

XXX

Variables aléatoires



En théorie des probabilités, une variable aléatoire est une variable dont la valeur est déterminée après la réalisation d'un phénomène, expérience ou événement, aléatoire : la valeur d'un dé entre 1 et 6, le côté de la pièce dans un pile ou face, le jour de semaine de naissance de la prochaine personne que vous rencontrez, le temps d'attente dans la queue du cinéma, ...



Mathématiquement, c'est une application définie sur l'ensemble des éventualités, c'est-à-dire l'ensemble des résultats possibles d'une expérience aléatoire.

Ce furent les jeux de hasard qui amenèrent à concevoir les variables aléatoires, en associant à une éventualité (résultat du lancer d'un ou plusieurs dés, d'un tirage à pile ou face, d'une roulette, etc.) un gain.

Contenu

| | |
|---|-----------|
| I. Variables aléatoires..... | 2 |
| I.1 Définition | 2 |
| I.2 Image d'une variable aléatoire, Image réciproque | 2 |
| I.3 Fonction d'une variable aléatoire | 4 |
| I.4 Loi d'une variable aléatoire | 5 |
| II. Espérance d'une variable aléatoire réelle..... | 8 |
| II.1 Définition | 8 |
| II.2 Linéarité | 9 |
| II.3 Formule de transfert | 11 |
| III. Variance et écart-type d'une variable aléatoire réelle..... | 12 |
| III.1 Définitions | 12 |
| III.2 Inégalité de Markov et Bienaymé-Tchebychev | 13 |
| IV. Lois usuelles..... | 14 |
| IV.1 Loi uniforme | 14 |
| IV.2 Loi de Bernoulli | 15 |
| IV.3 Loi binomiale | 17 |
| V. Couple de variables aléatoires..... | 21 |
| V.1 Lois conjointe et marginale | 22 |
| V.2 Loi conditionnelle | 23 |
| VI. Variables aléatoires indépendantes..... | 25 |
| VI.1 Caractérisation | 26 |
| VI.2 Variables aléatoires mutuellement indépendantes | 27 |
| VI.3 Application à la loi binomiale | 29 |
| VI.4 Espérance d'un produit de variables aléatoires indépendantes | 30 |
| VII. Covariance de deux variables aléatoires..... | 31 |

Dans tout ce chapitre ($\Omega; P$) représente un espace probabilisé fini et E désigne un ensemble quelconque.

I VARIABLES ALÉATOIRES

I.1 Définition

Définition 1 : On appelle *variable aléatoire* sur Ω à valeurs dans E toute application \mathbb{X} de Ω dans E .

$$\mathbb{X} : \Omega \mapsto E.$$

Si $E \subset \mathbb{R}$, on parle de *variable aléatoire réelle* raccourci en v.a.r .

Remarques :

- une *variable aléatoire* est une *fonction* !
- Comme Ω est fini, une variable aléatoire $\mathbb{X} : \Omega \mapsto E$ prend un nombre fini de valeurs. L'ensemble de ces valeurs peut s'écrire $\mathbb{X}(\Omega) = \{x_1, \dots, x_n\}$ où $n \in \mathbb{N}^*$ et $\forall k \in \llbracket 1; n \rrbracket$, $x_k \in E$.

Exemple 1 : On tire au hasard une carte d'un jeu de 32 cartes.

L'univers est $\Omega = \{7\heartsuit, 7\spadesuit, 7\clubsuit, 8\heartsuit, \dots, A\heartsuit, A\spadesuit, A\clubsuit\}$.

On gagne :

- 10 euros si la carte tirée est l'as de cœur ;
- 5 euros si la carte tirée est un autre as ;
- 2 euros si la carte tirée est un valet, une dame ou un roi ;
- -1 euro sinon. On perd donc !

On peut modéliser le gain par une variable aléatoire

$$\mathbb{X} : \begin{cases} \Omega & \mapsto & \mathbb{R} \\ A\heartsuit & \rightarrow & 10 \\ A\spadesuit & \rightarrow & 5 \\ \dots & \rightarrow & \dots \\ 7\clubsuit & \rightarrow & -1 \end{cases}$$

En particulier, $\mathbb{X}(\Omega) = \{-1, 2, 5, 10\}$ et $\mathbb{X}^{-1}(5) = \{A\spadesuit, A\heartsuit, A\clubsuit\}$.

I.2 Image d'une variable aléatoire, Image réciproque

Définition 2 : Soit \mathbb{X} une variable aléatoire sur Ω .

- L'ensemble $\mathbb{X}(\Omega)$, image directe de Ω par \mathbb{X} , est appelé *support* de la variable aléatoire.
- Soit A une partie de E .
 - L'événement $\mathbb{X}^{-1}(A) = \{\omega \in \Omega / \mathbb{X}(\omega) \in A\}$, image réciproque de A par \mathbb{X} , est habituellement et abusivement noté $(\mathbb{X} \in A)$.

Dans le cas où A est un singleton $\{x\}$ où $x \in E$, on emploie plutôt les notations $(\mathbb{X} = x)$ au lieu de $(\mathbb{X} \in \{x\})$.

- La probabilité $P(\mathcal{X}^{-1}(A))$ de l'événement $(\mathcal{X} \in A)$ est notée $P(\mathcal{X} \in A)$, celle de $(\mathcal{X} = x)$, $P(\mathcal{X} = x)$.

Dans le cas de v.a.r, on note $P(\mathcal{X} \leq x)$ au lieu de $P(\mathcal{X} \in]-\infty; x])$.

Exemple 2 : Dans l'exemple (1),

- le support de la variable aléatoire \mathcal{X} est $\mathcal{X}(\Omega) = \{-1, 2, 5, 10\}$.
- $(\mathcal{X} \leq 12)$ est l'événement certain et $(\mathcal{X} = 7)$ est l'événement impossible.
- $(\mathcal{X} = 5)$ désigne l'événement : $\{A\heartsuit, A\spadesuit, A\clubsuit\}$.

On peut aussi le lire : « Gagner 5 euros », et on a $P(\mathcal{X} = 5) = P(\{A\heartsuit, A\spadesuit, A\clubsuit\}) = \frac{3}{32}$.

Exemple 3 : On lance deux fois un dé bien équilibré à 6 faces.

- L'univers de cette expérience aléatoire est l'ensemble des couples de $\llbracket 1; 6 \rrbracket^2$ donc $\Omega = \llbracket 1; 6 \rrbracket^2$.
- Le dé étant bien équilibré, la probabilité sur Ω est uniforme.
- On peut définir la fonction S sur Ω qui donne la somme des valeurs obtenues lors des deux lancers.

S est une v.a.r sur Ω de support $S(\Omega) = \llbracket 2; 12 \rrbracket$.

Ainsi,

$$S : \llbracket 1; 6 \rrbracket^2 \rightarrow \llbracket 2; 12 \rrbracket .$$

$$(i; j) \mapsto i + j$$

- $(S = 2) = \{(1; 1)\}$.
- $(S = 3) = \{(1; 2), (2; 1)\}$.
- ...
- $(S = 7) = \{(1; 6), (2; 5), (3; 4), (4; 3), (5; 2), (6; 1)\}$.
- ...
- $(S = 12) = \{(6; 6)\}$.

On peut alors remarquer que les événements $(S = k)_{2 \leq k \leq 12}$ forment un système complet d'événements de Ω .

Plus généralement,

Théorème 1 : Soit $\mathcal{X} : \Omega \mapsto E$ une variable aléatoire de support $\mathcal{X}(\Omega) = \{x_1, \dots, x_n\}$ où $n \in \mathbb{N}^*$ et $\forall k \in \llbracket 1; n \rrbracket, x_k \in E$.

Les événements $(\mathcal{X} = x_1), (\mathcal{X} = x_2), \dots, (\mathcal{X} = x_n)$ forment un système complet d'événements de Ω appelé le système complet d'événements associé à \mathcal{X} .

En particulier, si on pose pour tout entier $k \in \llbracket 1; n \rrbracket, p_k = P(\mathcal{X} = x_k)$ alors :

$$p_k \geq 0 \quad \text{et} \quad \sum_{k=1}^n p_k = 1.$$

Preuve : $\bigcup_{k=1}^n (\mathcal{X} = x_k) = \bigcup_{k=1}^n \mathcal{X}^{-1}(x_k) = \mathcal{X}^{-1} \left(\bigcup_{k=1}^n \{x_k\} \right) = \mathcal{X}^{-1}(\mathcal{X}(\Omega)) = \Omega$ et il est clair que $(\mathcal{X} = x_i) \cap (\mathcal{X} = x_j) = \emptyset$ si $i \neq j$.

I.3 Fonction d'une variable aléatoire

Une variable aléatoire réelle est une fonction.

- On peut donc, grâce à deux variables aléatoires réelles X, Y , définir leur somme $X + Y$, leur produit XY , le produit par un scalaire λX qui sont de nouvelles variables aléatoires.

Par exemple,

Définition 3 (Somme de variables aléatoires) : Soient X et Y deux variables aléatoires définies sur l'univers Ω et $\lambda \in \mathbb{R}$.

- On appelle somme des variables aléatoires, notée $Z = X + Y$, la variable aléatoire définie sur Ω par :

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

- On appelle produit des variables aléatoires, notée $Z = XY$, la variable aléatoire définie sur Ω par :

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

- Pour une v.a.r X , on appelle produit par un scalaire, notée $Z = \lambda X$, la variable aléatoire définie sur Ω par :

$$\forall \omega \in \Omega, Z(\omega) = \lambda X(\omega).$$

Exemple 4 :

- On lance cinq dés équilibrés et on compte la somme des nombres obtenus.

Soit X la variable aléatoire correspondant à cette somme.

Alors, on peut écrire X sous la forme $X = X_1 + \dots + X_5$ où, pour tout $k \in \llbracket 1; 5 \rrbracket$, X_k correspond au résultat du dé numéro k .

ATTENTION

$X \neq 5 \times X_1$. L'ensemble des valeurs prises par X est $\{5; 6; 7; 8; 9; \dots; 30\}$ (somme possible des 5 dés) alors que $5 \times X_1$ ne peut prendre que les valeurs 5, 10, 15, 20, 25 et 30.

- On lance 20 fois une pièce de monnaie et on note X la variable aléatoire comptant le nombre de pile obtenu.

On peut écrire la variable aléatoire X sous la forme $X = X_1 + X_2 + \dots + X_{20}$ où, pour tout $k \in \llbracket 1; 20 \rrbracket$, $X_k = 1$ si on a obtenu pile au $k^{\text{ème}}$ lancer et $X_k = 0$ sinon.

Exercice 1 : Lors d'une soirée au casino, Nadège décide de tester différents jeux : une fois la roulette et deux fois les machines à sous.

Elle note X la variable aléatoire correspondant au gain total remporté.

- 1 Pour faciliter l'étude, elle écrit $X = X_1 + X_2 + X_3$.

À quoi les variables aléatoires X_1, X_2 et X_3 peuvent-elles alors correspondre ?

- 2 Yvann souhaite écrire X sous la forme $X = X_1 + 2X_2$.

A-t-il raison ? Justifier.

Correction :

- 1 Par exemple, X_1 correspond au gain remporté à la roulette, X_2 à celui remporté en jouant la première fois aux machines à sous et X_3 au gain remporté en y jouant une seconde fois.

- 2 Yvann a tort : le gain remporté la première aux machines à sous n'est pas nécessairement le même que celui remporté à la seconde tentative.

- On peut également composer la variable aléatoire réelle $\mathcal{X} : \Omega \mapsto \mathbb{R}$ par une fonction réelle $f : \mathbb{R} \mapsto \mathbb{R}$. On obtient alors une nouvelle variable aléatoire

$$f \circ \mathcal{X} : \Omega \mapsto \mathbb{R}.$$

Exemple 5 : Soit $f : x \mapsto ax + b$ une fonction affine. Alors $f(\mathcal{X}) : \Omega \rightarrow \mathbb{R}$ notée

$$\omega \mapsto a\omega + b$$

$f(\mathcal{X}) = a\mathcal{X} + b.$

ATTENTION | Cette variable aléatoire est généralement notée $f(\mathcal{X})$ mais il faudra bien comprendre $f \circ \mathcal{X}$.

I.4 Loi d'une variable aléatoire

Une variable \mathcal{X} aléatoire étant donnée, on cherche à déterminer la probabilité d'obtenir un élément de $\mathcal{X}(\Omega)$.

Dans l'exemple (1) :

| | | | | |
|----------------------|---------------|---------------|----------------|----------------|
| x | -1 | 2 | 5 | 10 |
| $P(\mathcal{X} = x)$ | $\frac{1}{2}$ | $\frac{3}{8}$ | $\frac{3}{32}$ | $\frac{1}{32}$ |

Mais on peut aussi définir $P(\mathcal{X} \in \{2, 5, 10\}) = \frac{1}{2}$.

La loi de probabilité de \mathcal{X} , notée $P_{\mathcal{X}}$ associée à toute partie A de $\mathcal{X}(\Omega)$ la probabilité $P(\mathcal{X} \in A)$.

Définition 4 : Soit $\mathcal{X} : \Omega \mapsto E$ une variable aléatoire sur un espace probabilisé fini (Ω, P) .

On appelle *loi de probabilité* de \mathcal{X} , notée $P_{\mathcal{X}}$, l'application

$$P_{\mathcal{X}} : \mathcal{X}(\Omega) \rightarrow [0; 1]$$

$$x \mapsto P(\mathcal{X} = x).$$

Exercice 2 : Un sac contient 6 jetons : deux numérotés 1, trois numérotés 2 et un numéroté 3. Chaque jeton apparaît avec la même probabilité. On tire simultanément trois jetons. On note \mathcal{X} la somme des numéros portés sur les trois jetons.

Déterminer la loi de \mathcal{X} .

Correction : Par exemple, $5 = 1 + 1 + 3 = 1 + 2 + 2$

$$P(\mathcal{X} = 5) = \frac{\binom{2}{2}\binom{3}{0}\binom{1}{1} + \binom{2}{1}\binom{3}{2}\binom{1}{0}}{\binom{6}{3}} = \frac{7}{20}$$

La loi de probabilité de \mathcal{X} est donc donnée par :

| | | | | |
|----------------------|----------------|----------------|----------------|----------------|
| x | 4 | 5 | 6 | 7 |
| $P(\mathcal{X} = x)$ | $\frac{3}{20}$ | $\frac{7}{20}$ | $\frac{7}{20}$ | $\frac{3}{20}$ |

Proposition 2 : Soit $\mathcal{X} : \Omega \mapsto E$ une variable aléatoire de loi de probabilité $P_{\mathcal{X}}$.

L'application $P_{\mathcal{X}}$ définit une probabilité sur Ω et on a :

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

Preuve : Le **théorème (1)** assure déjà que $P_{\mathcal{X}}$ est une probabilité. Faisons comme si nous ne l'avions pas vu et montrons-le directement :

- On a bien $\forall A \subset \mathcal{X}(\Omega), P_{\mathcal{X}}(A) \in [0; 1]$ puisque $P_{\mathcal{X}}(A) = P(\mathcal{X} \in A)$.
- $P_{\mathcal{X}}(\mathcal{X}(\Omega)) = P(\mathcal{X} \in \mathcal{X}(\Omega)) = P(\Omega) = 1$
- Soient A, B sont deux parties disjointes de $\mathcal{X}(\Omega)$. On a $P_{\mathcal{X}}(A \cup B) = P(\mathcal{X} \in A \cup B)$.

Or, $(\mathcal{X} \in A \cup B) = \mathcal{X}^{-1}(A \cup B) = \mathcal{X}^{-1}(A) \cup \mathcal{X}^{-1}(B)$.

De plus, $\mathcal{X}^{-1}(A) \cap \mathcal{X}^{-1}(B) = \mathcal{X}^{-1}(A \cap B) = \mathcal{X}^{-1}(\emptyset) = \emptyset$ entraîne que les événements $\mathcal{X}^{-1}(A)$ et $\mathcal{X}^{-1}(B)$ sont incompatibles.

On a alors :

$$\begin{aligned} P_{\mathcal{X}}(A \cup B) &= P(\mathcal{X} \in A \cup B) \\ &= P(\mathcal{X}^{-1}(A) \cup \mathcal{X}^{-1}(B)) \\ &= P(\mathcal{X}^{-1}(A)) + P(\mathcal{X}^{-1}(B)) \\ &= P(\mathcal{X} \in A) + P(\mathcal{X} \in B) \\ &= P_{\mathcal{X}}(A) + P_{\mathcal{X}}(B) \end{aligned}$$

En conclusion, $P_{\mathcal{X}}$ est bien une probabilité.

Remarques :

- 1** Dit autrement : on a changé d'univers ! Au lieu de considérer l'univers entier, on travaille désormais avec un univers constitué seulement du support de \mathcal{X} .

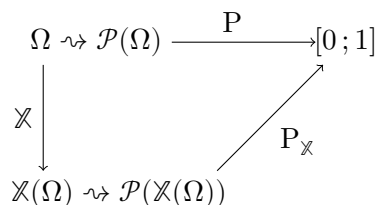


Figure XXX.1 – On associe une probabilité à un gain et non plus à un événement.

- 2** La donnée des $P(\mathcal{X} = x)$ pour tout $x \in \mathcal{X}(\Omega)$ suffit à définir $P_{\mathcal{X}}$.

On se contente en général de ces « germes de probabilités ».

Dans certains ouvrages, on définit même :

$$P_{\mathcal{X}} : \mathcal{X}(\Omega) \mapsto [0; 1]$$

au lieu de

$$P_{\mathcal{X}} : \mathcal{P}(\mathcal{X}(\Omega)) \mapsto [0; 1]$$

Exemple 6 : Reprenons l'exemple (3)

- $P(S = 2) = \frac{\text{card}(S = 2)}{\text{card}(\Omega)} = \frac{1}{36}$.
- $P(S = 3) = \frac{\text{card}(S = 3)}{\text{card}(\Omega)} = \frac{1}{18}$.
- ...
- $P(S = 7) = \frac{\text{card}(S = 7)}{\text{card}(\Omega)} = \frac{1}{6}$.
- ...
- $P(S = 12) = \frac{\text{card}(S = 12)}{\text{card}(\Omega)} = \frac{1}{36}$.

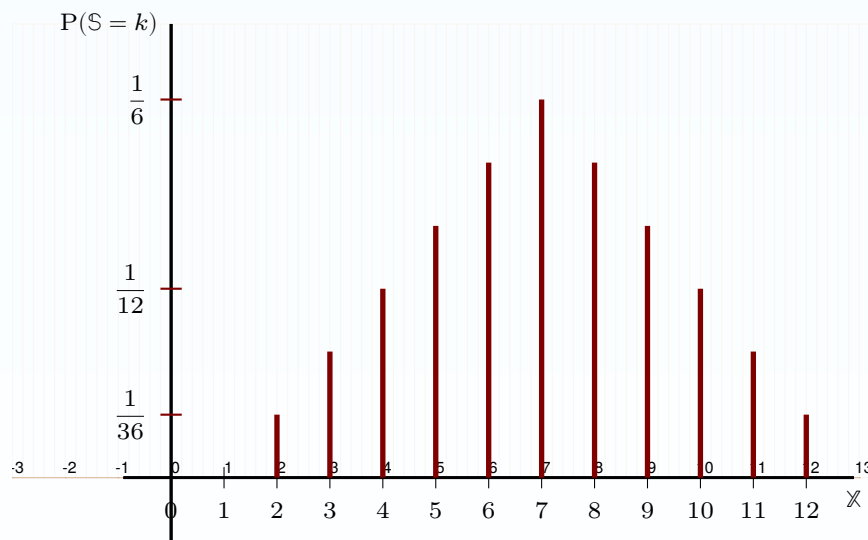
L'application $P_S : \llbracket 2 ; 12 \rrbracket \rightarrow [0 ; 1]$ est la loi de probabilité de S .

$$k \mapsto P(S = k)$$

On peut regrouper les résultats dans un tableau :

| k | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|------------|----------------|----------------|----------------|---------------|----------------|---------------|----------------|---------------|----------------|----------------|----------------|
| $P(S = k)$ | $\frac{1}{36}$ | $\frac{1}{18}$ | $\frac{1}{12}$ | $\frac{1}{9}$ | $\frac{5}{36}$ | $\frac{1}{6}$ | $\frac{5}{36}$ | $\frac{1}{9}$ | $\frac{1}{12}$ | $\frac{1}{18}$ | $\frac{1}{36}$ |

On peut aussi représenter la loi de S par un diagramme en bâtons :



Pour toute partie A de $\mathcal{S}(\Omega) = \llbracket 2 ; 12 \rrbracket$, on peut alors calculer la probabilité de l'événement $(S \in A)$.

Par exemple, $P(S \in \{4, 6, 11\}) = P(S = 4) + P(S = 6) + P(S = 11)$

$$= P_S(4) + P_S(6) + P_S(11) = \frac{1}{12} + \frac{5}{36} + \frac{1}{18} = \frac{10}{36} = \frac{5}{18}.$$

II ESPÉRANCE D'UNE VARIABLE ALÉATOIRE RÉELLE

Lorsqu'on s'intéresse à une variable aléatoire pouvant prendre un grand nombre de valeurs (et même dans les autres cas!), il peut être intéressant de donner, en plus de la loi de la variable qui ne sera pas toujours une donnée très lisible, des caractéristiques d'ensemble de cette loi, comme la moyenne des valeurs prises (pondérées par leur probabilité d'apparition).

Ces paramètres sont les mêmes que ceux qu'on étudie en statistiques, nous allons plus particulièrement nous intéresser à l'espérance (qui n'est autre que la moyenne évoquée plus haut, c'est un paramètre de position) et à l'écart-type (paramètre de dispersion, qui mesure la répartition des valeurs autour de l'espérance).

II.1 Définition

Définition 5 : Soit $\mathcal{X} : \Omega \mapsto \mathbb{R}$ une variable aléatoire **réelle** de support $\mathcal{X}(\Omega) = \{x_1, \dots, x_p\}$ où $p \in \mathbb{N}^*$ et $\forall k \in \llbracket 1; p \rrbracket, x_k \in \mathbb{R}$.

- On appelle *espérance* de \mathcal{X} , notée $E(\mathcal{X})$, le réel défini par :

$$E(\mathcal{X}) = \sum_{i=1}^p x_i P(\mathcal{X} = x_i) = \sum_{x \in \mathcal{X}(\Omega)} x P(\mathcal{X} = x).$$

- On dit que \mathcal{X} est une *variable centrée* si $E(\mathcal{X}) = 0$.

Remarques :

- 1 Dans les démonstrations et les résultats plus théoriques, on reviendra souvent à

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

- 2 L'espérance correspond à la moyenne des valeurs prises par \mathcal{X} , pondérée par leur probabilité. C'est donc un indicateur de position.
- 3 Plus particulièrement, chaque valeur de x , x décrivant $\mathcal{X}(\Omega)$, s'y trouve comptabilisé en proportion de sa probabilité d'occurrence $P(\mathcal{X} = x)$.

Ainsi, plus $P(\mathcal{X} = x)$ est proche de 1, plus x a d'importance dans le calcul de $E(\mathcal{X})$.

Exemple 7 : Reprenons l'exemple l' **exemple (1)**

| | | | | |
|----------------------|---------------|---------------|----------------|----------------|
| x | -1 | 2 | 5 | 10 |
| $P(\mathcal{X} = x)$ | $\frac{1}{2}$ | $\frac{3}{8}$ | $\frac{3}{32}$ | $\frac{1}{32}$ |

L'espérance de \mathcal{X} est $E(\mathcal{X}) = (-1) \times \frac{1}{2} + 2 \times \frac{3}{8} + 5 \times \frac{3}{32} + 10 \times \frac{1}{32} = \frac{33}{32} \simeq 1,03$ euros .

C'est la somme qu'on peut espérer gagner par partie en moyenne, en jouant de nombreuses fois.

Exemple 8 : Et si nous reprenons l' **exemple (3)** ,

$$E(\mathcal{S}) = 2 \times \frac{1}{36} + 3 \times \frac{1}{12} + 4 \times \frac{1}{12} + \dots + 11 \times \frac{1}{18} + 12 \times \frac{1}{36} = 7.$$

II.2 Linéarité

Proposition 3 : Soient $X : \Omega \mapsto \mathbb{R}$ et $Y : \Omega \mapsto \mathbb{R}$ deux variables aléatoires réelles sur Ω fini.

Linéarité : $\forall \alpha, \beta \in \mathbb{R}, E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$.

En particulier, $E(\alpha X + \beta) = \alpha E(X) + \beta$

Positivité : Si $X \geq 0$, alors $E(X) \geq 0$.

Croissance : Si $X \leq Y$, alors $E(X) \leq E(Y)$.

Inégalité triangulaire : $|E(X)| \leq E(|X|)$.

ATTENTION

La notation $X \geq 0$ est un tantinet abusive et signifie que la fonction X prend des valeurs positives sur Ω i.e. $\forall \omega \in \Omega, X(\omega) \geq 0$.

De même $X \leq Y \iff \forall \omega \in \Omega, X(\omega) \leq Y(\omega)$.

Preuve :

Linéarité : Soient $\alpha, \beta \in \mathbb{R}$.

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

Positivité : Si $X \geq 0$, alors $E(X)$ est une somme de nombres positifs donc ...

Croissance : $X \leq Y \iff Y - X \geq 0 \implies E(Y - X) \geq 0$
 $\implies E(Y) + E(-X) \geq 0 \implies E(Y) - E(X) \geq 0$.

Tout avait déjà été fait dans les points précédents.

Inégalité triangulaire : Par définition, de la valeur absolue, $\forall \omega \in \Omega, -|X(\omega)| \leq X(\omega) \leq |X(\omega)|$.

Par croissance et linéarité de l'espérance, on a donc :

$$-E(|X|) \leq E(X) \leq E(|X|).$$

Exercice 3 : Un forain a construit un appareil de jeu contenant six boules blanches et trois boules rouges. Lorsqu'on introduit un jeton dans l'appareil, trois boules prises au hasard tombent dans un panier. Si les trois boules sont rouges, le joueur gagne 100 € ; si deux des boules sont rouges, il gagne 15€, si une seule est rouge, il gagne un lot de 5 €. Le prix du jeton est fixé à 8 €.

1 Soit X la variable aléatoire désignant la somme gagnée par le joueur.

Déterminer la loi de X , et calculer son espérance. En déduire le gain moyen du forain.

2 L'appareil ne s'avérant pas suffisamment rentable, le forain envisage deux solutions : augmenter de 1 euro le pris du jeton ou bien ajouter une boule blanche dans l'urne. Quelle est la solution la plus rentable pour le forain ?

Correction :

1 On considère dans un premier temps la somme gagnée par le joueur, **sans tenir compte de l'achat du billet.**

a $(X = 100)$ est l'événement « les trois boules sont rouges ».

$$P(X = 100) = \frac{\binom{3}{3}\binom{6}{0}}{\binom{9}{3}} = \frac{1}{84}$$

b $(X = 15)$ est l'événement « deux boules sont rouges ».

$$P(X = 15) = \frac{\binom{3}{2}\binom{6}{1}}{\binom{9}{3}} = \frac{18}{84} = \frac{3}{14}$$

c $(X = 5)$ est l'événement « une boule est rouge ».

$$P(X = 5) = \frac{\binom{3}{1}\binom{6}{2}}{\binom{9}{3}} = \frac{45}{84} = \frac{15}{28}$$

d $(X = 0)$ est l'événement « aucune boule n'est rouge ».

$$P(X = 0) = \frac{\binom{3}{0}\binom{6}{3}}{\binom{9}{3}} = \frac{20}{84} = \frac{5}{21}$$

La loi de probabilité de X est donnée par

| | | | | |
|------------|----------------|----------------|-----------------|----------------|
| x | 100 | 15 | 5 | 0 |
| $P(X = x)$ | $\frac{1}{84}$ | $\frac{3}{14}$ | $\frac{15}{28}$ | $\frac{5}{21}$ |

Son espérance est

$$E(X) = 100 \times \frac{1}{84} + 15 \times \frac{3}{14} + 5 \times \frac{15}{28} + 0 \times \frac{5}{21}$$

$$E(X) = \frac{85}{12} \simeq 7,08\text{€}.$$

Si on note X' la variable aléatoire donnant le gain net du joueur (après achat du billet à 8€), on a

$$X' = X - 8$$

D'après les propriétés de l'espérance, $E(X') = E(X) - 8$.

$$E(X') = -\frac{11}{12}.$$

C'est la somme perdue en moyenne par partie par le joueur, en supposant qu'il joue de nombreuses fois.

Le gain moyen du forain est donc $\frac{11}{12}$, soit un peu moins de 0,92€.

2 a D'après la linéarité de l'espérance, si le forain augmente le prix du ticket d'un euro, son espérance de gain sera augmentée d'un euro soit à 1,92 €.

b Si on ajoute une boule blanche dans l'urne, on reprend tous les calculs en notant Y la nouvelle variable aléatoire du gain (brut) du joueur :

| | | | | |
|------------|--|---|---|---|
| x | 100 | 15 | 5 | 0 |
| $P(Y = x)$ | $\frac{\binom{3}{3}\binom{7}{0}}{\binom{10}{3}} = \frac{1}{120}$ | $\frac{\binom{3}{2}\binom{7}{1}}{\binom{10}{3}} = \frac{21}{120}$ | $\frac{\binom{3}{1}\binom{7}{2}}{\binom{10}{3}} = \frac{63}{120}$ | $\frac{\binom{3}{0}\binom{7}{3}}{\binom{10}{3}} = \frac{35}{120}$ |

L'espérance de gain brute est $E(Y) = \frac{73}{12}$.

On en déduit l'espérance de gain net : $E(Y') = E(Y) - 8 = -\frac{23}{12}$.

Le gain moyen du forain est donc $\frac{23}{12} \simeq 1,92\text{€}$.

Les deux solutions sont identiques.

II.3 Formule de transfert

La **proposition (3)** permet déjà d'écrire que $E(aX + b) = aE(X) + b$ pour toute fonction affine $x \mapsto ax + b$. On peut aller plus loin :

Théorème 4 : Soit $X : \Omega \mapsto \mathbb{R}$ une variable aléatoire **réelle** de support $X(\Omega) = \{x_1, \dots, x_p\}$ où $p \in \mathbb{N}^*$ et $\forall k \in \llbracket 1; p \rrbracket$, $x_k \in \mathbb{R}$.

Soit $f : X(\Omega) \mapsto \mathbb{R}$ une application à valeurs réelles définie sur $X(\Omega)$.

Alors :

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

Preuve : Par définition, $E(f(X)) = \sum_{\omega \in \Omega} P(\{\omega\}) f(X(\omega))$.

On regroupe les ω tels que $X(\omega) = x_i$ en posant $\Omega_i = \{\omega / X(\omega) = x_i\}$. Les Ω_i forment alors un système complet d'événements de Ω et on a $P(X = x_i) = P\left(\bigcup_{\omega \in \Omega_i} \{\omega\}\right) = \sum_{\omega \in \Omega_i} P(\{\omega\})$.

$$E(f(X)) = \sum_{\omega \in \Omega} P(\{\omega\}) (f \circ X)(\omega) = \sum_{\omega \in \Omega} P(\{\omega\}) f(X(\omega))$$

Or, $\Omega = \bigcup_{i=1}^p \Omega_i$ (réunion d'événements disjoints),

$$\begin{aligned} &= \sum_{i=1}^p \left(\sum_{\omega \in \Omega_i} P(\{\omega\}) f(X(\omega)) \right) = \sum_{i=1}^p \left(\sum_{\omega \in \Omega_i} P(\{\omega\}) f(x_i) \right) \\ &= \sum_{i=1}^p \left(f(x_i) \sum_{\omega \in \Omega_i} P(\{\omega\}) \right) = \sum_{i=1}^p f(x_i) P(\Omega_i) \\ &= \sum_{i=1}^p f(x_i) P(X = x_i). \end{aligned}$$

Remarque : Cette formule permet d'éviter le calcul de la loi de $f(X)$ qui peut être compliqué.

Exemple 9 : On reprend l'**exemple (3)** où S est la variable aléatoire qui donne la somme des valeurs obtenues lors des deux lancers.

Par exemple, on a alors :

$$E(S^2) = \sum_{s=2}^{12} s^2 \times P(S = s) = 2^2 \times \frac{1}{36} + 3^2 \times \frac{1}{18} + \dots + 12^2 \times \frac{1}{36} = \frac{329}{6}.$$

III VARIANCE ET ÉCART-TYPE D'UNE VARIABLE ALÉATOIRE RÉELLE

III.1 Définitions

Définition 6 (Variance) : Soit $X : \Omega \mapsto \mathbb{R}$ une variable aléatoire réelle.

On appelle *variance* de X , notée $V(X)$, le réel :

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

Remarque : La variance est un *indicateur de dispersion* de la variable aléatoire autour de son espérance.

Proposition 5 (Formule de König-Huygens) : Soit $X : \Omega \mapsto \mathbb{R}$ une variable aléatoire réelle.

Alors :

$$V(X) = E(X^2) - E(X)^2.$$

Preuve : Posons $m = E(X)$.

Comme $(X - m)^2 = X^2 - 2mX + m^2$, alors

$$E((X - m)^2) = E(X^2 - 2mX + m^2) = E(X^2) - 2mE(X) + m^2 = E(X^2) - m^2.$$

Proposition 6 : $\forall \alpha, \beta \in \mathbb{R}, \quad V(\alpha X + \beta) = \alpha^2 V(X).$

Preuve : Une fois n'est pas coutume, il suffit de développer :

$$\begin{aligned} V(\alpha X + \beta) &= E\left((\alpha X + \beta - E(\alpha X + \beta))^2\right) \\ &= E\left((\alpha X + \beta - \alpha E(X) - \beta)^2\right) \\ &= E(\alpha^2 (X - E(X))^2) \\ &= \alpha^2 E((X - E(X))^2) \\ &= \alpha^2 V(X). \end{aligned}$$

Comme $(X - m)^2 \geq 0$, la variance $V(X) = E((X - m)^2)$ est positive. Ce qui permet la définition suivante :

Définition 7 (Écart-type) : Soit X une variable aléatoire réelle.

On appelle *écart-type* de X , noté $\sigma(X)$, le réel :

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

Remarque : Une v.a.r $Z : \Omega \mapsto \mathbb{R}$ est dite *centrée* si $E(Z) = 0$ et *réduite* si $\sigma(Z) = 1$.

Par exemple, si $X : \Omega \mapsto \mathbb{R}$ est une v.a.r telle que $\sigma(X) > 0$ alors $Z = \frac{X - E(X)}{\sigma(X)}$ est centrée réduite

III.2 Inégalité de Markov et Bienaymé-Tchebychev

Théorème 7 (Inégalité de Markov [1]) : Soit $Z : \Omega \mapsto \mathbb{R}$ une variable aléatoire réelle **positive**.

Alors :

$$\forall a > 0, P(Z \geq a) \leq \frac{E(Z)}{a}.$$

Preuve : Soient $Z : \Omega \mapsto \mathbb{R}$ une variable aléatoire réelle positive et $a > 0$.

$$E(Z) = \sum_{z \in Z(\Omega)} zP(Z = z) = \sum_{\substack{z \in Z(\Omega) \\ z > a}} zP(Z = z) + \sum_{\substack{z \in Z(\Omega) \\ z < a}} zP(Z = z).$$

Or, Z est positive i.e. $\forall z \in Z(\Omega), z \geq 0$.

Donc, $\sum_{\substack{z \in Z(\Omega) \\ z < a}} zP(Z = z) \geq 0$.

Ainsi, $E(Z) \geq \sum_{\substack{z \in Z(\Omega) \\ z \geq a}} zP(Z = z) \geq \sum_{\substack{z \in Z(\Omega) \\ z \geq a}} aP(Z = z) = a \underbrace{\sum_{\substack{z \in Z(\Omega) \\ z \geq a}} P(Z = z)}_{P(Z \geq a)} \geq aP(Z \geq a)$.

D'où le résultat en divisant les deux membres par $a > 0$.

L'inégalité de Markov donne une majoration de la probabilité qu'une variable aléatoire réelle à valeurs positives soit supérieure ou égale à une constante positive.

Théorème 8 (Inégalité de Bienaymé-Tchebychev) : Soit X une variable aléatoire réelle d'espérance m et d'écart-type σ .

$$\forall \varepsilon > 0, P(|X - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Preuve : Soit X une variable aléatoire réelle d'espérance m et de variance $V(X) = \sigma^2$.

Soit $\varepsilon > 0$. On pose très naturellement $Z = (X - m)^2$ qui est une v.a.r positive et $a = \varepsilon^2$.

[1]. Andreï Andreïevitch Markov (1856-1922) est un mathématicien russe.

Il est considéré comme le fondateur de la théorie des *processus stochastiques*.

Un peu d'histoire : En 1874, il entra à la faculté de physique et de mathématiques de l'université impériale de Saint-Petersbourg. Il participa au séminaire dirigé par Korkine et Zolotarev, et assista à des conférences de Tchebychev, le directeur du département de mathématiques. Celles-ci étaient particulièrement stimulantes pour Markov, car Tchebychev encourageait souvent une atmosphère de recherche en posant de nouvelles questions et problèmes pour que ses étudiants y réfléchissent.

D'après le **théorème (7)**,

$$P(Z \geq a) \leq \frac{E(Z)}{a} \iff P((X - m)^2 \geq \varepsilon^2) \leq \frac{E((X - m)^2)}{\varepsilon^2}.$$

Or, $(X - m)^2 \geq \varepsilon^2 \iff |X - m| \geq \varepsilon$ et, $E((X - m)^2) = V(X) = \sigma^2$ par définition.

Donc, en remplaçant,

$$P(|X - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Remarque : Cette formule fournit un majorant de l'erreur quand on estime que $X \in]m - \varepsilon, m + \varepsilon[$.

Pour tout $\varepsilon > 0$, $P(|X - m| \geq \varepsilon)$ est la probabilité pour que X prenne des valeurs éloignées de $E(X)$ d'au moins ε . Cette probabilité est faible dès que $V(X)$ est petit et que ε est grand.

L'inégalité de Bienaymé-Tchebychev vise donc à montrer que la variable aléatoire X prend des valeurs proches de $E(X)$ avec une grande probabilité, mais elle donne, en général une majoration assez grossière.

Un peu d'histoire : En théorie des probabilités, l'inégalité de Bienaymé-Tchebychev, est une inégalité de concentration permettant de montrer qu'une variable aléatoire prendra avec une faible probabilité une valeur relativement lointaine de son espérance.

Ce résultat s'applique dans des cas très divers, nécessitant la connaissance de peu de propriétés (seules l'espérance et la variance doivent être connues), et permet de démontrer la loi faible des grands nombres.

Ce théorème doit son nom aux mathématiciens Irénée-Jules Bienaymé, qui fut le premier à le formuler, et Pafnouti Tchebychev qui le démontra.

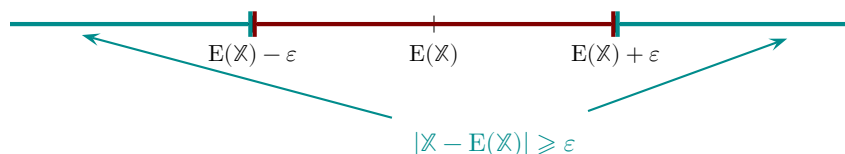


Figure XXX.2 – L'inégalité de Bienaymé-Tchebychev confirme que la variance est un indicateur de dispersion.

IV LOIS USUELLES

IV.1 Loi uniforme

Définition 8 : Soient $X : \Omega \mapsto \mathbb{R}$ une variable aléatoire **réelle** et $A = \{x_1, \dots, x_n\}$ où $n \in \mathbb{N}^*$ et $\forall k \in \llbracket 1; n \rrbracket$, $x_k \in \mathbb{R}$, un sous-ensemble fini non vide de \mathbb{R} .

On dit que X suit la *loi uniforme sur A* et on note $X \rightsquigarrow \mathcal{U}(A)$ lorsque :

- $X(\Omega) = A$.
- $\forall k \in \llbracket 1; n \rrbracket$, $P(X = x_k) = \frac{1}{n}$.

Remarques :

— C'est en fait la loi d'équiprobabilité sur A .

| | | | | |
|------------|---------------|---------------|---------|---------------|
| x | x_1 | x_2 | \dots | x_n |
| $P(X = x)$ | $\frac{1}{n}$ | $\frac{1}{n}$ | \dots | $\frac{1}{n}$ |

— On retrouve souvent dans les exercices la loi uniforme sur $\llbracket 1, n \rrbracket$:

| | | | | |
|------------|---------------|---------------|---------|---------------|
| x | 1 | 2 | \dots | n |
| $P(X = x)$ | $\frac{1}{n}$ | $\frac{1}{n}$ | \dots | $\frac{1}{n}$ |

Par exemple, le résultat d'un lancer de dé supposé honnête suit la loi $\mathcal{U}(\llbracket 1, 6 \rrbracket)$.

Proposition 9 : Si $X \rightsquigarrow \mathcal{U}(\llbracket 1, n \rrbracket)$ alors :

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

Preuve :

$$E(X) = \frac{1}{n} \times 1 + \frac{1}{n} \times 2 + \dots + \frac{1}{n} \times n = \frac{1}{n}(1 + 2 + \dots + n) = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Et,

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= \left(\frac{1}{n} \times 1^2 + \frac{1}{n} \times 2^2 + \dots + \frac{1}{n} \times n^2 \right) - \left(\frac{n+1}{2} \right)^2 \\ &= \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} \\ &= \frac{n^2-1}{12} \end{aligned}$$

IV.2 Loi de Bernoulli

Définition 9 : Soient $X : \Omega \mapsto \mathbb{R}$ une variable aléatoire **réelle** et $p \in [0; 1]$.

On dit que X suit la *loi de Bernoulli de paramètre p* et on note $X \rightsquigarrow \mathcal{B}(p)$ lorsque :

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

On pose souvent $q = 1 - p$.

Remarques :

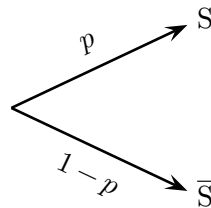
— En résumé :

| | | |
|------------|---------|-----|
| x | 0 | 1 |
| $P(X = x)$ | $1 - p$ | p |

— Cette loi modélise le succès ($X = 1$) ou l'échec ($X = 0$) à une expérience aléatoire donnée. p est la probabilité de succès.

On dira donc qu'une expérience aléatoire est une épreuve de Bernoulli lorsqu'elle admet seulement deux issues possibles, moralement appelées échec et succès.

— L'arbre correspondant s'appelle une *épreuve de Bernoulli*.



Exemple 10 :

- Le jeu du pile ou face : On considère par exemple comme succès « obtenir pile » et comme échec « obtenir face ».

$$p = \frac{1}{2}.$$

- On lance un dé et on considère par exemple comme succès « obtenir un six » et comme échec « ne pas obtenir un six ».

$$p = \frac{1}{6}.$$

- Interroger une personne dans la rue et lui demander si elle est gauchère est une épreuve de Bernoulli de succès S « la personne est gauchère ».

$$p \simeq 0,13.$$

Proposition 10 : Si $X \rightsquigarrow \mathcal{B}(p)$ alors :

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

Preuve : Il suffit d'appliquer simplement la définition :

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

$$= 1 \times p + 0 \times (1 - p)$$

$$= p.$$

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

$$= p \times (1 - p)^2 + (1 - p) \times p^2$$

$$= p(1 - p)((1 - \cancel{X}) + \cancel{X})$$

$$= p(1 - p).$$

IV.3 Loi binomiale

Définition 10 : Soit $n \in \mathbb{N}^*$.

On appelle *schéma de Bernoulli* (d'ordre n) toute répétition de n épreuves de Bernoulli identiques et indépendantes.

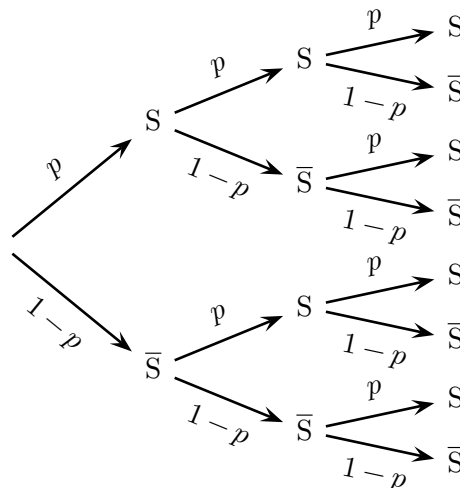


Figure XXX.3 – Schéma de Bernoulli d'ordre 3.

Exemple 11 :

- Soit n un entier naturel non nul.
Répéter n fois chacune des expériences de l' [exemple \(10\)](#) est un schéma de Bernoulli (d'ordre n).
- On peut aussi considérer une urne opaque dans laquelle ont été placées une boule verte et deux boules bleues, toutes indiscernables au toucher.
 - 1 On prélève alors une boule dans cette urne, on note sa couleur, puis on remet la boule dans l'urne.
 - 2 On répète ainsi dix fois l'expérience et on s'intéresse aux boules bleues obtenues.
 - 3 Chaque tirage est une épreuve de Bernoulli de succès S : « La boule est bleue » dont la probabilité est $p = \frac{2}{3}$.
 - 4 Comme les dix tirages se font avec remise, les tirages sont identiques et indépendants : on a bien un schéma de Bernoulli (d'ordre $n = 10$).

La variable aléatoire \mathcal{X} comptant les succès peut prendre toutes les valeurs entre 0 et n donc $\mathcal{X}(\Omega) = \llbracket 0; n \rrbracket$.

Soit $k \in \llbracket 0; n \rrbracket$. L'événement $(\mathcal{X} = k)$ est réalisé si on obtient exactement k succès et $n - k$ échecs. La probabilité de chacune de ces branches est $p^k q^{n-k}$.

Enfin, lors de ces n épreuves, il y a autant de manières d'obtenir k succès (et donc $n - k$ échecs) que d'anagrammes d'un mot de n lettres contenant k lettres S (pour succès) et $n - k$ lettres E

(pour échec) *i.e.* $\frac{n!}{k!(n-k)!} = \binom{n}{k}$.

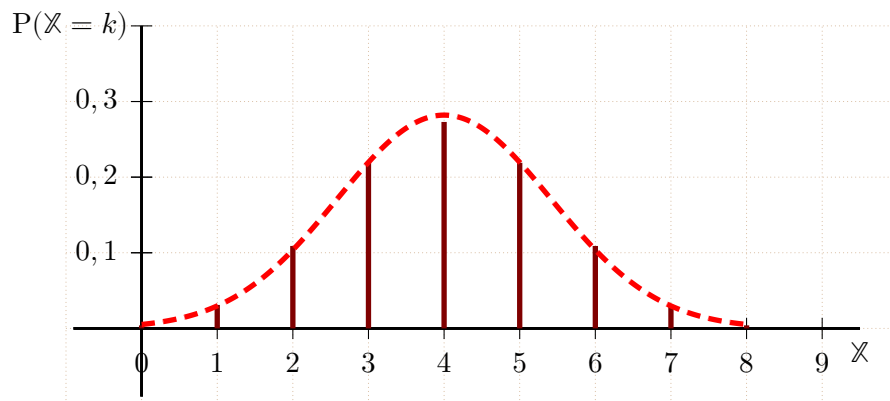
Finalement, $\forall k \in \llbracket 0; n \rrbracket, P(\mathcal{X} = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Remarque : Le coefficient binomial $\binom{n}{k}$ représente donc aussi le nombre de chemins à k succès parmi un schéma de Bernoulli d'ordre n .

Définition II : Soient $X : \Omega \mapsto \mathbb{R}$ une variable aléatoire **réelle**, $n \in \mathbb{N}$ tel que $n \geq 2$ et $p \in [0; 1]$.

On dit que X suit la *loi binomiale de paramètres n et p* et on note $X \rightsquigarrow \mathcal{B}(n, p)$ lorsque :

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



$$P(X = k) = \binom{8}{k} 0,5^k 0,5^{8-k} = \binom{8}{k} 0,5^8.$$

| | | | | | | | | | |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| X | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $P(X = k)$ | 0,004 | 0,031 | 0,109 | 0,219 | 0,273 | 0,219 | 0,109 | 0,031 | 0,004 |

Figure XXX.4 – Représentation (symétrique) de $\mathcal{B}(8; 0,5)$.

Remarque : On peut vérifier que :

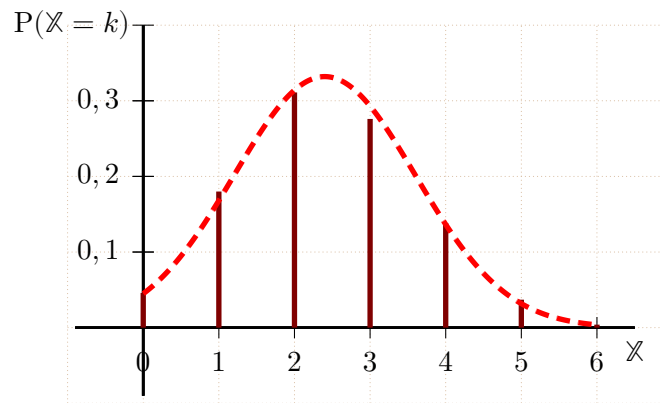
$$\begin{aligned} \sum_{k=0}^n P(X = k) &= \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} \\ &= [p + (1 - p)]^n \\ &= 1. \end{aligned}$$

Exemple 12 (Classique) : Une urne contient des boules blanches et noires avec une proportion p de boules blanches.

On pioche n boules avec remise et on note X la variable aléatoire donnant le nombre de boules blanches tirées après les n tirages.

X suit une loi $\mathcal{B}(n, p)$.

Exercice 4 : Une personne sur 1500 est daltonienne. Combien de personnes doit-on choisir pour être sûr à 95% d'avoir au moins un daltonien ?



$$P(X = k) = \binom{6}{k} 0,4^k 0,6^{6-k}.$$

| X | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|------------|-------|-------|-------|-------|-------|-------|-------|
| $P(X = k)$ | 0,046 | 0,187 | 0,311 | 0,276 | 0,138 | 0,037 | 0,004 |

Figure XXX.5 – Représentation (asymétrique) de $\mathcal{B}(6; 0,4)$.

Correction : On considère n personnes, chacune ayant une probabilité de $\frac{1}{1500}$ d'être daltonienne, indépendamment du daltonisme des autres personnes choisies.

Le nombre de personnes daltoniennes X dans le groupe suit donc la loi binomiale de paramètres n et $\frac{1}{1500}$: $X \rightsquigarrow \mathcal{B}\left(n, \frac{1}{1500}\right)$.

La probabilité d'avoir au moins un daltonien dans le groupe est :

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \left(\frac{1499}{1500}\right)^n$$

$$\begin{aligned} P(X \geq 1) \geq 0,95 &\Leftrightarrow 1 - \left(\frac{1499}{1500}\right)^n \geq 0,95 \\ &\Leftrightarrow \left(\frac{1499}{1500}\right)^n \leq 0,05 \\ &\Leftrightarrow n \geq \frac{\ln(0,05)}{\ln 1499 - \ln 1500} \\ &\Leftrightarrow n \geq 4493. \end{aligned}$$

Proposition II : Soit X une variable aléatoire suivant une loi binomiale de paramètres n et p . Alors,

$$E(X) = np.$$

Preuve : Soit X une variable aléatoire suivant une loi binomiale de paramètre n et p .

Calcul direct :

$$E(X) = \sum_{k=0}^n kP(X = k) = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k}$$

D'après la formule du « capitaine »,

$$= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} = n \sum_{k=1}^n \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

Un petit changement de variable sur les indices,

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

On factorise par p ,

$$= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

Enfin, on reconnaît le développement de $(p + (1-p))^{n-1} = 1^{n-1} = 1$:

$$\text{Donc, } E(X) = np.$$

X vue comme une somme de v.a.r : Appelons X_1, X_2, \dots, X_n les variables aléatoires associées à chacune des épreuves de Bernoulli.

D'après la **proposition (3)**, on a :

$$E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = p + p + \dots + p = np.$$

Remarque : Il est à ce stade impossible, avec nos connaissances, de calculer la variance d'une loi binomiale (confer le **corollaire (19.1)** ...).

Exercice 5 : On considère qu'à un concours, un candidat a 20% de chances de réussir. On prend un groupe de 25 candidats au hasard.

- 1] Quelle est la probabilité qu'au moins un candidat réussisse ?
- 2] Quelle est la probabilité qu'au plus deux candidats réussissent ?
- 3] Quelle est la probabilité que dix candidats réussissent ?
- 4] Calculer le nombre moyen de candidats qui réussissent sur 25 qui passent le concours.

Correction : Le nombre de candidats X réussissant suit la loi binomiale de paramètres 23 et 0,2.

$$X \rightsquigarrow B(25; 0,2)$$

En effet, la probabilité de réussite de chaque candidat est la même : 0,2, et la réussite d'un candidat est indépendante de la réussite des autres.

$$P(X = k) = \binom{25}{k} (0,2)^k (0,8)^{25-k}.$$

- 1] $P(X \geq 1) = 1 - P(X = 0) = 1 - (0,8)^{25} \simeq 99,6\%$.

2

$$\begin{aligned}
 P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\
 &= (0,8)^{25} + \binom{25}{1}(0,2)^1(0,8)^{25-1} + \binom{25}{2}(0,2)^2(0,8)^{25-2} \\
 &\simeq 9,82\%.
 \end{aligned}$$

3 $P(X = 10) = \binom{25}{10}(0,2)^{10}(0,8)^{25-10} \simeq 1,18\%.$

4 D'après le cours, l'espérance de la loi binomiale $\mathcal{B}(25; 0,2)$ est $E(X) = 25 \times 0,2 = 5.$



COUPLE DE VARIABLES ALÉATOIRES

Introduction

Une urne contient 6 boules numérotées de 1 à 6.

- On en tire une et on note son numéro X .
- On retire alors toutes les boules dont le numéro est strictement supérieur à X .
- On tire une nouvelle boule dont le numéro est Y .

Le support de chacune de ces variables aléatoires réelles est $X(\Omega) = Y(\Omega) = \llbracket 1; 6 \rrbracket$. Ω désignant l'ensemble des tirages sans remise de l'urne.

Ω est donc l'ensemble des 2-arrangements de l'ensemble formé des 6 boules de l'urne.

Le couple (X, Y) est une nouvelle variable aléatoire de support $\llbracket 1; 6 \rrbracket \times \llbracket 1; 6 \rrbracket$.

On veut déterminer la loi de (X, Y) .

Cela revient à chercher, pour tout couple $(i; j) \in \llbracket 1; 6 \rrbracket^2$,

$$P((X = i) \cap (Y = j)).$$

$$\text{Or, } P((X = i) \cap (Y = j)) = P(X = i) \times P_{(X=i)}(Y = j) = \begin{cases} \frac{1}{6} \times \frac{1}{i} & \text{si } j \leq i \\ 0 & \text{si } j > i \end{cases}.$$

| $X \setminus Y$ | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|-----------------|------------------|------------------|------------------|------------------|------------------|----------------|---------------|
| 1 | $\frac{1}{6}$ | 0 | 0 | 0 | 0 | 0 | $\frac{1}{6}$ |
| 2 | $\frac{1}{12}$ | $\frac{1}{12}$ | 0 | 0 | 0 | 0 | $\frac{1}{6}$ |
| 3 | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | 0 | 0 | 0 | $\frac{1}{6}$ |
| 4 | $\frac{1}{24}$ | $\frac{1}{24}$ | $\frac{1}{24}$ | $\frac{1}{24}$ | 0 | 0 | $\frac{1}{6}$ |
| 5 | $\frac{1}{30}$ | $\frac{1}{30}$ | $\frac{1}{30}$ | $\frac{1}{30}$ | $\frac{1}{30}$ | 0 | $\frac{1}{6}$ |
| 6 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| Total | $\frac{49}{120}$ | $\frac{29}{120}$ | $\frac{19}{120}$ | $\frac{37}{360}$ | $\frac{11}{180}$ | $\frac{1}{36}$ | 1 |

Commentaires :

- Ce tableau donne la loi, dite *conjointe*, du couple (X, Y) .
- Les marges (obtenues en sommant les valeurs) comportent les lois de X et de Y . On les appellera donc les lois *marginales*.

V.1 Lois conjointe et marginale

Définition 12 : Soient \mathcal{X} et \mathcal{Y} deux variables aléatoires définies sur un univers fini Ω à valeurs, respectivement, dans des ensembles E et F quelconques.

- Le couple $(\mathcal{X}, \mathcal{Y})$ est également une variable aléatoire définie sur Ω à valeurs dans $E \times F$.
- La loi de probabilité du couple $(\mathcal{X}, \mathcal{Y})$ est appelée *loi conjointe* de \mathcal{X} et \mathcal{Y} .
- Les lois de \mathcal{X} et \mathcal{Y} sont appelées *lois marginales* du couple $(\mathcal{X}, \mathcal{Y})$.

Proposition 12 (Lois marginales et loi conjointe) : Soient $\mathcal{X} : \Omega \mapsto E$ et $\mathcal{Y} : \Omega \mapsto F$ deux variables aléatoires sur des ensembles E et F quelconques.

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

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

Moralité : Si on connaît la loi conjointe d'un couple de variables aléatoires $(\mathcal{X}, \mathcal{Y})$, on peut déterminer les lois marginales.

Remarque : En revanche, la réciproque est fautive : si on connaît les lois marginales, on ne peut déterminer la loi conjointe.

Preuve : Soit $\mathcal{X} : \Omega \mapsto E$ une v.a. où E est un ensemble quelconque.

Soit $\mathcal{Y} : \Omega \mapsto F$ une v.a. de support $\mathcal{Y}(\Omega) = \{y_1, \dots, y_n\}$ où $n \in \mathbb{N}^*$ et, $\forall k \in \llbracket 1; n \rrbracket$, $y_k \in F$.

$\forall y \in \mathcal{Y}(\Omega)$, les événements $(\mathcal{Y} = y)$ forment un système complet d'événements de Ω .

$\forall x \in \mathcal{X}(\Omega)$, on peut écrire

$$(\mathcal{X} = x) = (\mathcal{X} = x) \cap \Omega = (\mathcal{X} = x) \cap \left(\bigcup_{y \in \mathcal{Y}(\Omega)} (\mathcal{Y} = y) \right) = \bigcup_{y \in \mathcal{Y}(\Omega)} ((\mathcal{X} = x) \cap (\mathcal{Y} = y)).$$

Comme les événements $(\mathcal{X} = x) \cap (\mathcal{Y} = y)$ sont deux à deux disjoints, on a alors :

$$P(\mathcal{X} = x) = P\left(\bigcup_{y \in \mathcal{Y}(\Omega)} ((\mathcal{X} = x) \cap (\mathcal{Y} = y)) \right) = \sum_{y \in \mathcal{Y}(\Omega)} P((\mathcal{X} = x) \cap (\mathcal{Y} = y)).$$

Le raisonnement est identique pour $P(\mathcal{Y} = y)$.

On représente souvent la loi de $(\mathcal{X}, \mathcal{Y})$ sous la forme d'un tableau à deux entrées donnant $P((\mathcal{X} = x) \cap (\mathcal{Y} = y))$ en fonction des valeurs possibles de $x \in \mathcal{X}(\Omega)$ et $y \in \mathcal{Y}(\Omega)$.

Le tableau ci-dessous illustre le calcul des lois marginales. On note :

- $\mathcal{X} : \Omega \mapsto E$ une v.a. de support $\mathcal{X}(\Omega) = \{x_1, \dots, x_m\}$ où $m \in \mathbb{N}^*$ et, $\forall k \in \llbracket 1; m \rrbracket$, $x_k \in E$.
- $\mathcal{Y} : \Omega \mapsto F$ une v.a. de support $\mathcal{Y}(\Omega) = \{y_1, \dots, y_n\}$ où $n \in \mathbb{N}^*$ et, $\forall k \in \llbracket 1; n \rrbracket$, $y_k \in F$.

| $X \setminus Y$ | y_1 | y_2 | ... | y_n | Total |
|-----------------|-------------------------------------|-------------------------------------|----------|-------------------------------------|-------------------------------------|
| x_1 | $p_{1,1}$ | $p_{1,2}$ | ... | $p_{1,n}$ | $P(X = x_1) = \sum_{j=1}^n p_{1,j}$ |
| x_2 | $p_{2,1}$ | $p_{2,2}$ | ... | $p_{2,n}$ | $P(X = x_2) = \sum_{j=1}^n p_{2,j}$ |
| \vdots | \vdots | \ddots | \vdots | \vdots | \vdots |
| x_m | $p_{m,1}$ | $p_{m,2}$ | ... | $p_{m,n}$ | $P(X = x_m) = \sum_{j=1}^n p_{m,j}$ |
| Total | $P(Y = y_1) = \sum_{i=1}^m p_{i,1}$ | $P(Y = y_2) = \sum_{i=1}^m p_{i,2}$ | ... | $P(Y = y_n) = \sum_{i=1}^m p_{i,n}$ | 1 |

V.2 Loi conditionnelle

Définition 13 : Soient E et F deux ensembles quelconques.

Soient $X : \Omega \mapsto E$ et $Y : \Omega \mapsto F$ deux variables aléatoires. Soit $x \in X(\Omega)$.

On appelle *loi conditionnelle sachant que* ($X = x$) la loi définie par :

$$P_{(X=x)}(Y = y) = \frac{P((X = x) \cap (Y = y))}{P(X = x)} \quad \text{pour } y \in Y(\Omega)$$

De même, pour $y \in Y(\Omega)$, on définit la *loi conditionnelle sachant que* ($Y = y$) par :

$$P_{(Y=y)}(X = x) = \frac{P((X = x) \cap (Y = y))}{P(Y = y)} \quad \text{pour } x \in X(\Omega)$$

Exercice 6 : On considère une urne contenant 4 boules blanches et 3 boules noires. On tire successivement 2 boules. On note X la variable aléatoire définie par $X = 0$ si la première boule est noire et $X = 1$ si elle est blanche. On définit Y de la même façon pour la deuxième boule.

- 1 On effectue les tirages avec remise. Déterminer la loi de probabilité du couple (X, Y) ainsi que les lois marginales de X et Y . Déterminer la loi de X conditionnée par Y ainsi que la loi de Y conditionnée par X .
- 2 Mêmes questions pour un tirage sans remise.

Correction :

- 1 On effectue les tirages **avec remise**.

La loi conjointe du couple (X, Y) est

| | | | |
|-----------|--|--|---------------|
| | $(Y = 0)$ | $(Y = 1)$ | Total |
| $(X = 0)$ | $\frac{3}{7} \times \frac{3}{7} = \frac{9}{49}$ | $\frac{3}{7} \times \frac{4}{7} = \frac{12}{49}$ | $\frac{3}{7}$ |
| $(X = 1)$ | $\frac{4}{7} \times \frac{3}{7} = \frac{12}{49}$ | $\frac{4}{7} \times \frac{4}{7} = \frac{16}{49}$ | $\frac{4}{7}$ |
| Total | $\frac{3}{7}$ | $\frac{4}{7}$ | 1 |

Les lois marginales sont donc $P(X = 0) = \frac{3}{7}$ et $P(X = 1) = \frac{4}{7}$.

| | | |
|------------|---------------|---------------|
| x | 0 | 1 |
| $P(X = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

 et

| | | |
|------------|---------------|---------------|
| x | 0 | 1 |
| $P(Y = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

On a, par exemple,

$$P_{(X=0)}(Y = 0) = \frac{P((X = 0) \text{ et } (Y = 0))}{P(X = 0)} = \frac{\frac{9}{49}}{\frac{3}{7}} = \frac{3}{7}$$

La loi de probabilité de Y

- conditionnée par $(X = 0)$ est :

| | | |
|--------------------|---------------|---------------|
| x | 0 | 1 |
| $P_{(X=0)}(Y = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

- conditionnée par $(X = 1)$ est :

| | | |
|--------------------|---------------|---------------|
| x | 0 | 1 |
| $P_{(X=1)}(Y = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

De même, la loi de probabilité de X

- conditionnée par $(Y = 0)$ est :

| | | |
|--------------------|---------------|---------------|
| x | 0 | 1 |
| $P_{(Y=0)}(X = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

- conditionnée par $(Y = 1)$ est :

| | | |
|--------------------|---------------|---------------|
| x | 0 | 1 |
| $P_{(Y=1)}(X = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

NB : Les variables aléatoires X et Y sont indépendantes.

2 On effectue maintenant les tirages **sans remise**.

La loi de probabilité du couple (X, Y) est

| | | | |
|-----------|--|--|---------------|
| | $(Y = 0)$ | $(Y = 1)$ | Total |
| $(X = 0)$ | $\frac{3}{7} \times \frac{2}{6} = \frac{1}{7}$ | $\frac{3}{7} \times \frac{4}{6} = \frac{2}{7}$ | $\frac{3}{7}$ |
| $(X = 1)$ | $\frac{4}{7} \times \frac{3}{6} = \frac{2}{7}$ | $\frac{4}{7} \times \frac{3}{6} = \frac{2}{7}$ | $\frac{4}{7}$ |
| Total | $\frac{3}{7}$ | $\frac{4}{7}$ | 1 |

Les lois marginales sont identiques :

| x | 0 | 1 |
|------------|---------------|---------------|
| $P(X = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

et

| x | 0 | 1 |
|------------|---------------|---------------|
| $P(Y = x)$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

La loi de probabilité de Y

- conditionnée par $(X = 0)$ est :

| x | 0 | 1 |
|--------------------|---------------|---------------|
| $P_{(X=0)}(Y = x)$ | $\frac{1}{3}$ | $\frac{2}{3}$ |

- conditionnée par $(X = 1)$ est :

| x | 0 | 1 |
|--------------------|---------------|---------------|
| $P_{(X=1)}(Y = x)$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

De même, la loi de X

- conditionnée par $(Y = 0)$ est :

| x | 0 | 1 |
|--------------------|---------------|---------------|
| $P_{(Y=0)}(X = x)$ | $\frac{1}{3}$ | $\frac{2}{3}$ |

- conditionnée par $(Y = 1)$ est :

| x | 0 | 1 |
|--------------------|---------------|---------------|
| $P_{(Y=1)}(X = x)$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

NB : Les variables aléatoires X et Y ne sont pas indépendantes.

VI VARIABLES ALÉATOIRES INDÉPENDANTES

Définition 14 : Soit (Ω, P) un espace probabilisé et X, Y deux variables aléatoires sur Ω .

On dit que les deux *variables aléatoires* X et Y sont *indépendantes* si et seulement si :

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

Exemple 13 :

| $X \setminus Y$ | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|---------------|
| 1 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| 2 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| 3 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| 4 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| 5 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| 6 | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{6}$ |
| Total | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | 1 |

X et Y sont indépendantes

| $X \setminus Y$ | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|-----------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| 1 | $\frac{1}{6}$ | 0 | 0 | 0 | 0 | 0 | $\frac{1}{6}$ |
| 2 | 0 | $\frac{1}{6}$ | 0 | 0 | 0 | 0 | $\frac{1}{6}$ |
| 3 | 0 | 0 | $\frac{1}{6}$ | 0 | 0 | 0 | $\frac{1}{6}$ |
| 4 | 0 | 0 | 0 | $\frac{1}{6}$ | 0 | 0 | $\frac{1}{6}$ |
| 5 | 0 | 0 | 0 | 0 | $\frac{1}{6}$ | 0 | $\frac{1}{6}$ |
| 6 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{6}$ | $\frac{1}{6}$ |
| Total | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | 1 |

X et Y ne sont pas indépendantes

Exercice 1 : Soient deux variables aléatoires X et Y dont les tableaux des lois conjointes sont donnés ci-dessous. Lequel correspond à des variables indépendantes ?

| $X \setminus Y$ | 1 | 2 | Total |
|-----------------|--------------------------|--------------------------|--------------------------|
| 1 | $\frac{1}{3}$ | $\frac{1}{3}$ | $P(X = 1) = \frac{2}{3}$ |
| 2 | $\frac{1}{3}$ | 0 | $P(X = 2) = \frac{1}{3}$ |
| Total | $P(Y = 1) = \frac{2}{3}$ | $P(Y = 2) = \frac{1}{3}$ | 1 |

| $X \setminus Y$ | 1 | 2 | Total |
|-----------------|--------------------------|--------------------------|--------------------------|
| 1 | $\frac{1}{4}$ | $\frac{1}{4}$ | $P(X = 1) = \frac{1}{2}$ |
| 2 | $\frac{1}{4}$ | $\frac{1}{4}$ | $P(X = 2) = \frac{1}{2}$ |
| Total | $P(Y = 1) = \frac{1}{2}$ | $P(Y = 2) = \frac{1}{2}$ | 1 |

VI.1 Caractérisation

Proposition 13 : Soient (Ω, P) un espace probabilisé et X, Y deux variables aléatoires sur Ω . X et Y sont indépendantes si, et seulement si

$$\forall A \subset X(\Omega), \forall B \subset Y(\Omega), \quad P((X \in A) \cap (Y \in B)) = P(X \in A) \times P(Y \in B)$$

Preuve :

(\Leftarrow) : Immédiat en posant si nécessaire $A = \{x\}$ et $B = \{y\}$.

(\Rightarrow) : Supposons que les variables aléatoires X, Y soient indépendantes.

Considérons deux ensembles $A = \{x_1, x_2, \dots, x_r\}$ et $B = \{y_1, y_2, \dots, y_s\}$.

$$P((X \in A) \cap (Y \in B)) = P\left(\bigcup_{i=1}^r \bigcup_{j=1}^s (X = x_i) \cap (Y = y_j)\right)$$

Or, les événements $(\mathbb{X} = x_i) \cap (\mathbb{Y} = y_j)_{\substack{i \in \llbracket 1; r \rrbracket \\ j \in \llbracket 1; s \rrbracket}}$ sont deux à deux incompatibles, donc :

$$\begin{aligned} &= \sum_{i=1}^r \sum_{j=1}^s \mathbb{P}((\mathbb{X} = x_i) \cap (\mathbb{Y} = y_j)) \\ &= \sum_{i=1}^r \sum_{j=1}^s \mathbb{P}(\mathbb{X} = x_i) \times \mathbb{P}(\mathbb{Y} = y_j), \text{ par indépendance de } \mathbb{X} \text{ et } \mathbb{Y}. \\ &= \left(\sum_{i=1}^r \mathbb{P}(\mathbb{X} = x_i) \right) \left(\sum_{j=1}^s \mathbb{P}(\mathbb{Y} = y_j) \right) \\ &= \mathbb{P}(\mathbb{X} \in A) \times \mathbb{P}(\mathbb{Y} \in B) \end{aligned}$$

Corollaire 131 : Si \mathbb{X}, \mathbb{Y} sont deux variables aléatoires indépendantes, et f, g sont deux applications définies respectivement sur $\mathbb{X}(\Omega)$ et $\mathbb{Y}(\Omega)$, alors les variables aléatoires $f(\mathbb{X})$ et $g(\mathbb{Y})$ sont indépendantes.

Preuve : Soient $f : \mathbb{X}(\Omega) \mapsto E$, $g : \mathbb{Y}(\Omega) \mapsto F$, $x \in E$ et $y \in F$.

$$\begin{aligned} \text{On a : } \mathbb{P}(f(\mathbb{X}) = x \cap g(\mathbb{Y}) = y) &= \mathbb{P}(\mathbb{X} \in f^{-1}(\{x\}) \cap \mathbb{Y} \in g^{-1}(\{y\})) \\ &= \mathbb{P}(\mathbb{X} \in f^{-1}(\{x\})) \times \mathbb{P}(\mathbb{Y} \in g^{-1}(\{y\})) \\ &= \mathbb{P}(f(\mathbb{X}) = x) \times \mathbb{P}(g(\mathbb{Y}) = y). \end{aligned}$$

Ainsi, $f(\mathbb{X})$ et $g(\mathbb{Y})$ sont indépendantes.

VI.2 Variables aléatoires mutuellement indépendantes

Définition 15 : Les variables aléatoires $\mathbb{X}_1, \mathbb{X}_2, \dots, \mathbb{X}_n$ sont dites *mutuellement indépendantes* si, pour toute famille $(x_1, x_2, \dots, x_n) \in \mathbb{X}_1(\Omega) \times \mathbb{X}_2(\Omega) \times \dots \times \mathbb{X}_n(\Omega)$ les événements $(\mathbb{X}_1 = x_1), (\mathbb{X}_2 = x_2), \dots, (\mathbb{X}_n = x_n)$ sont mutuellement indépendants, *i.e.* si, et seulement si

$$\forall (x_1, x_2, \dots, x_n) \in \prod_{i=1}^n \mathbb{X}_i(\Omega), \quad \mathbb{P}\left(\bigcap_{i=1}^n (\mathbb{X}_i = x_i)\right) = \prod_{i=1}^n \mathbb{P}(\mathbb{X}_i = x_i).$$

Exemples 14 :

- Dans le cas d'un tirage avec remise dans une urne, si \mathbb{X}_i est le numéro de la $i^{\text{ème}}$ boule tirée, $\mathbb{X}_1, \dots, \mathbb{X}_n$ sont mutuellement indépendantes.
- De manière plus générale, si on effectue n fois la même expérience, de manière indépendante, et si \mathbb{X}_i est le résultat de la $i^{\text{ème}}$, $\mathbb{X}_1, \dots, \mathbb{X}_n$ sont mutuellement indépendantes.

Proposition 14 : Soient $\mathbb{X}_1, \dots, \mathbb{X}_n$ des variables mutuellement indépendantes.

- 1 $\forall (A_1, A_2, \dots, A_n) \in \prod_{i=1}^n \mathcal{P}(\mathcal{X}_i(\Omega))$, les événements $(\mathcal{X}_1 \in A_1), (\mathcal{X}_2 \in A_2), \dots, (\mathcal{X}_n \in A_n)$ sont mutuellement indépendants *i.e.*

$$\mathbb{P} \left(\bigcap_{i=1}^n (\mathcal{X}_i \in A_i) \right) = \prod_{i=1}^n \mathbb{P}(\mathcal{X}_i \in A_i).$$

- 2 Toute sous-famille de $\mathcal{X}_1, \dots, \mathcal{X}_n$ est indépendante.

Preuve :

1

$$\begin{aligned} \mathbb{P} \left(\bigcap_{i=1}^n (\mathcal{X}_i \in A_i) \right) &= \mathbb{P} \left(\bigcap_{i \in \llbracket 1; n \rrbracket} \left(\bigcup_{x_i \in A_i} \{\mathcal{X}_i = x_i\} \right) \right) \\ &= \mathbb{P} \left(\bigcup_{(x_1, \dots, x_n) \in A_1 \times \dots \times A_n} \left(\bigcap_{i \in \llbracket 1; n \rrbracket} \{\mathcal{X}_i = x_i\} \right) \right) \\ &= \sum_{(x_1, \dots, x_n) \in A_1 \times \dots \times A_n} \mathbb{P} \left(\bigcap_{i \in \llbracket 1; n \rrbracket} \{\mathcal{X}_i = x_i\} \right) \\ &= \sum_{(x_1, \dots, x_n) \in A_1 \times \dots \times A_n} \prod_{i \in \llbracket 1; n \rrbracket} \mathbb{P}(\{\mathcal{X}_i = x_i\}) \\ &= \sum_{x_1 \in A_1} \dots \sum_{x_n \in A_n} \prod_{i \in \llbracket 1; n \rrbracket} \mathbb{P}(\{\mathcal{X}_i = x_i\}) \\ &= \left(\sum_{x_1 \in A_1} \mathbb{P}(\{\mathcal{X}_1 = x_1\}) \right) \dots \left(\sum_{x_n \in A_n} \mathbb{P}(\{\mathcal{X}_n = x_n\}) \right) \\ &= \prod_{i=1}^n \mathbb{P}(\mathcal{X}_i \in A_i). \end{aligned}$$

- 2 Quitte à réordonner $(\mathcal{X}_1, \dots, \mathcal{X}_n)$, on se donne une sous-famille $(\mathcal{X}_1, \dots, \mathcal{X}_p)$ avec $p \leq n$.

Pour $(x_1, \dots, x_p) \in \mathcal{X}_1(\Omega) \times \dots \times \mathcal{X}_p(\Omega)$, on pose $A_i = \{x_i\}$ si $i \in \llbracket 1; p \rrbracket$ et $A_i = \Omega$ sinon.

D'après l'assertion précédente, on a alors :

$$\begin{aligned} \mathbb{P} \left((\mathcal{X}_1 = x_1) \cap \dots \cap (\mathcal{X}_p = x_p) \right) &= \mathbb{P} \left((\mathcal{X}_1 \in A_1) \cap \dots \cap (\mathcal{X}_n \in A_n) \right) \\ &= \mathbb{P}(\mathcal{X}_1 \in A_1) \dots \mathbb{P}(\mathcal{X}_n \in A_n) \\ &= \mathbb{P}(\mathcal{X}_1 = x_1) \dots \mathbb{P}(\mathcal{X}_p = x_p) \times 1 \dots \end{aligned}$$

Donc $\mathcal{X}_1, \dots, \mathcal{X}_p$ sont indépendantes.

En considérant toutes les sous-familles à deux éléments d'une famille de variables aléatoires mutuellement indépendantes, on en déduit aisément le résultat suivant :

Corollaire 14.1 : Toute famille de variables aléatoires mutuellement indépendantes sont deux à deux indépendantes.

ATTENTION | Comme dans le cas des événements, la réciproque est fausse.

Proposition 15 (Lemme des coalitions) : Soit $n \in \mathbb{N}$ tel que $n \geq 2$ et $p \in \llbracket 1; n-1 \rrbracket$.
 Soient X_1, \dots, X_n une famille de n variables aléatoires mutuellement indépendantes sur Ω .
 Pour toutes application $f : \mathbb{R}^p \mapsto \mathbb{R}$ et $g : \mathbb{R}^{n-p} \mapsto \mathbb{R}$, $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.

Preuve : Admis.

VI.3 Application à la loi binomiale

Théorème 16 : Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes d'un espace probabilisé (Ω, \mathbb{P}) suivant la même loi de Bernoulli de paramètre $p \in [0; 1]$.
 Alors $X_1 + \dots + X_n$ suit la loi binomiale de paramètres n et p .

Comme X_i comptabilise le succès ou l'échec à la $i^{\text{ème}}$ épreuve de Bernoulli, la variable aléatoire $X_1 + \dots + X_n$ représente le nombre de succès de n expériences indépendantes ayant probabilité p de réussir. Une telle variable suit une loi binomiale de paramètre n et p .

On vérifie ici cela par le calcul.

Preuve : Posons $Y = X_1 + \dots + X_n$.

Pour tout $i \in \llbracket 1; n \rrbracket$, $X_i(\Omega) = \{0, 1\}$ et $Y(\Omega) = \llbracket 0; n \rrbracket$.

Soit $k \in \llbracket 0; n \rrbracket$, notons

$$A_k = \{(x_1, \dots, x_n) \in \{0, 1\}^n / x_1 + \dots + x_n = k\}.$$

A_k est de cardinal $\binom{n}{k}$ puisqu'il faut choisir k des x_i parmi les n qui ont la valeur 1 (les autres prenant la valeur 0) pour avoir un élément de A_k .

Il est alors facile de constater que $((X_1 = x_1) \cap \dots \cap (X_n = x_n))$ avec $(x_1, \dots, x_n) \in A_k$ sont deux à deux incompatibles.

Ainsi,

$$\begin{aligned} P(Y = k) &= P\left(\bigcup_{(x_1, \dots, x_n) \in A_k} \left(\bigcap_{i \in \llbracket 1; n \rrbracket} \{X_i = x_i\}\right)\right) \\ &= \sum_{(x_1, \dots, x_n) \in A_k} P\left(\bigcap_{i \in \llbracket 1; n \rrbracket} \{X_i = x_i\}\right) \\ &= \sum_{(x_1, \dots, x_n) \in A_k} \prod_{i \in \llbracket 1; n \rrbracket} P(\{X_i = x_i\}). \end{aligned}$$

Parmi les x_i , k valent 1 donc k des $\mathbb{P}(X_i = x_i)$ valent p , et $n - k$ valent 0 donc les $n - k$ autres $\mathbb{P}(X_i = x_i)$ valent $1 - p$.

$$\begin{aligned}
 &= \sum_{(x_1, \dots, x_n) \in A_k} p^k (1 - p)^{n-k} \\
 &= \binom{n}{k} p^k (1 - p)^{n-k}.
 \end{aligned}$$

Ainsi, $Y \rightsquigarrow \mathcal{B}(n, p)$.

VI.4 Espérance d'un produit de variables aléatoires indépendantes

Théorème 17 : Si X et Y sont deux variables aléatoires réelles indépendantes, alors

$$E(XY) = E(X)E(Y).$$

Preuve : Posons $X(\Omega) = \{x_1, \dots, x_p\}$, les x_i étant tous distincts et $Y(\Omega) = \{y_1, \dots, y_q\}$, les y_j étant tous distincts.

On pose $Z = XY$.

On a $Z(\Omega) = \{x_i y_j, (i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket\} = \{z_1, \dots, z_r\}$, les z_k étant tous distincts.

$$\begin{aligned}
 E(Z) &= \sum_{k=1}^r z_k \mathbb{P}(Z = z_k) \\
 &= \sum_{k=1}^r z_k \mathbb{P} \left(\bigcup_{x_i y_j = z_k} ((X = x_i) \cap (Y = y_j)) \right) \\
 &= \sum_{k=1}^r z_k \left[\sum_{x_i y_j = z_k} \mathbb{P}((X = x_i) \cap (Y = y_j)) \right] \text{ car les événements sont incompatibles} \\
 &= \sum_{k=1}^r z_k \left[\sum_{x_i y_j = z_k} \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j) \right] \text{ car les v.a. sont indépendantes} \\
 &= \sum_{k=1}^r \sum_{x_i y_j = z_k} x_i y_j \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j) \\
 &= \sum_{i=1}^p \sum_{j=1}^q x_i y_j \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j) \\
 &= \left(\sum_{i=1}^p x_i \mathbb{P}(X = x_i) \right) \left(\sum_{j=1}^q y_j \mathbb{P}(Y = y_j) \right) = E(X)E(Y).
 \end{aligned}$$

Remarque : On peut aussi voir ce résultat comme une application du **théorème (4)** de transfert appliqué à $f : (x; y) \mapsto xy$ et à la variable $Z = (X; Y)$.

La réciproque est fautive en général comme le montre l'exemple suivant :

Contre-Exemple 15 : Soient U et V deux urnes, que l'on remplit, de manière aléatoire avec deux boules.

On note X la variable aléatoire qui compte le nombre de boules dans U , Y la variable aléatoire qui compte le nombre d'urnes vides.

ATTENTION

| | $(X = 0)$ | $(X = 1)$ | $(X = 2)$ | Total |
|-----------|---------------|---------------|---------------|---------------|
| $(Y = 0)$ | 0 | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |
| $(Y = 1)$ | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ | $\frac{1}{2}$ |
| Total | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | 1 |

| z | $P(XY = z)$ |
|-----|---------------|
| 0 | $\frac{1}{2}$ |
| 1 | $\frac{1}{2}$ |
| 2 | 0 |

Alors $P(X = 0) = P(X = 2) = \frac{1}{4}$, $P(X = 1) = \frac{1}{2}$ et $P(Y = 0) = P(Y = 1) = \frac{1}{2}$.

On calcule $E(X) = 1$, $E(Y) = \frac{1}{2}$ d'où $E(XY) = \frac{1}{2} = E(X)E(Y)$ mais $P((X = 0) \cap (Y = 0)) = 0 \neq \frac{1}{8} = P((X = 0))P((Y = 0))$. Les variables X et Y ne sont pas indépendantes.

Corollaire 171 : Si X_1, \dots, X_n sont des variables mutuellement indépendantes, alors

$$E(X_1, \dots, X_n) = E(X_1) \dots E(X_n).$$

Preuve : On montre par récurrence sur $n \in \mathbb{N}^*$ la propriété $\mathcal{P}(n)$ correspondant à l'énoncé.

$\mathcal{P}(1)$ est trivialement vraie.

Soit $n \geq 2$ tel que $\mathcal{P}(n)$ et soient X_1, \dots, X_n, X_{n+1} des variables aléatoires indépendantes.

Alors X_{n+1} est indépendante de X_1, \dots, X_n donc par hypothèse de récurrence :

$$E(X_1 \dots X_n X_{n+1}) = E(X_1 \dots X_n)E(X_{n+1}) = E(X_1) \dots E(X_n)E(X_{n+1}).$$

Et, on a $\mathcal{P}(n + 1)$.

En conclusion, $\forall n \in \mathbb{N}^*$, $\mathcal{P}(n)$.

VII COVARIANCE DE DEUX VARIABLES ALÉATOIRES

Définition 16 : Soient X et Y des variables aléatoires réelles.

On appelle *covariance* de X et Y , notée $Cov(X, Y)$, le réel

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)E(Y) \\ &= E\left((X - E(X))(Y - E(Y))\right). \end{aligned}$$

On dit que X et Y sont *décorrélées* si $Cov(X, Y) = 0$.

La covariance est donc la moyenne des produits des écarts des valeurs à la moyenne de chaque série. Elle évalue la dépendance linéaire entre X et Y . D'après le **théorème (17)**, si X et Y sont indépendantes alors $Cov(X, Y) = 0$.

ATTENTION | La réciproque est fausse.

Proposition 18 : Soient X et Y des variables aléatoires réelles.

Alors,

$$V(X + Y) = V(X) + V(Y) + 2Cov(X, Y).$$

Si, X et Y sont décorrélées alors $V(X + Y) = V(X) + V(Y)$.

Preuve : Il suffit de développer et d'utiliser la **proposition (5)** :

$$\begin{aligned} V(X + Y) &= E\left((X + Y - E(X + Y))^2\right) \\ &= E\left((X + Y)^2 - 2(X + Y)E(X + Y) + E(X + Y)^2\right) \\ &= E\left((X + Y)^2\right) - E(X + Y)^2 \\ &= E(X^2) + E(Y^2) + 2E(XY) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 \\ &= V(X) + V(Y) + 2\left(E(XY) - E(X)E(Y)\right) \\ &= V(X) + V(Y) + 2Cov(X, Y). \end{aligned}$$

Corollaire 18.1 : Si X et Y sont deux variables aléatoires réelles indépendantes alors

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

Preuve : Le **théorème (17)** donne le résultat.

Le **corollaire (18.1)** se généralise aisément au cas de n variables aléatoires mutuellement indépendantes :

Proposition 19 : Si X_1, \dots, X_n sont des variables aléatoires réelles mutuellement indépendantes alors

$$V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n).$$

« La variance d'une somme de variables aléatoires indépendantes est la somme de leur variance. »

Preuve : Par linéarité de l'espérance, $E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n)$

Ainsi,

$$\begin{aligned}
 V(X_1 + \dots + X_n) &= E\left((X_1 + \dots + X_n)^2\right) - \left(E(X_1) + \dots + E(X_n)\right)^2 \\
 &= E\left(\sum_{i=1}^n X_i^2 + 2 \sum_{i<j} X_i X_j\right) - \sum_{i=1}^n E(X_i)^2 - 2 \sum_{i<j} E(X_i)E(X_j) \\
 &= \sum_{i=1}^n \underbrace{\left(E(X_i^2) - E(X_i)^2\right)}_{=V(X_i)} + 2 \sum_{i<j} \underbrace{\left(E(X_i X_j) - E(X_i)E(X_j)\right)}_{=0 \text{ par indépendance}} \\
 &= \sum_{i=1}^n V(X_i).
 \end{aligned}$$

Corollaire 191 : Soit X une variable aléatoire suivant une loi binomiale de paramètres n et p . Alors,

$$V(X) = np(1 - p).$$

Preuve : On redécompose X sous la forme d'une somme de n variables aléatoires indépendantes X_i indépendantes. On a alors :

$$V(X) = V(X_1) + V(X_2) + \dots + V(X_n) = p(1 - p) + p(1 - p) + \dots + p(1 - p) = np(1 - p).$$

Exercice 8 : Soit U et V deux variables aléatoires indépendantes de même loi, à valeurs dans $\{1, \dots, N\}$. On pose $X = U - V$ et $Y = U + V$.

Déterminer la covariance entre X et Y .

Correction : $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(XY)$ car $E(X) = E(U) - E(V) = 0$ (U, V ont même loi). Puis, $E(XY) = E((U - V)(U + V)) = E(U^2) - E(V^2) = 0$ (U, V ont même loi).

D'où $\text{Cov}(X, Y) = 0$.

Un peu d'histoire : Corrélation vs Indépendance

L'indépendance et la corrélation ou plus exactement la dépendance et la corrélation sont deux notions en statistique qui sont souvent confondues. Formellement, l'indépendance entre deux variables X et Y indique que la connaissance de la réalisation de l'une des variables n'a aucune incidence sur la probabilité de réalisation de l'autre variable. La corrélation quant à elle renvoie à l'existence d'une liaison linéaire entre les deux variables.

Donc l'indépendance est une notion liée à la « chance » de réalisation des événements des variables X et Y tandis que la corrélation est une notion liée à la fonction linéaire qui existerait entre les valeurs des deux variables.

Il est théoriquement démontré que si deux variables sont indépendantes alors elles ne sont pas corrélées. Autrement dit, si les probabilités de réalisation des événements liés aux variables X et Y vérifient la condition d'indépendance alors les valeurs de ces variables n'ont pas de liaison linéaire entre elles. Par conséquent, si deux variables sont corrélées alors elles sont nécessairement dépendantes (c'est-à-dire elles ne sont pas indépendantes).

Ce résultat laisse alors la possibilité aux situations où des variables seraient dépendantes sans être corrélées.

Index

- Bernoulli
 - Schéma de, 17
- Covariance, 31, 32
- Espérance, 8
 - Propriétés, 9
- Gain, 1
- Image
 - directe, 2
 - récioproque, 2
- Indicateur
 - de dispersion, 12, 14
 - de position, 8
- Inégalité
 - de Bienaymé-Tchebychev, 13
 - de Markov, 13
- Loi
 - binomiale, 17
 - conditionnelle, 23
 - conjointe, 22
 - d'une probabilité, 5, 14
 - de Bernoulli, 15
 - de probabilité, 5
 - marginale, 22
 - uniforme, 14
- Somme
 - de variables aléatoires, 4
- Système
 - complet d'événements, 3
 - associé à une v.a.r , 3
- Variable aléatoire, 2
 - décorélées, 31
 - centrée, 13
 - indépendante, 25
 - réduite, 13
 - réelle, 2
 - Support d'une, 2
- Variance, 12
 - Propriétés, 12