, est défini par

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbf{V}(X) \mathbf{V}(Y)}}.$$

Exemple 18

En reprenant l'exemple précédent,

$$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\mathbf{V}(X_1)\mathbf{V}(X_2)}} = \frac{\frac{-1}{16}}{\sqrt{\frac{3}{16} \times \frac{3}{16}}} = -\frac{1}{3}.$$

Proposition 11

- $-1 \leq \rho(X, Y) \leq 1$.
- $|\rho(X, Y)| = 1$ signifie qu'il existe trois réels a , b et c tels que $\mathbf{P}(aX + bY + c = 0) = 1$. Les variables X et Y sont reliées presque sûrement par une relation *affine*.