CHAPITRE 7
Probabilités sur un univers fini

Plan du chapitre

Introduction

La théorie des probabilités est une mathématisation de l'incertitude et du caractère imprévisible de certains phénomènes. Pendant l'antiquité, le hasard n'a pas été l'objet d'étude à proprement parler, mais il a été utilisé pour le divertissement dans les jeux de hasard comme les jeux de dés en terre cuite dont on trouve des traces trois millénaires avant notre ère en Égypte, en Mésopotamie et en Inde.

L'apparition de la notion de risque, préalable à l'étude des probabilités, n'est apparue qu'au XII^{e} siècle, pour l'évaluation de contrats commerciaux avec le *Traité des contrats* de Pierre de Jean Olivi, et elle s'est développée au XVI^{e} siècle, avec la généralisation des contrats d'assurance maritime.

Le célèbre problème des parties est à l'origine de la théorie des probabilités. Il avait pour objet de déterminer la proportion suivant laquelle l'enjeu doit être partagé entre les joueurs en fonction des points de chacun lorsqu'ils conviennent de ne point achever une partie. Il a été étudié par les mathématiciens italiens Luca Pacioli, Niccolo Fontana et Jérôme Cardan au XVI^{e} siècle, mais leurs réflexions manquaient de rigueur et demeuraient empiriques. Le problème fut entièrement résolu par Blaise Pascal et Pierre de Fermat dans une de leur correspondance en 1654 suite à une question posée par Antoine Gombaud (dit chevalier de Méré). À la suite d'un séjour à Paris en 1655, Christian Huygens prend connaissance de cette discussion à l'Académie Parisienne et publie en 1657 le premier traité sur la théorie probabiliste : *De ratiociniis in ludo aleae*. Cependant, Huygens attribue lui-même la paternité de la théorie des probabilités à Pascal et Fermat.

On trouve les premières traces d'une loi des grands nombres dans l'œuvre posthume *Ars Conjectandi* publiée en 1713 du mathématicien suisse Jacques Bernoulli. Abraham de Moivre publie en 1718 son traité *The Doctrine of Chances* contenant des problèmes combinatoires dont la formule de Stirling, des probabilités conditionnelles ainsi qu'une première version du théorème central limite. Aux alentours de 1770, Joseph-Louis Lagrange publie un mémoire étudiant des problèmes de durée d'un jeu de hasard, dans lequel il introduit pour la première fois des lois continues. Par la suite, il donne une première table de la loi normale dans son ouvrage publié en 1781. Ces évolutions permettent à Laplace d'énoncer une nouvelle version du théorème central limite en démontrant l'actuel théorème de Moivre-Laplace dans son traité *Théorie analytique des probabilités* publié 1812.

Finalement, la théorie moderne des probabilités est introduite par le mathématicien russe Kolmogorov au début du XXe siècle. Son axiomatisation repose sur la théorie de la mesure développée par Henri Lebesgue dans sa dissertation *Intégrale, longueur, aire* publiée à l'université de Nancy en 1902.

Les probabilités jouent un rôle important en sciences. En physique, elles apparaissent par exemple dans la description du mouvement d'une particule immergée dans un fluide (mouvement brownien). Plus généralement, elles jouent un rôle central dans la branche de la physique qui étudie et décrit les phénomènes fondamentaux à à l'échelle atomique et subatomique : la mécanique quantique. De plus, elles interviennent dans de nombreuses autres disciplines, notamment en biologie, en théorie des jeux ou en sciences humaines où les outils statistiques sont omniprésents.

Dans ce chapitre, nous commencerons par effectuer des rappels de première année sur les espaces probabilisés finis et sur la notion de variables aléatoires. Dans un second temps, nous étudierons les couples de variables aléatoires. Nous définirons notamment la notion d'indépendance pour ces dernières, ainsi que leur covariance et leur coefficient de corrélation linéaire.

I - Rappels sur les espaces probabilisés finis

I.A - Vocabulaires

Définition (Expérience aléatoire) : Une expérience aléatoire est une expérience qui ne donne pas nécessairement le même résultat quand on la renouvelle dans des conditions identiques.

Exemples 1 :

a) Un lancer d'une pièce,

b) Un lancer d'un dé classique,

c) La durée de vie en année d'un composant électronique,

d) La durée de désintégration d'un isotope radioactif.

Définition (Univers) : L'univers est l'ensemble des issues possibles d'une expérience aléatoire.

Exemples 2 :
En reprenant les exemples précédents, on a

a) $\Omega = \{\text{Pile}, \text{Face}\}$,

b) $\Omega = [\![1, 6]\!]$,

c) $\Omega = \mathbb{N}$,

d) $\Omega = \mathbb{R}_+$.

Vocabulaires : La théorie moderne des probabilités utilisent le langage des ensembles pour modéliser une expérience aléatoire. Le tableau ci-dessous établit la correspondance entre le langage ensembliste et le langage probabiliste.

Notation	**Vocabulaire ensembliste**	**Vocabulaire probabiliste**
Ω	ensemble plein	évènement certain
$\varnothing$	ensemble vide	évènement impossible
ω	élément de Ω	évènement élémentaire
A	sous-ensemble de Ω	évènement
A^c	complémentaire de A dans Ω	évènement contraire de A
$A \cup B$	réunion de A et B	évènement A ou B
$A \cap B$	intersection de A et B	évènement A et B
$A \cap B = \varnothing$	A et B sont disjoints	A et B sont incompatibles

L'événement contraire d'un évènement A est noté $\bar{A}$.

I.B - Probabilités sur un univers fini

Jusqu'à la fin du chapitre, on fixe un univers Ω **fini** modélisant une expérience aléatoire.

Définition (Probabilité) : On appelle probabilité sur l'univers Ω toute application $P : \mathscr{P}(\Omega) \to [0,1]$ vérifiant

(i) $P(\Omega) = 1$,

(ii) Pour tous évènements $A_1, \dots, A_n$ deux à deux incompatibles

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k).$$

Définition (Espace probabilisé) : Un espace probabilisé (fini) est un couple (Ω, P) où Ω est un univers fini et P une probabilité sur Ω.

Proposition 1 : Soit $P : \Omega \to [0,1]$ une probabilité sur l'univers Ω. Alors

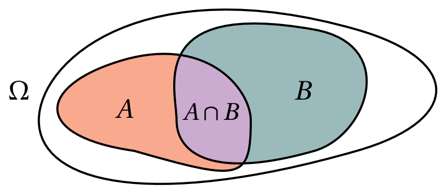
(i) $P(\varnothing) = 0$,

(ii) $\forall A \in \mathscr{P}(\Omega)$, $P(\bar{A}) = 1 - P(A)$,

(iii) $\forall (A,B) \in \mathscr{P}(\Omega)^2$, $A \subset B \Rightarrow P(A) \leqslant P(B)$ et $P(B \setminus A) = P(B) - P(A)$,

(iv) $\forall (A,B) \in \mathscr{P}(\Omega)^2$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Illustration : Le schéma ci-dessous permet de comprendre l'origine de la formule du point (iv) de la proposition.



Proposition 2 : Si $\Omega = \{\omega_1, \dots, \omega_n\}$ et si $p_1, \dots, p_n \in \mathbb{R}$ sont des réels positifs de somme 1, alors il existe une et une seule probabilité $P : \mathscr{P}(\Omega) \to [0,1]$ vérifiant

$$\forall k \in [\![1,n]\!], \quad P(\{\omega_k\}) = p_k.$$

Remarque 1 : Autrement dit, pour définir une probabilité P sur un univers fini Ω, il suffit de définir P sur chaque singleton de Ω de sorte que la somme donne 1. La formule générale pour P est alors simplement donnée par

$$\forall A \in \mathscr{P}(\Omega), \quad P(A) = \sum_{\omega \in A} P(\{\omega\}).$$

Exemple 3 : Si $\Omega = [\![1,n]\!]$, il existe une unique probabilité sur Ω vérifiant

$$\forall k \in \Omega, \quad P(\{k\}) = \frac{1}{n}.$$

Elle est appelée **probabilité uniforme** sur Ω et on a

$$\forall A \in \mathscr{P}(\Omega), \quad P(A) = \frac{\mathrm{Card}(A)}{n}.$$

I.C - Conditionnement par un évènement

Définition (Probabilité conditionnelle) : Soient A et B deux évènements telles que $P(B) > 0$. On appelle probabilité conditionnelle de A sachant B le nombre

$$P(A \mid B) = P_B(A) = \frac{P(A \cap B)}{P(B)}.$$

Exemple 4 : On jette un dé équilibré. On désigne par A l'évènement le chiffre du dé est 6 et par B l'évènement le chiffre du dé est pair. On a

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}.$$

Le résultat est intuitif : si on sait que le dé affiche une chiffre pair, on a une chance sur trois que le chiffre soit six.

Proposition 3 : Si B est un évènement de probabilité non nulle, alors l'application $P_B : \mathscr{P}(\Omega) \to [0,1]$ est une probabilité.

Remarque 2 : En particulier, on peut utiliser toutes les propriétés de la proposition 1 avec la probabilité P_B. Par exemple, on a $P_B(\bar{A}) = 1 - P_B(A)$.

Formule des probabilités composées : Si $A_1, \dots, A_m$ sont des événements tels que $P(A_1 \cap \dots \cap A_m) \neq 0$, alors

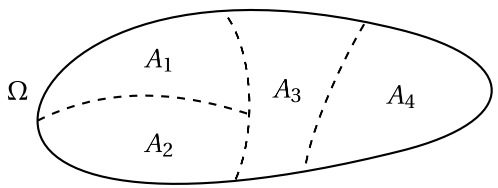
$$P\left(\bigcap_{i=1}^{m} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1) \cdots P(A_m|A_{m-1} \cap \dots \cap A_1).$$

Définition (Système complet d'évènements) : On dit que $(A_1, \dots, A_n)$ est un système complet d'événements si

(i) les évènements $A_1, \dots, A_n$ sont deux à deux incompatibles,

(ii) les événements $A_1, \dots, A_n$ recouvrent Ω, i.e. $\Omega = \bigcup_{k=1}^{n} A_k$.

Illustration : On peut représenter la notion de système complet d'évènement par le schéma ci-dessous.

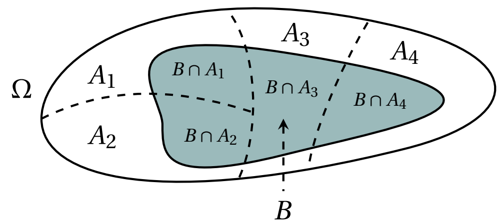


Exemple 5 : On étudie le résultat du lancer d'un dé classique. On note A l'évènement « le résultat du dé est pair » et B l'évènement « le résultat du dé est impair ». Le couple (A, B) est un système complet d'évènements.

Formule des probabilités totales : Soit B un événement. Si $(A_1, \dots, A_n)$ est un système complet d'évènements de probabilités non nulles, alors

$$P(B) = \sum_{k=1}^{n} P(B \cap A_k) = \sum_{k=1}^{n} P(B \mid A_k)P(A_k).$$

Illustration : En reprenant l'illustration précédente, le schéma ci-dessous permet de comprendre la formule des probabilités totales.



Formule de Bayes : Si A et B sont deux événements de probabilités non nulles, alors

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

I.D - Indépendance d'évènements

Définition (Évènements indépendants) : Deux évènements A et B sont dits indépendants si $P(A \cap B) = P(A)P(B)$.

Remarque 3 : Si $P(B) > 0$, l'indépendance de A et B équivaut à $P(A \mid B) = P(A)$. Autrement dit, avoir des informations sur la réalisation de B ne renseigne pas la réalisation de A.

Définition (Évènements mutuellement indépendants) : On dit que des évènements $A_1, \ldots, A_n$ sont dits mutuellement indépendants si

$$\forall I \subset [\![1, n]\!], \quad P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

Remarque 4 : Si n évènements sont mutuellement indépendants, alors ils sont indépendants deux à deux. La réciproque est fausse comme le montre l'exemple suivant.

Exemple 6 : On considère $\Omega = \{1, 2, 3, 4\}$ muni de la probabilité uniforme et on note

$$A = \{1, 2\}, \quad B = \{1, 3\}, \quad C = \{1, 4\}.$$

On vérifie facilement que ces évènements sont indépendants deux à deux, mais ils ne sont pas mutuellement indépendants car

$$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C).$$

II - Rappels sur les variables aléatoires réelles

II.A - Généralités

Définition (Variable aléatoire réelle) : Une variable aléatoire réelle est une application $X : \Omega \to \mathbb{R}$.

Jusqu'à la fin du chapitre, on fixe une variable aléatoire réelle X sur Ω.

Notation : Si $U \subset \mathbb{R}$, on note l'évènement

$$(X \in U) = X^{-1}(U) = \{\omega \in \Omega \mid X(\omega) \in U\}$$

En particulier, si $x \in \mathbb{R}$, on a

$$(X = x) = \{\omega \in \Omega \mid X(\omega) = x\} \quad \text{et} \quad (X \leqslant x) = \{\omega \in \Omega \mid X(\omega) \leqslant x\}.$$

Définition (Image d'une variable aléatoire par une fonction) : Soit $f : \mathbb{R} \to \mathbb{R}$ une application. On définit la variable aléatoire $f(X)$ comme la composée $f \circ X$, i.e.

$$\forall \omega \in \Omega, \quad f(X)(\omega) = f(X(\omega)).$$

II.B - Loi d'une variable aléatoire réelle

Proposition 4 : L'application $P_X : \mathscr{P}(X(\Omega)) \to [0, 1]$, définie par $P_X(U) = P(X \in U)$ est une probabilité sur l'ensemble $X(\Omega)$.

Remarque 5 : L'ensemble $X(\Omega)$ est l'ensemble des valeurs que prend X. Par exemple, si X donne le résultat du lancer d'un dé classique, alors $X(\Omega) = [\![1, 6]\!]$.

Définition (Loi d'une variable aléatoire) : L'application P_X est appelée la loi de la variable aléatoire X.

Proposition 5 : La loi P_X est entièrement déterminé par les nombres $P(X = x)$ pour $x \in X(\Omega)$.

Remarque 6 : Pour donner la loi d'une variable aléatoire réelle, on en déduit qu'il suffit de donner les $P(X = x)$ pour tout $x \in X(\Omega)$.

II.C - Fonction de répartition d'une variable aléatoire réelle

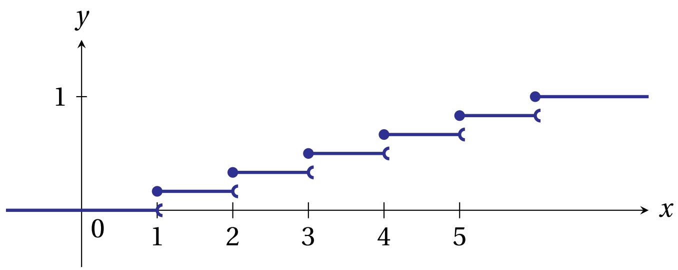
Définition (Fonction de répartition d'une variable aléatoire) : La fonction de répartition de X est l'application $F_X : \mathbb{R} \to \mathbb{R}$ définie par $F_X(x) = P(X \leqslant x)$.

Proposition 6 : On a les propriétés suivantes.

(i) La fonction F_X est croissante et continue à droite en tout point de $\mathbb{R}$.

(ii) La fonction F_X vérifie $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(x) = 1$.

Exemple 7 : Si X désigne le résultat du lancer d'un dé classique équilibré, alors le graphe de la fonction de répartition F_X est le suivant.



Proposition 7 : La loi P_X est entièrement déterminée par la fonction F_X.

Remarque 7 : Si $X(\Omega) = \{x_1, \dots, x_n\}$ avec $x_1 < x_2 < \cdots < x_n$, alors on a

$$\forall i \in [\![1, n]\!], \quad P(X = x_i) = P(X \leqslant x_i) - P(X \leqslant x_{i-1}) = F_X(x_i) - F_X(x_{i-1}).$$

II.D - Espérance d'une variable aléatoire réelle

Comme Ω est fini, il en est de même de l'ensemble $X(\Omega)$. Il n'y a donc aucune difficulté à manipuler des sommes indexées par $X(\Omega)$.

Définition (Espérance d'une variable aléatoire) : L'espérance de X est le nombre

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

Remarque 8 : L'espérance correspond à la moyenne théorique des valeurs prises par X.

Proposition (Linéarité de l'espérance) : Si Y est une variable aléatoire réelle sur Ω et si $a \in \mathbb{R}$, alors

$$\mathrm{E}(aX + Y) = a\,\mathrm{E}(X) + \mathrm{E}(Y).$$

Théorème du transfert : Si $f : \mathbb{R} \to \mathbb{R}$ est une application, alors

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

II.E - Variance et écart type d'une variable aléatoire réelle

Définition (Variance d'une variable aléatoire) : On appelle variance de X le réel

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

Définition (Écart type d'une variable aléatoire) : On appelle écart type de X le réel $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarques 9 :

a) L'écart type est aussi appelée l'écart quadratique moyen à la moyenne. Il permet d'évaluer la dispersion de X autour de sa moyenne $\mathrm{E}(X)$.

b) La variable aléatoire X est constante si et seulement si $\mathrm{V}(X) = \sigma(X) = 0$.

Théorème de Koenig-Huyghens : On a $\mathrm{V}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$.

Proposition 8 : Pour tout $(a, b) \in \mathbb{R}^2$, on a $\mathrm{V}(aX + b) = a^2\,\mathrm{V}(X)$.

Inégalité de Bienaymé-Tchebychev : Pour tout $\varepsilon > 0$, on a l'inégalité

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

II.F - Lois usuelles

II.F.1 - Loi uniforme

Définition (Loi uniforme) : Soit $n \in \mathbb{N}^*$. On dit que X suit la loi uniforme sur un ensemble $\{x_1, \dots, x_n\}$, que l'on note $\mathscr{U}(\{x_1, \dots, x_n\})$, si

$$X(\Omega) = \{x_1, \dots, x_n\} \quad \text{et} \quad \forall k \in [\![1, n]\!], \quad P(X = x_k) = \frac{1}{n}.$$

Exemple 8 : On lance un dé classique équilibré et on note X le résultat. La variable aléatoire X suit une loi uniforme sur $[\![1, 6]\!]$.

Proposition 9 : Si X suit la loi $\mathscr{U}([\![1, n]\!])$ avec $n \in \mathbb{N}^*$, alors

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

II.F.2 - Loi de Bernoulli

Définition (Loi de Bernoulli) : Soit $p \in [0, 1]$. On dit que X suit la loi de Bernoulli de paramètre p, que l'on note $\mathscr{B}(p)$, si

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

Exemple 9 : On considère une épreuve de Bernoulli : c'est une expérience aléatoire ne possédant que deux issues que l'on note 1 pour « succès » et 0 pour « échec ». La variable aléatoire X donnant le résultat de cette expérience suit la loi de Bernoulli de paramètre p qui est la probabilité d'obtenir un succès.

Proposition 10 : Si X suit la loi $\mathscr{B}(p)$ avec $p \in [0, 1]$, alors

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

II.F.3 - Loi binomiale

Définition (Loi binomiale) : Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. On dit que X suit la loi binomiale de paramètre (n, p), que l'on note $\mathscr{B}(n, p)$, si

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

Exemple 10 : On répète une épreuve de Bernoulli dont la probabilité de succès est p de manière identique et indépendante n fois. La variable aléatoire X donnant le nombre de succès total suit la loi binomiale de paramètre n et p.

Remarque 10 : Si $n = 1$, on retrouve la loi de Bernoulli de paramètre p.

Proposition 11 : Si X suit la loi $\mathscr{B}(n, p)$ avec $n \in \mathbb{N}$ et $p \in [0, 1]$, alors

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

III - Couples de variables aléatoires

Jusqu'à la fin du chapitre, on fixe deux variables aléatoires réelles X et Y sur Ω. Le couple (X, Y) est un couple de variables aléatoires.

III.A - Généralités

Définition (Loi conjointe) : La loi conjointe de (X, Y) (ou la loi du couple) est la donnée des nombres réels $P((X = x) \cap (Y = y))$ pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$.

Définition (Lois marginales) : La loi de X et la loi de Y sont appelées les lois marginales du couple (X, Y).

Remarque 11 : La loi conjointe détermine les lois marginales. En effet, par la formule des probabilités totales, on a les relations

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

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

ATTENTION : La réciproque est fausse en général : les lois marginales ne déterminent pas la loi conjointe. Par exemple, les tableaux ci-dessous donnent deux lois conjointes différentes possibles pour un couple (X, Y) avec des lois marginales identiques.

$P((X=x) \cap (Y=y))$	$y=0$	$y=1$	$P(X=x)$
$x=0$	$1/2$	0	$1/2$
$x=1$	0	$1/2$	$1/2$
$P(Y=y)$	$1/2$	$1/2$	

$P((X=x) \cap (Y=y))$	$y=0$	$y=1$	$P(X=x)$
$x=0$	$1/4$	$1/4$	$1/2$
$x=1$	$1/4$	$1/4$	$1/2$
$P(Y=y)$	$1/2$	$1/2$	

Remarque 12 : Les définitions précédentes s'étendent naturellement aux n-uplets de variables aléatoires.

Définition (Loi conditionnelle) : Soit $y \in Y(\Omega)$ tel que $P(Y=y) > 0$. La loi conditionnelle de X sachant $(Y=y)$ est la donnée des nombres $P(X=x \mid Y=y)$ pour tout $x \in X(\Omega)$.

III.B - Indépendance de deux variables aléatoires

Définition (Indépendance de deux variables aléatoires) : On dit que X et Y sont des variables aléatoires indépendantes si

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

Proposition 12 : Soient X et Y deux variables aléatoires indépendantes.

(i) Pour tout $(A,B) \in \mathscr{P}(X(\Omega)) \times \mathscr{P}(Y(\Omega))$, les évènements $(X \in A)$ et $(Y \in B)$ sont indépendants, i.e.

$$P((X \in A) \cap (Y \in B)) = P(X \in A)P(Y \in B).$$

(ii) Si $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ sont deux applications, alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

III.C - Indépendance mutuelle de variables aléatoires

On fixe un entier $n \in \mathbb{N}^*$ et n-uplet $(X_1, \ldots, X_n)$ de variables aléatoires sur Ω.

Définition (Indépendance mutuelle de variables aléatoires) : On dit que des variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes si

$$\forall (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega), \quad P\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = \prod_{i=1}^{n} P(X_i = x_i).$$

Proposition 13 : Soient $X_1, \ldots, X_n$ mutuellement indépendantes.

(i) Pour tout $(A_1, \ldots, A_n) \in \mathscr{P}(X_1(\Omega)) \times \cdots \times \mathscr{P}(X_n(\Omega))$, les évènements $(X_i \in A_i)$ pour $i \in [\![1, n]\!]$ sont mutuellement indépendants.

(ii) Si $f : \mathbb{R}^p \to \mathbb{R}$ et $g : \mathbb{R}^{n-p} \to \mathbb{R}$ sont deux applications avec $p \in [\![1, n-1]\!]$, alors les variables aléatoires $f(X_1, \ldots, X_p)$ et $g(X_{p+1}, \ldots, X_n)$ sont indépendantes.

Théorème 1 : Si $X_1, \dots, X_n$ sont des variables aléatoires mutuellement indépendantes de loi $\mathscr{B}(p)$, alors la variable aléatoire $X_1 + \cdots + X_n$ suit la loi $\mathscr{B}(n, p)$.

III.D - Covariance de deux variables aléatoires

Définition (Covariance de deux variables aléatoires) : La covariance des variables aléatoires X et Y est le nombre réel

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

Remarque 13 : On peut interpréter la covariance de la façon suivante.

a) Si $\mathrm{Cov}(X, Y) > 0$, alors Y a tendance à augmenter lorsque X augmente.

b) Si $\mathrm{Cov}(X, Y) < 0$, alors Y a tendance à diminuer lorsque X augmente.

c) Si $\mathrm{Cov}(X, Y) = 0$, alors il n'y a pas de lien entre les variations des variables aléatoires X et Y.

Théorème 2 : On a $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$.

Remarque 14 : En pratique, pour calculer l'espérance de XY, on utilise la relation

$$E(XY) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xyP((X = x) \cap (Y = y)).$$

Proposition 14 : Si $(a, b) \in \mathbb{R}^2$, alors on a

$$\mathrm{V}(aX + bY) = a^2\mathrm{V}(X) + 2ab\mathrm{Cov}(X, Y) + b^2\mathrm{V}(Y).$$

Théorème 3 : Si X et Y sont des variables aléatoires indépendantes, alors

(i) $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$,

(ii) $\mathrm{Cov}(X, Y) = 0$,

(iii) $\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y)$.

ATTENTION : La réciproque est fausse en général. Par exemple, si X et Y sont deux variables aléatoires indépendantes dont la loi est donnée par

$$P(X = 0) = P(X = 1) = P(Y = 1) = P(Y = -1) = \frac{1}{2},$$

alors les variables aléatoires X et $Z = XY$ ne sont pas indépendantes car

$$P((X = 1) \cap (Z = 1)) \neq P(X = 1)P(Z = 1).$$

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

III.E - Coefficient de corrélation linéaire

Définition (Coefficient de corrélation linéaire) : Le coefficient de corrélation linéaire de deux variables aléatoires non constantes X et Y est le nombre réel

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

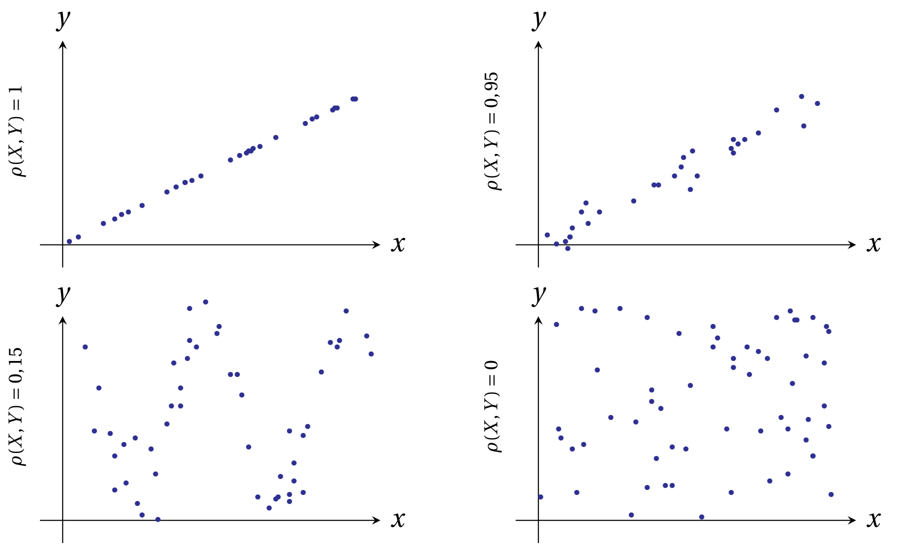
Proposition 15 : Soient X et Y deux variables aléatoires non constantes.

(i) Si X et Y sont indépendantes, alors $\rho(X, Y) = 0$.

(ii) On a $|\rho(X, Y)| \leqslant 1$.

(iii) On a $|\rho(X, Y)| = 1$ si et seulement si il existe $(a, b) \in \mathbb{R}^2$ tel que $Y = aX + b$.

Illustration : Avec un ordinateur, on simule de nombreuses fois un couple de variables aléatoires (X, Y) et on place les points obtenus dans le plan. Voici les résultats des simulations de quatre expériences différentes.



Remarque 15 : On peut interpréter le coefficient de corrélation linéaire de la façon suivante.

a) S'il vaut 1 ou -1, les variables aléatoires sont linéairement liées.

b) Plus il est proche de 1 ou de -1, plus les variables aléatoires sont linéairement corrélées.

c) Lorsqu'il est nul, il n'y a pas de corrélation linéaire entre les deux variables aléatoires.