Université Abdelmalek Essaadi
Ecole Nationale des Sciences Appliquées
Al Hoceima, Maroc

# Mathématiques Appliquées

–Cours–
Probabilité & Statistique descriptive

**Mohamed ADDAM**
Professeur de Mathématiques

École Nationale des Sciences Appliquées d'Al Hoceima
–ENSAH–
addam.mohamed@gmail.com
m.addam@uae.ac.ma



16 Mars 2020

# Table des matières

# Notations

- $\mathbb{N} := \{0, 1, 2, \ldots\}$ l'ensemble des naturels,
- $(-\mathbb{N}) := \{\ldots, -2, -1, 0\}$ l'ensemble des opposés des naturels,
- $\mathbb{N}^* = \mathbb{N}\backslash\{0\} := \{n \in \mathbb{N} \ / \ n \neq 0\}$,
- $\mathbb{Z} := \mathbb{N} \cup (-\mathbb{N})$ l'ensemble des entiers,
- $\mathbb{D}$ l'ensemble des décimaux,
- $\mathbb{Q} := \{\frac{p}{q} \ / \ p \in \mathbb{Z}, \ q \in \mathbb{N}^*\}$ l'ensemble des rationnels,
- $\mathbb{R}$ l'ensemble des nombres réels,
- $\mathbb{C}$ l'ensemble des nombres complexes.

On suppose connues les propriétés élémentaires de ces ensembles.

# Chapitre 1

# Statistique descriptive

## 1.1 Introduction et vocabulaire

Avant d'entamer le sujet de ce chapitre, on pose la question suivante : Qu'est-ce que la statistique ?
La statistique est une science qui se propose de décrire, de présenter, d'analyser et d'interpréter des phénomènes collectifs susceptibles d'être quantifiés.
Depuis l'Antiquité jusqu'au Moyen-Âge, la statistique avait surtout pour objectif de recenser la population afin de renseigner les Etats (même racine que "statistique" : Status, ... ).
Le vocabulaire employé en statistque en découle (population, individus...).
De nos jours, les états modernes possèdent des organismes chargés de mesurer l'activité économique sous ses différents aspects. Au Maroc, c'est la délégation nationale de la statistique et de planification
Depuis plusieurs décennies, la statistique a évolué : elle a pris une ampleur considérable dans les secteurs publics et privés qui possèdent leurs propres services de statistique pour mesurer, prévoir et prendre des décisions au bon moment.
La démarche statistique commence toujours par une question que l'on se pose. C'est à partir de cette question qu'est élaborée l'enquête.

## 1.2 Présentation de la statistique descriptive

Les statistiques descriptives permettent seulement, à partir des données, de dégager des pourcentages, de représenter l'intensité d'un caractère,..., sans réel questionnement.
La statistique mathématique permet d'estimer les paramètres (moyenne, variance,...), de vérifier la validité des hypothèses..., et ainsi de prévoir et de prendre des décisions.

Depuis les nouvelles techniques et la puissance des ordinateurs, cette partie des mathématiques a pris une grande place. Géographie, médecine, sciences humaines, sciences économiques, linguistique, biologie, politique,... aucun domaine qui ne soit épargné.

– **En sociologie** : la statistique permet une comparaison des pyramides des âges de divers pays et prévisions démographyques ; problèmes d'emplois, départ en retraites...
– **En médecine** : La statistique joue un rôle de plus en plus imprtant. Les première études statistiques de **Florence NIGHTINGALE**, infirmière anglaise durant la guerre de Crimée de 1854 à 1856, permirent d'identifier les causes de mortalité des soldats et conduirent à l'amélioration des conditions d'hygiène des hôpitaux militaires anglais.

L'idée de la statistique est de présenter des données par quelques valeurs sans trop déformer les informations : d'où la recherche de paramètres.
Dans le monde de l'industrie, les normes de qualité, la fiabilité d'une production, la bonne utilisation d'une machine... sont basées sur la statistique.

Les prises de décision dans de nombreuses entreprises sont faites à partir de traitement de données statistiques analysées, interprétées, généralisées... : études de marché, analyse des risques, étude de stratégies. Il en est de même dans le monde politique.

La recherche de modélisation a dégagé un modèle que l'on rencontre très fréquemment : c'est la **loi normale**. Sa représentation est une courbe en cloche ou gaussienne [1].

Cette loi s'est imposée dans de nombreux domaines : physique, physiologie, génétique, biologie, médecine...
Quel que soit l'examen médical effectué, on compare toujours votre cas à un cas **normal**.
La taille des enfants, relevée dans les carnets de santé, le poids, le taux de choléstérol, le nombre de plaquettes, de globules rouges,.... Et que dire de la normalité en météo !

## 1.3 Indicateurs de position

### 1.3.1 Vocabulaire

La statistique commence toujours par une problématique : c'est à partir de la **question** que l'on se pose que s'élabore l'enquête et que l'on récolte les données.

L'ensemble sur lequel porte l'étude est la **population**. C'est l'ensemble auquel on fera référence : il doit donc être délimité avec précision. Un élément de la population est un **individu**.

Le caractère étudié est l'aspect que l'on observe sur les individus. Ce caractère prend des valeurs différentes suivant les individus. C'est la **variabilité**.

Pour un même individu, cette variabilité peut dépendre aussi du moment où cette étude est effectuée. Il est donc nécessaire de préciser les conditions de l'enquête.

Lorsque la population est trop importante, l'étude est faite sur un ou plusieurs **échantillons** de la population.
Pour étendre les résultats collectés à l'ensemble de la population, il est nécessaire que l'échantillon soit **représentatif** et **non biaisé**. Une procédure fiable consiste à prendre un échantillon parfaitement au hasard dans la population complète.
Mais l'être humain ne choisit pas vraiment au hasard : pour obtenir un échantillon de façon aléatoire, on utilise des **simulations**. Le problème est alors de prendre un échantillon de taille suffisante.

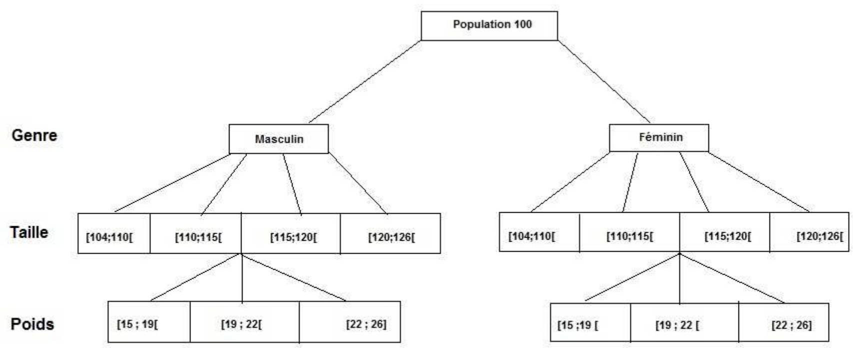
Pour chaque valeur (ou modalité) du caractère, on compte le nombre d'individus ayant cette valeur : on obtient les effectifs et en divisant par l'effectif total, on obtient la**fréquance** de cette valeur.
Suivant les échantillons, la fréquence de cette valeur n'est pas toujours identique : c'est la **fluctuation d'échantillonnage**. Rien n'est sûr ! !
Pour le calcul des probabilités, les statisticiens ont établi des **intervalles de confiance** (les fourchettes du langage courant), avec des risques d'erreur de 5 $\%$, ou de 1 $\%$...
On ne peut appliquer ces formules que si l'échantillon est tiré au hasard dans la population.
Ainsi la fréquence d'une valeur d'un caractère sera donnée dans un intervalle, avec risque d'erreur connu : **c'est l'apport fondamental de la Statistique**.

**Exemple 1.3.1** *On pose la question de savoir si, pour un enfant, il y a un lien entre sa taille et son poids, et si on peut établir un critère de normalité.*
*L'enquête va porter sur des enfants ; il faudra limiter la population à des enfants de même âge pour que l'on puisse répondre à la question que l'on se pose.*
*Il est nécessaire d'étudier la taille et le poids de chaque enfant.*
*Si la variabilité n'existait pas, tous les enfants auraient exactement le même poids et la même taille !*
*Si on observe un enfant alors qu'il vient d'être malade, on ne trouvera pas le même poids : suivant l'instant, il y a aussi variabilité pour le mêm individu.*

1. **Francis Galton(1822–1911)**, physiologiste anglais disait de la loi normale :"Je ne connais rien d'aussi impressionnant que la merveilleuse forme de l'ordre cosmique exprimée par cette loi..., elle règne sereinement et avec retenue au milieu de la plus folle confusion."

*Il n'est pas possible de faire l'étude pour tous les enfants. L'échantillon ne sera pas représentatif si l'étude est faite sur des enfants venant en consulation dans un centre pour l'enfance maltraitée, ou si elle est faite dans un quartier à population asiatique, plus petite en taille...*
*L'étude est faite sur 719 enfants de 6 à 6,5 ans, venant en consultation pédiatrique dans un cHU pour l'entrée à l'école primaire.*
*Pour prendre un échantillon non biaisé de 100 enfants, on donne un numéro d'ordre à chaque enfant et on choisit 100 nombres au hasard entre 1 et 719. C'est ce qui a été réalisé dans la problématique suivante :*
**Problématique** *: La collecte de données statistiques nécessite un tri. Quels critères choisir ? Comment répartir la population suivant les informations données ?*
*L'utilisation d'arbres, de tableaux et des commandes de tri des tableurs peuvent être très utiles.*

*Une étude porte sur des enfants de 6 ans. On peut imaginer un tri par le genre (masculin, féminin), puis par la taille, puis par le poids.*
*Le travail suivant peut se faire sans l'utilisation d'un logiciel(se partager le travail par dépouillement) ou avec les commandes de tri sous EXCEL.*



*Il arrive que, pour la lisibilité, on donne les résultats par classe. Le graphique présente la distribution des fréquences des poids pour deux échantillons de 100 enfants parmi les 719 enfants.*

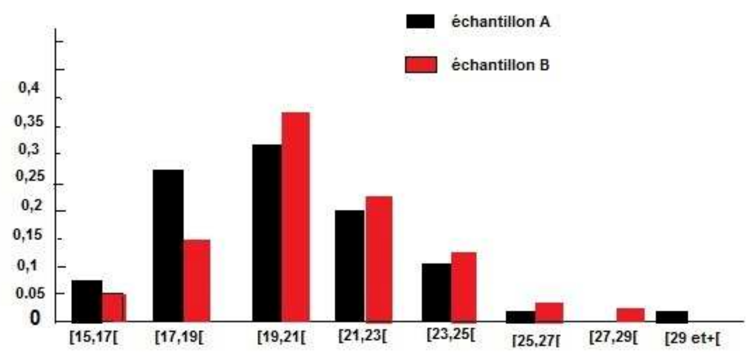



FIGURE 1.1 – Cette figure représente l'histogramme de la fluctuation des deux échantillons $A$ et $B$.

### 1.3.2 Valeurs extrêmes, étendue, et mode

***Définition 1.3.1***

1. *La valeur maximale d'une série statistique est notée $X_{max}$ et La valeur minimale d'une série statistique est notée $X_{min}$.*
2. *La distance $d = X_{max} - X_{min}$ s'appelle l'***étendue** *de la population ou de la série statistique. L'étendue est un* **indicateur de dispersion**.
3. *Le mode ou la classe modale d'une série statistique est la valeur la plus fréquente de cette série, c'est-à-dire la valeur ayant le plus grand effectif. Le mode est un* **indicateur de position**.

### 1.3.3 Moyenne, médiane

**Définition 1.3.2**

1. *La* **moyenne empirique** *exprime la valeur moyenne de l'échantillon :*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{\displaystyle\sum_{i=1}^{p} n_i X_i}{\displaystyle\sum_{i=1}^{p} n_i} = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_p x_p}{n}$$

*où $n_i$ est l'effectif de type $X_i$ et $n = \sum_{i=1}^{p} n_i$ est la somme des effectifs. La moyenne est* **indicateur de position**.

2. *La* **mediane** $X_M$ *est la valeur qui partage l'échantillon, rangé dans l'ordre croissant* $X_1, X_2, \ldots, X_n$ *(ou décroissant), en deux parties égales :*

$$\mathcal{M} = \begin{cases} X_{\frac{n+1}{2}}, & \text{si } n \text{ est impair}, \\ \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}, & \text{sinon}. \end{cases}$$

*La médiane* $X_M = M_e$ *est la valeur telle que* $50\%$ *au moins de la population prend une valeur inférieure ou égale à* $X_M = M_e$ *et* $50\%$ *au moins de la population prend une valeur supérieure ou égale à* $X_M = M_e$*. La médiane divise la population en son milieu. La médiane est un* **indicateur de position**.

Le mathématicien Français, Gauss (1777–1855), a montré que c'est autour de la moyenne que se répartissent des mesures faites au cours d'une expérience, de telle sorte que la somme des carrées des erreurs soit minimale.

**Propriété 1.3.1** *On considère une série statistique* $x_1, x_2, \ldots, x_n$*, on suppose que chaque* $x_i$ *a une fréquence* $f_i$*. Alors*

1. *La moyenne* $\overline{x}$ *de cette série s'écrit sous la forme :*

$$\overline{x} = \sum_{i=1}^{n} f_i x_i$$

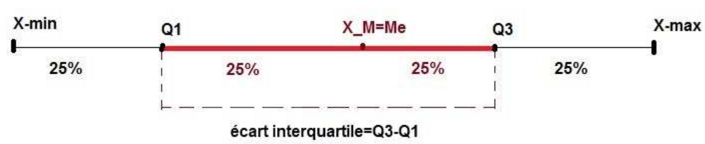
*où* $f_i = \frac{n_i}{n}$ *est l'expression de la fréquence de type* $i$ *et* $n_i$ *est l'effectif de* $x_i$*.*

2. *Si une série de valeurs* $x_i$ *a pour moyenne* $\overline{x}$*, alors la série de valeurs* $ax_i + b$ *a pour moyenne* $a\overline{x} + b$*. (Ici on parle de la linéarité de la moyenne)*
3. *Si on connaît les moyennes* $\overline{x}$ *et* $\overline{y}$ *de deux partielles d'une série, d'effectifs* $N$ *et* $P$*, alors la moyenne de la série est la moyenne des moyennes partielles* $x$ *et* $y$*, affectées des effectifs* $N$ *et* $P$ *:*

$$\overline{z} = \frac{N\overline{x} + P\overline{y}}{N + P}$$

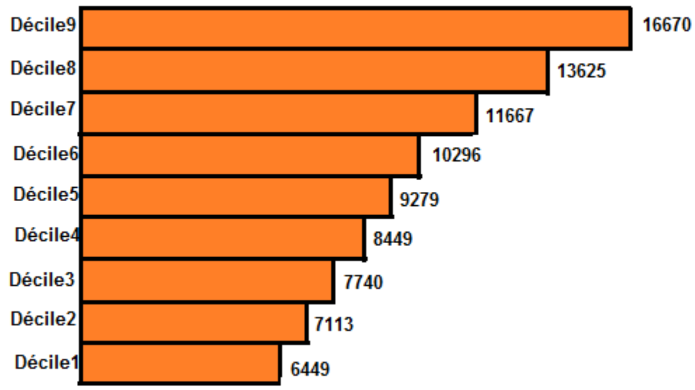
### 1.3.4 Quartiles, déciles

**Définition 1.3.3** 1. *Lorsque les individus sont rangés dans l'ordre croissant des valeurs du caractère, les* **quartiles***, notés* $Q_1$ *et* $Q_3$*, séparent la population en 4 parties de même effectif.*
$[Q_1, Q_3]$ *est l'intervalle interquartile. La distance* $Q_3 - Q_1$ *est l'écart* **interquartile**.




2. *La distance* $d = X_{max} - X_{min}$ *s'appelle l'***étendue** *de la population ou de la série statistique.*
3. *Les déciles séparent la population en 10 parties de même effectif, soit* $10\%$ *de la population. On utilise souvent le rapport* $\frac{\text{décile 9}}{\text{décile 1}}$ *qui donne une idée de la disparité.*

**Exemple 1.3.2** *Toujours en triant les individus, lorsque l'étendue des valeurs est très grande et la population très importante, on a la découpe en 10 parties de même effectif.*
*La répartition des salaires est souvent donnée en déciles, par tranches de* $10\%$ *des salariés.*
*Voici la répartition des femmes salariées selon le salaire annuel (en Dirhams Dh) en 2010 :*



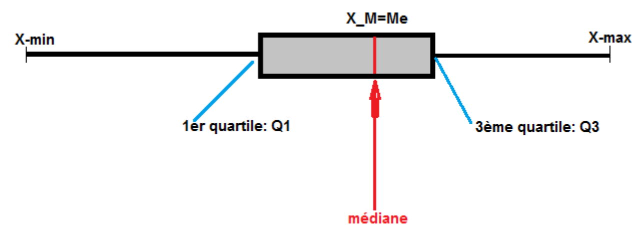

- *Le premier décile est 6449, cela signifie que* $10\%$ *des salariées ont un salaire inférieur à* $6449$ *Dh ;*
- *le neuvième décile est 16670 Dh, cela signifie que* $90\%$ *ont un salaire inférieur à 16670 Dh, donc il reste* $10\%$ *de salariées qui ont un salaire supérieur à* $16670$ *Dh.*
- *le décile 5 est la médiane :* $50\%$ *des salariés ont un salaire inférieur à 9279 Dh ; on parle du salaire médian.*

*Dans cette répartition, on ne connaît ni la valeur minimale, ni la valeur maximale.*

1. *Donner une signification au décile 7...*
2. *A-t-on toujours le même écart, en Dirhams, entre deux déciles consécutifs ?*
3. *Le salaire moyen est de 10892 Dh, Où se situe-t-il par rapport aux déciles ?...*
4. *Calculer le quotient* $\frac{\text{décile 9}}{\text{décile 1}} = \ldots.$

### 1.3.5 Diagramme en boîte ou BoxPlot

**Définition 1.3.4** *Pour représenter la répartition en quartiles d'une série statistique, on peut utiliser le* **diagramme** *en boîte : L'écrat interquartile est la longueur du rectangle (la boîte)*




**Exemple 1.3.3** *L'étude du taux de réussite au BAC Sciences Maths-section A dans les départements marocains a permis d'obtenir une répartition en 4 parties. On peut en faire une représentation sur un axe. Pour donner plus de poids aux $50\%$ des départements situés dans l'intervalle interquartile, $[Q_1, Q_3]$, on trace un rectangle, coupé au niveau de la médiane.*
*Un tel diagramme est un* **diagramme en boîte** *(ou «boîte à moustaches»).*


## 1.4 Variance et écart-type

### 1.4.1 Définition

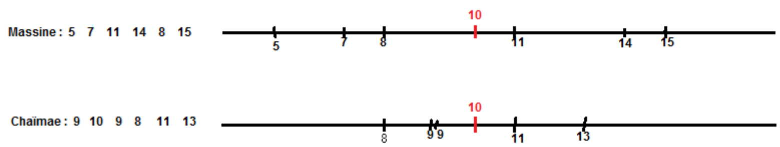
La médiane, les quartiles et les déciles ont permis de donner une "image" d'une série statistique en séparant la population par moitiés, par quarts et par dixièmes. Il est cependant souhaitable que la population soit de taille suffisante ! ! Les valeurs extrêmes n'ont alors pas d'influence sur ces paramètres.
Pour donner de l'importance aux valeurs extrêmes et donner un autre regard sur une série, on utilise la moyenne et l'écart-type, très utile pour de petites populations.

**Définition 1.4.1** 1. *La* **variance empirique** *exprime la valeur moyenne de l'échantillon :*

$$\mathrm{V} = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2.$$

2. *L'écart-type $\sigma$ est égal à la racine carrée de la variance : $\sigma = \sqrt{\mathrm{V}}$.*

**Exemple 1.4.1** *Massine et Chaïmae comportent leurs notes aux tests de Mathématiques de l'année*




*On vérifie que les deux amis ont la même moyenne :$\overline{x} = 10$.*
*Cependent leurs notes ne sont pas réparties de la même façon autour de la moyenne : Chaïmae est plus "régulière", plus proche de la moyenne à chaque test.*
*Pour chiffrer cette "régularité", on calcule l'écart de chaque note à la moyenne $\overline{x}$ et on élève cet écart au carré*

*Plus la note est éloignée de la moyenne, plus le carré de l'écart est grand : c'est le segment entre l'axe des abscisses et la parabole.*
*On calcule la moyenne de ces carrées : cette moyenne est la variance de la série, et, pour revenir dans la même unité que les valeurs, on en prend la racine carrée : c'est l'écart-type.*

- *Pour Massine : $\sigma^2 = \frac{80}{6} = 13,33$ et $\sigma = \sqrt{\frac{80}{6}} = 3,65$.*
- *Pour Chaïmae : $\sigma^2 = \frac{16}{6} = 2,66$ et $\sigma = \sqrt{\frac{16}{6}} = 1,63$.*

*Les deux distributions ne donnent pas la même impression. Pour la première, la donnée de la moyenne et de l'écart-type ne rend pas la dissymétrie. Le diagramme en boîte est plus significatif.*
*Pour la seconde, la symétrie se retrouve la courbe et le diagramme en boîte.*

### 1.4.2 Interprétation de la courbe en cloche : Plage de normalité

On étudie un caractère sur une population très nombreuse. Lorsque la distribution des individus est représentée par un diagramme de forme symétrique autour de la moyenne, avec de moins en moins de valeurs lorsque l'on éloigne de la moyenne, on parle de **courbe en cloche** ou **gaussienne**.
On considère une population formée par le poids de 500 enfants entre 5 et 7 ans, de moyenne $\mu = 20$ et d'écart-type $\sigma = 3$.
Il s'agit du cas de nombreux caractères étudiés sur les personnes : taille ou poids dans une tranche d'âge donnée, mesures pratiquées dans les analyses médicales,..., mais aussi la répartition des notes à un concours ; dans le domaine de la biologie, pour l'étude des espèces (le nombre de graines dans un fruit de coquelicot, par exemple).
Le nombre d'industrie n'est pas de reste ! Nombre de véhicules non utilisables dans une grande société de location de voitures, contrôle de qualité à la sortie d'une chaîne de production, ou vérification d'une machine à empaqueter....

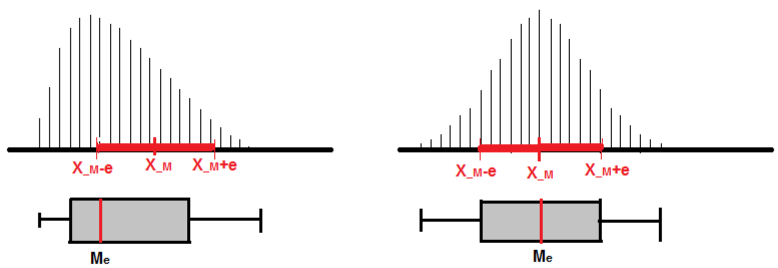
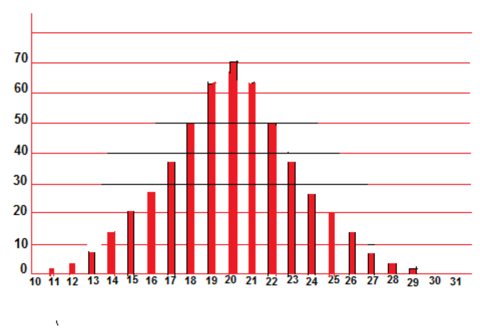



FIGURE 1.2 – A gauche la courbe et le diagramme en boîte de Massine et à droite la courbe et le diagramme en boîte de Chaïmae. Les deux courbes montrent que les notes sont plus normales pour Chaïmae que Massine.



**Propriété 1.4.1** *Les courbes en cloche, ou gaussiennes, ont des caractéristiques bien particulières. De part la symétrie de la courbe,* **médiane** *et* **moyenne** *sont confondues : $\bar{X} = M_e = \mu$, le diagramme en boîte est aussi symétrique, et les quartiles $Q_1$ et $Q_3$ sont tels que :*

$$M_e - Q_1 = Q_3 - M_e.$$

*En pourcentage de la population, on a*

*(i)* $68\%$ *sur l'intervalle* $[\mu - \sigma, \mu + \sigma]$*,*

*(ii)* $95\%$ *sur l'intervalle* $[\mu - 2\sigma, \mu + 2\sigma]$*,*

*(ii)* $99,7\%$ *sur l'intervalle* $[\mu - 3\sigma, \mu + 3\sigma]$*.*

On définit ainsi des **plages de normalité** :

1. $95\%$ sur l'intervalle $[\mu - 2\sigma, \mu + 2\sigma]$,
2. $99\%$ sur l'intervalle $[\mu - 3\sigma, \mu + 3\sigma]$.

**Définition 1.4.2** *On appelle* **intervalle** *de normalité ou plage de normalité, l'intervalle de type $[\mu - k\sigma, \mu + k\sigma]$ avec $k \in \mathbb{N}^*$ dans lequel on a une plage de normalité $\geq 95\%$.*

### 1.4.3 Intervalle de confiance : Loi des grandes nombres

La clé de la réussite des sondages est la loi des grandes nombres : si on répète un grand nombre de fois une même expérience, dans bien des cas la moyenne des résultats tend à s'équilibrer et devenir, à long terme, prévisible.

1. **Cas d'une pièce bien équilibrée** : Par exemple, si on lance deux fois une pièce, même si chaque côté a la même chance d'apparaître, il n'est pas rare que l'on obtienne deux face ou deux pile. Mais si on la lance 1000 fois, il y a de grandes chances pour que la proportion de pile et de face soit à peu près la même (en fait, il y a plus de 99 chances sur 100 pour que chacune de ces proportions soit comprise entre $47\%$ et $53\%$.
2. **Cas d'une pièce mal équilibrée** : Supposons qu'on lance 1000 fois une pièce et qu'on observe sur cette série de lancers que la fréquence d'apparition de pile est $60\%$ .
   Lors d'une autre série de 1000 lancers, la fréquence d'apparition de pile ne sera pas à coup sûr égale à $60\%$. Cependant, les statisticiens estiment qu'il y a $95\%$ de chances pour que cette fréquence soit comprise entre $57\%$ et $63\%$.
   Pour trouver cette fourchette en pourcentage, ils calculent $\frac{100}{\sqrt{n}}$ où $n$ est le nombre d'expérience. Ici, $n = 1000$ et $\frac{100}{\sqrt{1000}} = 3$, donc ils écrivent que la fourchette est
   $$[60-3; 60+3] = [57; 63]$$
   cet intervalle est appelé **intervalle de confiance** au seuil de $95\%$. On voit que plus le nombre d'observations est grand plus cette fourchette est étroite.

## 1.5 Tableaux croisés

**Exemple 1.5.1** *Répartition d'un groupe de 40 élèves de BAC ayant un ordinateur personnel ou non.*

- *Le tableau.1.6 donne les effectifs suivant la section SM, SVT ou SE, complétés par les marges en lignes et en colonnes Ainsi, 6 élèves de SM ont un ordinateur personnel, 4 élèves de SE n'en ont pas, 10 élèves sont SVT ... On eput établir le tableau.1.7 donnant les* **fréquences par rapport à l'effectif total** *: chaque cellule est alors divisée par la somme de toutes les cellules.*
  - i) *Les élèves de SM ayant un ordinateur représentant* $15\%$ *du groupe d'élèves.*
  - ii) $5\%$ *des élèves sont des élèves de SVT ayant un ordinnateur.*
  - iii) $10\%$ *des élèves sont des élèves de SVT.*
- *En privilégiant l'un des deux caractères étudiés, on obtient les tableaux suivants :*
  - **Fréquences par rapport aux lignes** *: chaque cellule est divisée par la somme des cellules d'une même ligne. Le tableau.1.8 donne la fréquence de la section selon la possession ou non d'un ordinateur.*
  - **Fréquences par rapport aux colonnes** *: chaque cellule est divisée par la somme des cellules d'une même colonne. Le tableau.1.9 donne la fréquence de la possession d'un ordinateur selon la section.*

a) *Parmi les élèves ayant un ordinateur,* $30\%$ *sont en SM et* $60\%$ *sont en SE.*

b) $20\%$ *des élèves de SE n'ont pas d'ordinateur ; ils sont* $40\%$ *en SVT.*

c) *Les élèves en SVT n'ayant pas d'ordinateur représentent* $40\%$ *des élèves de SVT,* $80\%$ *des élèves n'ayant pas d'ordinateur et* $25\%$ *des élèves du groupe.*

**Propriété 1.5.1** *A partir d'un tableau de répartition selon deux caractères, Ⓡ et Ⓢ, on peut établir trois tableaux (Voir le tableau1.10) :*

- **Tableau des fréquences** *: chaque* **cellule** *est divisée par la somme de toutes les cellules ;*
- **Tableau des fréquences en lignes** *: chaque cellule est divisée par la somme des cellules de la même ligne ; ce tableau donne la fréquence du caractère Ⓢ selon le critère Ⓡ.*
- **Tableau des fréquences en colonnes** *: chaque cellule est divisée par la somme des cellules de la même colonne ; ce tableau donne la fréquence du caractère Ⓡ selon le critère Ⓢ.*

TABLE 1.1 – Données statistiques.

| Numéro | genre | taille | poids | Numéro | genre | taille | poids |
|---|---|---|---|---|---|---|---|
| 1 | M | 121 | 25 | 50 | F | 116 | 20 |
| 2 | M | 122 | 19 | 51 | M | 115 | 20 |
| 3 | M | 121 | 22 | 52 | M | 114,5 | 23 |
| 4 | F | 23,5 | 22 | 53 | F | 110,5 | 20 |
| 5 | M | 117 | 22 | 54 | M | 112 | 20 |
| 6 | F | 108 | 19,5 | 55 | F | 122 | 25 |
| 7 | M | 108 | 24,5 | 56 | F | 122 | 24 |
| 8 | F | 108 | 18 | 57 | F | 114 | 20 |
| 9 | F | 126 | 22 | 58 | M | 112,5 | 20 |
| 10 | F | 121 | 19 | 60 | F | 113,5 | 22 |
| 11 | F | 121 | 20 | 61 | F | 119 | 21 |
| 12 | F | 110,5 | 22 | 62 | M | 113 | 21 |
| 13 | F | 106 | 17 | 63 | M | 115 | 19 |
| 14 | M | 108,5 | 19 | 64 | M | 115,5 | 21 |
| 15 | M | 112 | 20 | 65 | F | 122 | 22 |
| 16 | M | 122,5 | 22 | 66 | F | 114,5 | 18 |
| 17 | M | 118 | 20 | 67 | M | 118,5 | 22 |
| 18 | M | 109 | 17 | 68 | F | 119 | 26 |
| 19 | M | 109 | 18 | 69 | F | 124 | 24 |
| 20 | F | 104 | 17 | 70 | F | 121,5 | 20 |
| 21 | F | 106 | 19 | 71 | F | 104 | 18 |
| 22 | F | 120 | 26 | 72 | M | 108 | 19 |
| 23 | M | 118 | 22 | 73 | F | 109,5 | 20 |
| 24 | F | 114 | 20 | 74 | F | 118 | 24 |
| 25 | M | 114 | 19 | 75 | F | 120 | 24 |
| 26 | M | 111 | 18 | 76 | F | 114 | 22 |
| 27 | M | 116,5 | 18 | 77 | M | 124 | 25 |
| 28 | F | 109,5 | 18 | 78 | M | 125 | 24 |
| 29 | F | 114 | 24 | 79 | M | 108 | 18,5 |
| 30 | M | 121 | 22 | 80 | F | 117 | 20 |
| 31 | F | 116 | 25 | 81 | M | 115 | 19 |
| 32 | F | 109 | 19 | 82 | M | 104 | 16 |
| 33 | F | 114 | 23 | 83 | F | 123 | 25 |
| 34 | F | 113 | 23,5 | 84 | F | 112 | 17 |
| 35 | M | 115 | 19 | 85 | M | 106 | 15 |
| 36 | M | 120 | 22 | 86 | M | 117 | 20 |
| 37 | F | 113 | 19 | 87 | M | 122 | 24 |
| 38 | M | 117 | 18 | 88 | M | 112,5 | 20 |
| 39 | M | 114 | 19 | 89 | M | 120 | 22 |
| 40 | F | 120 | 20 | 90 | M | 117 | 22,5 |
| 41 | F | 109 | 19 | 91 | M | 121 | 23 |
| 42 | F | 111 | 23 | 92 | F | 113 | 17 |
| 43 | F | 111 | 20 | 93 | F | 119 | 20 |
| 44 | F | 107 | 16,5 | 94 | M | 119 | 22 |
| 45 | M | 111 | 17,5 | 95 | M | 117 | 20 |
| 46 | F | 115 | 20 | 96 | F | 121 | 24 |
| 47 | M | 114 | 18 | 97 | F | 109 | 19 |
| 48 | M | 115 | 18,5 | 98 | M | 117 | 22 |
| 49 | M | 116 | 22 | 99 | F | 111 | 17 |
| 50 | F | 116 | 20 | 100 | F | 112 | 15 |

TABLE 1.2 – Répartition des enfants masculins.

| poids \ taille | $[104; 110[$ | $[110; 115[$ | $[115; 120[$ | $[120; 126[$ |
|---|---|---|---|---|
| $[15; 19[$ | | | | |
| $[19; 22[$ | | | | |
| $[22; 26[$ | | | | |

TABLE 1.3 – Répartition des enfants féminins.

| poids \ taille | $[104; 110[$ | $[110; 115[$ | $[115; 120[$ | $[120; 126[$ |
|---|---|---|---|---|
| $[15; 19[$ | | | | |
| $[19; 22[$ | | | | |
| $[22; 26[$ | | | | |

TABLE 1.4 – Tableau des valeurs représentant les calculs pour Massine

| $L_1$ | $L_1 - \overline{X}$ | $(L_1 - \overline{X})^2$ |
|---|---|---|
| 5 | -5 | 25 |
| 7 | -3 | 9 |
| 8 | -2 | 4 |
| 11 | 1 | 1 |
| 14 | 4 | 16 |
| 15 | 5 | 25 |

TABLE 1.5 – Tableau des valeurs représentant les calculs pour Chaïmae

| $L_1$ | $L_1 - \overline{X}$ | $(L_1 - \overline{X})^2$ |
|---|---|---|
| 8 | -2 | 4 |
| 9 | -1 | 1 |
| 9 | -1 | 1 |
| 10 | 0 | 0 |
| 11 | 1 | 1 |
| 13 | 3 | 9 |

TABLE 1.6 – Tableau des profils-lignes et profils-colonnes

| | SM | SVT | SE | total |
|---|---|---|---|---|
| Ordinateur | 6 | 2 | 12 | 20 |
| Non | 8 | 8 | 4 | 20 |
| Total | 14 | 10 | 16 | 40 |

TABLE 1.7 – Tableau représentant les fréquences par rapport à l'effectif total

| | SM | SVT | SE | total |
|---|---|---|---|---|
| Ordinateur | $\frac{6}{40} = 0,15$ | $\frac{2}{40} = 0,05$ | $\frac{12}{40} = 0,30$ | $\frac{20}{40} = 0,50$ |
| Non | $\frac{8}{40} = 0,20$ | $\frac{8}{40} = 0,20$ | $\frac{4}{40} = 0,10$ | $\frac{20}{40} = 0,50$ |
| Total | $\frac{14}{40} = 0,35$ | $\frac{10}{40} = 0,25$ | $\frac{16}{40} = 0,40$ | $\frac{40}{40} = 1$ |

TABLE 1.8 – Tableau représentant les fréquences par rapport aux effectifs totals des lignes

| | SM | SVT | SE | total |
|---|---|---|---|---|
| Ordinateur | $\frac{6}{20}.100 = 30\%$ | $\frac{2}{20}.100 = 10\%$ | $\frac{12}{20}.100 = 60$ | $100\%$ |
| Non | $\frac{8}{20}.100 = 40\%$ | $\frac{8}{20}.100 = 40\%$ | $\frac{4}{20}.100 = 20\%$ | $100\%$ |
| Total | $\frac{14}{40}.100 = 35\%$ | $\frac{10}{40}.100 = 25\%$ | $\frac{16}{40}.100 = 40\%$ | $100\%$ |

TABLE 1.9 – Tableau représentant les fréquences par rapport aux effectifs totals des colonnes

| | SM | SVT | SE | total |
|---|---|---|---|---|
| Ordinateur | $\frac{6}{14}.100 = 42,85\%\%$ | $\frac{2}{10}.100 = 20\%$ | $\frac{12}{16}.100 = 75$ | $50\%$ |
| Non | $\frac{8}{14}.100 = 57,15\%$ | $\frac{8}{10}.100 = 80\%$ | $\frac{4}{16}.100 = 25\%$ | $50\%$ |
| Total | $100\%$ | $100\%$ | $100\%$ | 100 |

TABLE 1.10 – Représentation des tableaux croisés

| S / R | | | | |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |

# Chapitre 2

# Eléments de la théorie des ensembles et dénombrements

## 2.1 Comment décrire un ensemble

Par la suite, nous convenons de noter les ensembles par des lettres majuscules ($\mathbf{A}$, $\mathbf{B}$, $\Omega$,. . .) et les éléments par des lettres miniscules ($a$,$b$,$x$,$y$,$\alpha$,$\beta$,$\gamma$,. . .). Certains ensembles sont déjà connus et reçoivent des dénominations fixes : ainsi, les lettres $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ désignent respectivement l'ensemble des entiers naturels, celui des entiers relatifs, celui des nombres rationnels, celui des réels et celui des nombres complexes.

**Définition 2.1.1** *Un ensemble est une collection d'objets appelés éléments. On note $x \in \mathbf{E}$, $x$ est un élément de $\mathbf{E}$. Il existe deux manières usuelles de décrire les ensembles :*

1. *La définition d'un ensemble par compréhension est celle qui consiste à donner la propriété caractérisant ses éléments.*
   **Par exemple**, *l'ensemble évoqué dans l'introduction s'écrit :*
   $$\mathbf{A} = \{x \in \mathbb{R}/ \quad |x| = 1\} \quad ou \quad \mathbf{A} = \{x \in \mathbb{R}; \quad |x| = 1\},$$
   *ce qui se lit : "$\mathbf{A}$ est l'ensemble des éléments $x$ de $\mathbb{R}$ tels que la valeur absolue de $x$ soit égale à 1".*
   *On notera que $\{u \in \mathbb{R}; \quad |u| = 1\}$ est une autre écriture pour l'ensemble $\mathbf{A}$ : on dit que $x$ (ou $u$) est une* **variable muette**.
2. *La définition d'un ensemble par* **extension** *est celle qui consiste à faire la liste complète de ses éléments.*
   **Par exemple**, *pour écrire l'ensemble des cinq continents, on note ses éléments entre deux accolades :*
   $$\mathbf{C} = \{\text{Afrique}, \text{Amerique}, \text{Asie}, \text{Europe}, \text{Oceanie}\}.$$
   *Cette définition suppose en général que les ensembles en question soient finis et qu'on ait une conception clair de la propriété envisagée, comme dans le cas suivant*
   $$D = \{1, 2, 3, 4, 6, 8, 12, 24\},$$
   *qui est l'ensemble des entiers naturels qui divisent 24.*

**Remarque 2.1.1**
1. *Un élément donné ne doit figurer qu'une seule fois dans l'écriture d'un ensemble défini par extension. Au lieu de $\{1, 5, 6, 1\}$, il faut écrire $\{1, 5, 6\}$.*
2. *Lorsque l'on écrit un ensemble par extension, l'ordre dans lequel figurent les éléments n'a pas d'importance. Ainsi, $\{1, 2, 3\} = \{3, 1, 2\}$.*

## 2.2 Exemples d'ensembles

### 2.2.1 Ensemble vide

**Définition 2.2.1** *On dit qu'un ensemble* $\mathbf{E}$ *qui ne possède aucun élément est un* **ensemble vide**. *On note :* $\mathbf{E} = \emptyset = \{\ \}$.

**Exemple 2.2.1** *Dans* $\mathbb{R}$, *l'ensemble de solutions de l'équation* $x^2 - 2x + 4 = 0$ *est l'ensemble vide, car* $\triangle = b^2 - 4ac = 2^2 - 4 \times 4 = -12 < 0$.
*On écrit* $\mathbb{S} = \emptyset$.

### 2.2.2 Singletons

**Définition 2.2.2** *On appelle* **singleton** *tout ensemble contenant un et un seul élément. Si* $x$ *est un élément dans un ensemble de* $\mathrm{A}$, *le "singleton* $x$*" est noté par* $\{x\}$.

**Exemple 2.2.2** $\mathrm{A} = \{x \in \mathbb{R};\quad e^x = 5\} = \{\ln(5)\}$ *est l'ensemble de solutions de l'équation* $e^x = 5$. *On écrit* $\mathrm{A} = \{\ln(5)\}$.

## 2.3 Sous-ensemble d'un ensemble

**Définition 2.3.1** *On appelle* **sous-ensemble F** *d'un ensemble* $\mathbf{E}$, *toute partie d'éléments contenant dans* $\mathbf{E}$.

**Exemple 2.3.1** $\mathbf{A} = \{1.5, 2, 3, 4, 5\}$ *et* $\mathbf{B} = \{1.5, 5\}$. $\mathbf{B}$ *est un sous -ensemble de* $\mathbf{A}$ *car* $1.5$ *et* $5$ *sont tous les deux dans* $\mathbf{A}$.

**Exercice 2.3.1** *Décrire explicitement tous les sous-ensembles de* $\mathbf{A}$ *de l'exemple 2.3.1.*

**Proposition 2.3.1** *Tout sous-ensemble est un ensemble.*

## 2.4 Ensemble des parties d'un ensemble

### 2.4.1 Définitions et notations

**Définition 2.4.1** *Soit* $\mathbf{E}$ *un ensemble. On appelle ensemble des parties de* $\mathbf{E}$, *l'ensemble noté par* $\mathcal{P}(\mathbf{E})$, *défini par*

$$\mathcal{P}(\mathbf{E}) = \{\mathbf{F}/\mathbf{F}\ est\ un\ sous-ensemble\ de\ \mathbf{E}\} = \{\mathbf{F}/\mathbf{F} \subset \mathbf{E}\}.$$

*Les éléments de l'ensemble* $\mathcal{P}(\mathbf{E})$ *sont des ensembles cette fois-ci. Et, on a* $a \in \mathbf{E} \Leftrightarrow \{a\} \in \mathcal{P}(\mathbf{E})$. *En outre* $\mathcal{P}(\emptyset) = \{\emptyset\}$.

**Remarque 2.4.1** *Pour tout ensemble* $\mathbf{E}$, *on a*

$$\emptyset \in \mathcal{P}(\mathbf{E}) \quad et \quad \mathbf{E} \in \mathcal{P}(\mathbf{E}).$$

**Exemple 2.4.1** *Soit* $\mathbf{E} = \{0, 3, 7\}$. *L'ensemble des parties de* $\mathbf{E}$ *est*

$$\mathcal{P}(\mathbf{E}) = \{\emptyset, \{0\}, \{3\}, \{7\}, \{0, 3\}, \{0, 7\}, \{3, 7\}, \mathbf{E}\}$$

### 2.4.2 Complémentaire d'un ensemble dans un autre

**Définition 2.4.2** *On appelle complèmentaire de* $\mathbf{A}$ *dans* $\mathbf{E}$ *le sous-ensemble, noté* $\overline{\mathbf{A}}$ *ou* $C_{\mathbf{E}}(\mathbf{A})$ *, défini par*

$$\overline{\mathbf{A}} = \{x \in \mathbf{E}/ \quad x \notin \mathbf{A}\}.$$

**Exemple 2.4.2** *Si* $\mathbf{E}$ *est l'ensemble des jours de la semaine et* $\mathbf{A}$ *le sous-ensemble des jours où les enfants vont à l'école, alors*

$$\overline{\mathbf{A}} = \{\text{mercredi}, \text{samedi}, \text{dimanche}\}.$$

**Théorème 2.4.1** *Soient* $\mathbf{A}$ *et* $\mathbf{B}$ *deux ensembles inclus dans un ensemble* $\mathbf{E}$*. Alors, on a*

1. $\mathbf{B} = \overline{\mathbf{A}}$ *si, et seulement si* $\mathbf{A} \cap \mathbf{B} = \emptyset$ *et* $\mathbf{A} \cup \mathbf{B} = \mathbf{E}$*.*
2. *Si* $\mathbf{A} \subset \mathbf{B}$*, alors* $\mathbf{A} \cap \mathbf{B} = \mathbf{A}$ *et* $\mathbf{A} \cup \mathbf{B} = \mathbf{B}$*.*

**Théorème 2.4.2** *(***Lois de De Morgan***)*
*Soient* $\mathbf{A}$ *et* $\mathbf{B}$ *deux ensembles inclus dans un ensemble* $\mathbf{E}$*. Soit* $(\mathrm{E}_i)_{1 \leq i \leq n}$ *une famille de sous-ensembles de* $\mathbf{E}$*. Alors, on a*

1. $\overline{\mathbf{A} \cap \mathbf{B}} = \overline{\mathbf{A}} \cup \overline{\mathbf{B}}$*.*
2. $\overline{\mathbf{A} \cup \mathbf{B}} = \overline{\mathbf{A}} \cap \overline{\mathbf{B}}$*.*
3. $\overline{\overline{\mathbf{A}}} = \mathbf{A}$*.*

*Et plus généralement pour une famille* $(\mathrm{E}_i)_{1 \leq i \leq n}$ *de sous-ensembles de* $\mathbf{E}$*, on a*

$$\overline{\bigcap_{i=1}^{n} \mathrm{E}_i} = \bigcup_{i=1}^{n} \overline{\mathrm{E}_i} \quad et \quad \overline{\bigcup_{i=1}^{n} \mathrm{E}_i} = \bigcap_{i=1}^{n} \overline{\mathrm{E}_i}.$$

**Démonstration.**

□

### 2.4.3 Différence et différence symétrique

**Définition 2.4.3** *On appelle différence de* $\mathbf{A}$ *par* $\mathbf{B}$ *le sous-ensemble, noté* $\mathbf{A} \backslash \mathbf{B}$ *, défini par*

$$\mathbf{A} \backslash \mathbf{B} = \{x \in \mathbf{E}/ \quad x \in \mathbf{A} \text{ et } x \notin \mathbf{B}\}.$$

**Remarque 2.4.2** *La différence de* $\mathbf{B}$ *par* $\mathbf{A}$ *est noté* $\mathbf{B} \backslash \mathbf{A}$*, définie par*

$$\mathbf{B} \backslash \mathbf{A} = \{x \in \mathbf{E}/ \quad x \in \mathbf{B} \text{ et } x \notin \mathbf{A}\}.$$

*En général,* $\mathbf{A} \backslash \mathbf{B} \neq \mathbf{B} \backslash \mathbf{A}$*.*

**Exemple 2.4.3**
1. *L'ensemble des nombres complexes non réels se note* $\mathbb{C} \backslash \mathbb{R}$*. Il ne faut sur tout pas le confondre avec l'ensemble* $\{\mathrm{i}x | x \in \mathbb{R}\}$ *des imaginaires purs, noté* $\mathrm{i}\mathbb{R}$*, qui est inclus dans* $\mathbb{C} \backslash \mathbb{R}$*, mais ne lui est pas égal. Par exemple,* $1 + 2\mathrm{i} \in \mathbb{C} \backslash \mathbb{R}$ *mais* $1 + 2\mathrm{i} \notin \mathrm{i}\mathbb{R}$*.*
2. *L'ensemble des nombres réels non rationnel se note* $\mathbb{R} \backslash \mathbb{Q}$ *qui est l'ensemble des irrationnels.*

**Définition 2.4.4** *On appelle différence symétrique ou réunion disjonctive de deux ensembles* $\mathbf{A}$ *et* $\mathbf{B}$ *le sous-ensemble, noté* $\mathbf{A} \triangle \mathbf{B}$ *, défini par*

$$\mathbf{A} \triangle \mathbf{B} = (\mathbf{A} \cup \mathbf{B}) \backslash (\mathbf{A} \cap \mathbf{B}) = (\mathbf{A} \backslash \mathbf{B}) \cup \mathbf{B} \backslash \mathbf{A}).$$

**Théorème 2.4.3** *Soient* $\mathbf{A}$*,* $\mathbf{B}$ *et* $\mathbf{C}$ *trois ensembles. Alors, on a*
$\mathbf{A} \triangle \emptyset = \mathbf{A}$*,* $\mathbf{A} \triangle \mathbf{B} = \mathbf{B} \triangle \mathbf{A}$*,* $\mathbf{A} \triangle \mathbf{A} = \emptyset$ *et* $(\mathbf{A} \triangle \mathbf{B}) \triangle \mathbf{C} = \mathbf{A} \triangle (\mathbf{B} \triangle \mathbf{C})$*.*

## 2.5 Cardinal d'un ensemble

**Définition 2.5.1** *Soit* $\mathbf{E}$ *un ensemble. On appelle cardinal de* $\mathbf{E}$*, le nombre des éléments de l'ensemble* $\mathbf{E}$*, on le note par* $\mathrm{Card}(\mathbf{E})$ *ou encore* $|\mathbf{E}|$.

**Exemple 2.5.1** *Soit* $\mathbf{E} = \{0, 3, 7\}$. *Le cardinal de* $\mathbf{E}$ *est égal à 3. On écrit* $\mathrm{Card}(\mathbf{E}) = |\mathbf{E}| = 3$.

**Théorème 2.5.1** *Soit* $\mathbf{E}$ *un ensemble. Soient* $\mathbf{A}$ *et* $\mathbf{B}$ *deux sous-ensembles de* $\mathbf{E}$. *Si* $\mathbf{A}$ *et* $\mathbf{B}$ *sont des ensembles finis, alors :*

1. *La réunion* $\mathbf{A} \cup \mathbf{B}$ *est aussi un ensemble fini et on a*
$$\mathrm{Card}(\mathbf{A} \cup \mathbf{B}) \geq \max(\mathrm{Card}(\mathbf{A}), \mathrm{Card}(\mathbf{B})).$$
2. *L'intersection* $\mathbf{A} \cap \mathbf{B}$ *est aussi un ensemble fini et on a*
$$\mathrm{Card}(\mathbf{A} \cap \mathbf{B}) \leq \min(\mathrm{Card}(\mathbf{A}), \mathrm{Card}(\mathbf{B})).$$
3. $\mathrm{Card}(\mathbf{A} \cup \mathbf{B}) = \mathrm{Card}(\mathbf{A}) + \mathrm{card}(\mathbf{B}) - \mathrm{Card}(\mathbf{A} \cap \mathbf{B})$.

**Démonstration.**

□

**Théorème 2.5.2** *Soit* $\mathbf{E}$ *un ensemble fini. Soient* $\mathbf{A}$ *et* $\mathbf{B}$ *deux sous-ensembles de* $\mathbf{E}$. *alors :*

1. *Toute partie* $\mathbf{A}$ *de* $\mathbf{E}$ *est un ensemble fini et on a*
$$\mathrm{Card}(\mathbf{A}) \leq \mathrm{Card}(\mathbf{E}).$$
2. *Si* $\mathrm{Card}(\mathbf{E}) = n$*, alors* $\mathcal{P}(\mathbf{E})$ *est un ensemble fini à* $2^n$ *éléments. Autrement dit*
$$\mathrm{Card}(\mathcal{P}(\mathbf{E})) = 2^{\mathrm{Card}(\mathbf{E})} = 2^{|\mathbf{E}|}.$$

**Démonstration.**

□

**Exemple 2.5.2** *Soit* $\mathbf{E} = \{0, 3, 7\}$. *L'ensemble des parties de* $\mathbf{E}$ *est*
$$\mathcal{P}(\mathbf{E}) = \{\emptyset, \{0\}, \{3\}, \{7\}, \{0, 3\}, \{0, 7\}, \{3, 7\}, \mathbf{E}\}$$
*Alors,* $\mathrm{Card}(\mathbf{E}) = 3$, $\mathrm{Card}(\mathcal{P}(\mathbf{E})) = 8$ *et* $2^{\mathrm{Card}(\mathbf{E})} = 2^3 = 8$,
*d'où* $\mathrm{Card}(\mathcal{P}(\mathbf{E})) = 2^{\mathrm{Card}(\mathbf{E})}$.

## 2.6 Fonction caractéristique d'un ensemble

**Définition 2.6.1** *On appelle fonction caractéristique d'un sous-ensemble* $\mathbf{A}$ *de* $\mathbf{E}$*, la fonction, noté* $\chi_{\mathbf{A}}$*, définie par*
$$\begin{array}{rcl} \chi_{\mathbf{A}} : \mathbf{A} & \longrightarrow & \{0, 1\} \\ x & \longmapsto & \chi_{\mathbf{A}}(x) = \left\{ \begin{array}{ll} 1, & si \quad x \in \mathbf{A}, \\ 0, & si \quad x \notin \mathbf{A}. \end{array} \right. \end{array}$$

**Propriété 2.6.1** *Soient* $\mathbf{A}$ *et* $\mathbf{B}$ *deux ensembles inclus dans un ensemble* $\mathbf{E}$.

1. $\mathbf{A} = \mathbf{B} \Leftrightarrow \chi_{\mathbf{A}} = \chi_{\mathbf{B}}$.
2. $\mathbf{A} \subset \mathbf{B} \Leftrightarrow \chi_{\mathbf{A}} \leq \chi_{\mathbf{B}}$.
3. $\chi_{\overline{\mathbf{A}}} = 1 - \chi_{\mathbf{A}}$.
4. $\chi_{\mathbf{A} \cap \mathbf{B}} = \chi_{\mathbf{A}} \chi_{\mathbf{B}}$.
5. $\chi_{\mathbf{A} \cup \mathbf{B}} = \chi_{\mathbf{A}} + \chi_{\mathbf{B}} - \chi_{\mathbf{A}} \chi_{\mathbf{B}}$.

**Démonstration.**

□

## 2.7 Dénombrements

**Définition 2.7.1** *Soit $E$ un ensemble fini de cardinal $n$ et $I_p$ l'ensemble $\{1, 2, 3, \dots, p\}$ ($p \in \mathbb{N}^*$).*

1. *On appelle arrangement $p$ à $p$ des $n$ éléments de $E$ toute injection de $I_p$ dans $E$.*
2. *L'image de $I_p$ par un arrangement s'appelle combinaison $p$ à $p$ des $n$ éléments de $E$.*
3. *Une combinaison est donc une partie $E_p$ de $E$ à $p$ éléments. On note $\mathrm{A}_n^p$ le nombre des arrangements de $n$ éléments $p$ à $p$ , et $\mathrm{C}_n^p$ le nombre des combinaisons.*

**Propriété 2.7.1** *Soit $p$ et $n$ dans $\mathbb{N}^*$.*

1. $(p > n) \quad \Rightarrow \quad (\mathrm{A}_n^p = 0 \quad \textit{et} \quad \mathrm{C}_n^p = 0)$ ;
2. $(p \leq n) \quad \Rightarrow \quad \mathrm{A}_n^p = n(n-1)\dots(n-p+1) = \dfrac{n!}{(n-p)!}$ ;

$$\mathrm{C}_n^p = \frac{\mathrm{A}_n^p}{\mathrm{A}_p^p} = \frac{n(n-1)\dots(n-p+1)}{p(p-1)\dots 2.1} = \frac{n!}{p!(n-p)!}$$

**Théorème 2.7.1** *Soit $p$ et $q$ deux entiers non nuls. Le nombre des applications d'un ensemble $E_p$ quelconque de cardinal $p$ dans un ensemble $F_q$ quelconque de cardinal $q$ est $q^p$ :*

$$\mathrm{Card}(\mathcal{F}(E_p, F_q)) = q^p.$$

**Démonstration.**

□

# Chapitre 3

# Calcul des probabilités

## 3.1 Espace de probabilité discret

### 3.1.1 Espace probabilisable

**Définition 3.1.1** *Soit $\Omega$ un ensemble et $\mathcal{P}(\Omega)$ l'ensemble des parties de $\Omega$.*
*On appelle algèbre de Boole de parties de $\Omega$ tout sous-ensemble non vide $\mathcal{B}$ de $\mathcal{P}(\Omega)$ stable par l'***intersection** *et la* **complémentation**.
*Autrement dit*

1. *$(\forall (A,B) \in \mathcal{B}^2)$, $A \cap B \in \mathcal{B}$ ;*
2. *$(\forall A \in \mathcal{B})$, $\overline{A} = \complement_\Omega^A \in \mathcal{B}$.*

**Théorème 3.1.1** *Une algèbre de Boole de parties d'un ensemble $\Omega$ est stable pour l'opération $(A,B) \longmapsto A \cup B$. De plus, il contient la partie pleine de $\Omega$ et la partie vide $\emptyset$.*

**Démonstration.**

1. $\mathcal{B}$ étant non vide, contient un élément $A$, contient donc $\overline{A}$, et donc $A \cap \overline{A} \in \mathcal{B}$, c'et-à-dire que $\emptyset \in \mathcal{B}$.
2. $\emptyset \in \mathcal{B}$, alors $\overline{\emptyset} = \complement_\Omega^\emptyset = \Omega \in \mathcal{B}$.
3. Soit $A$ et $B$ deux éléments de $\mathcal{B}$. $\mathcal{B}$ contient $\overline{A}$ et $\overline{B}$, donc $\overline{A} \cap \overline{B} \in \mathcal{B}$, donc $\overline{\overline{A} \cap \overline{B}} \in \mathcal{B}$, qui n'est autre que $\overline{\overline{A} \cap \overline{B}} = A \cup B$, d'où $A \cup B \in \mathcal{B}$ (Loi de De Morgan).

$\square$

**Exercice 3.1.1** *Montrer que l'intersection (et la réunion) d'une famille finie d'éléments d'une algèbre de Boole $\mathcal{B}$ appartient à $\mathcal{B}$.*

**Exemple 3.1.1**
1. *$\mathcal{P}(\Omega)$ est une algèbre de Boole de parties de $\Omega$. On l'appelle l'***algèbre de Boole discrète**.
2. *$\{\emptyset, \Omega\}$ est une algèbre de Boole de parties de $\Omega$. On l'appelle l'***algèbre de Boole grossière**.
3. *Si $\Omega$ a au moins deux éléments, soit $A$ une partie propre (c'est-à-dire distincte de $\Omega$ et de $\emptyset$) de $\Omega$. Alors $\{\emptyset, A, \overline{A}, \Omega\}$ est une algèbre de Boole de parties de $\Omega$.*

**Définition 3.1.2** *Soit $\Omega$ un ensemble et soit $\tau \subset \mathcal{P}(\Omega)$.*
*$\tau$ est dite* **tribu** *où bien $\sigma$-***algèbre** *sur $\Omega$ si elle vérifie les conditions suivantes :*

*i)* $\Omega \in \tau$,

*ii)* *si $A \in \tau$, alors $\overline{A} = \complement_\Omega^A \in \tau$,* *iii) si* $(A_i)_{i\in\mathbb{N}} \subset \tau$*, alors* $(\bigcup_{i\in\mathbb{N}} A_i) \in \tau$.

**Exemple 3.1.2** *1.* $\mathcal{P}(\Omega)$ *est une* $\sigma$*-algèbre de parties de* $\Omega$*. On l'appelle la* $\sigma$**-algèbre discrète**.

*2.* $\{\emptyset, \Omega\}$ *est une* $\sigma$*-algèbre de parties de* $\Omega$*. On l'appelle la* **tribu grossière**.

*3. Si* $\Omega$ *a au moins deux éléments, soit* $A$ *une partie propre (c'est-à-dire distincte de* $\Omega$ *et de* $\emptyset$*) de* $\Omega$*. Alors* $\tau = \{\emptyset, A, \overline{A}, \Omega\}$ *est une tribu de parties de* $\Omega$.

*En particulier, la tribu grossière est la plus petite tribu et la tribu discrète est la plus grande tribu sur un ensemble* $\Omega$*. Et si* $\tau$ *était une tribu sur* $\Omega$ *alors*

$$\{\emptyset, \Omega\} \subset \tau \subset \mathcal{P}(\Omega).$$

**Définition 3.1.3** *Soit* $\Omega$ *un ensemble et* $\tau$ *une tribu sur* $\Omega$.
*On appelle* **espace probabilisable** *le couple* $(\Omega, \tau)$ *formé par un ensemble* $\Omega$ *et une* $\sigma$*-algèbre* $\tau$ *de parties de* $\Omega$.

- $\Omega$ *est appelé* **univers**.
- *Les éléments de* $\Omega$ *sont appelés* **éventualités**.
- *Les éléments de* $\tau$ *sont appelés* **événements**.
- *L'événement* $\emptyset$ *est appelé événement* **impossible**.
- *L'événement* $\Omega$ *est appelé événement* **certain**.
- *Deux événements* $A$ *et* $B$ *tels que* $A \cap B = \emptyset$ *sont dits* **incompatibles**.

**Un espace probabilisable* $(\Omega, \tau)$ *est dit* **fini** *si* $\Omega$ *est un* **ensemble fini**.

### 3.1.2 Atomes

**Définition 3.1.4** *Soit* $(\Omega, \tau)$ *un espace probabilisablee fini. Deux éventualités* $\omega$ *et* $\omega'$ *sont dites inséparables par* $\tau$ *si et seulement si, tout événement qui contient l'une contient l'autre.*

**Exemple 3.1.3** *Soit* $\Omega = \{a, b, c\}$*,* $\tau = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$ *:* $\tau$ *et une tribu sur* $\Omega$ *et* $b$ *et* $c$ *sont inséparables par* $\tau$.

**Théorème 3.1.2** *Soit* $(\Omega, \tau)$ *un espace probabilisable fini.*
*La relation définie sur* $\Omega$ *par "*$\omega$ *et* $\omega'$ *sont inséparables par* $\tau$*" est une relation d'équivalence.*
*Les classes d'équivalence sont des événements, qu'on appelle les atomes de* $\tau$.

**Remarque 3.1.1** *Les atomes de l'espace probabilisable* $(\Omega, \mathcal{P}(\Omega))$ *ne sont autres que les événements* $\{\omega\}$ *réduits à une éventualité.*

## 3.2 Espace probabilisés

**Définition 3.2.1** *Soit* $(\Omega, \tau)$ *un espace probabilisable.*
*On appelle* **probabilité** *sur* $(\Omega, \tau)$ *toute application* $P$ *de* $\tau$ *dans l'intervalle* $[0, 1]$ *de* $\mathbb{R}$ *vérifiant les axiomes suivants :*

$(A_1)$ $P(\Omega) = 1$.

$(A_2)$ $P$ *est* $\sigma$*-additive :*

$$(A_n)_n \subset \tau, \quad A_i \cap A_j = \emptyset,\ i \neq j \quad \Rightarrow \quad P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

**Si* $\Omega$ *est un ensemble fini, alors le triplet* $(\Omega, \tau, P)$ *est appelé un* **espace probabilisé** *fini.*

### 3.2.1 Propriétés élémentaires

Les axiomes entraînent immédiatement un certain nombre de propriétés élémentaires, vérifiées dans tout espace probabilisé $(\Omega, \tau, P)$ :

1. La probabilité de l'événement impossible est nulle, c'est-à-dire que

$$P(\Omega) = 1 = P(\emptyset) + P(\mathrm{C}_{\Omega}^{\emptyset}) = P(\emptyset) + 1 \quad \Rightarrow \quad P(\emptyset) = 0$$

2. Soit $\{A_1, A_2, \ldots, A_n\}$ un ensemble fini d'événements **deux à deux incompatibles** ; on a

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

ceci traduit l'axiome $(A_2)$ dans le cas d'un ensemble fini d'événements.

3. Si $A$ et $B$ sont deux événements contraires ($B = \overline{A}$), alors la somme de leurs probabilités est égale à 1 :

$$P(\Omega) = 1 = P(A) + P(\overline{A})$$

car $\Omega = A \cup \overline{A}$.

4. Nous avons la propriété suivante :
**Théorème 3.2.1** *Soitent $A$ et $B$ deux événements quelconques d'un espace probabilisé $(\Omega, \tau, P)$, on a*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

**Démonstration.**Quels que soient les événements $A$ et $B$, on peut écrire

$$A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (\overline{A} \cap B)$$

et les trois événements $A \cap \overline{B}$, $A \cap B$ et $\overline{A} \cap B$ sont deux à deux incompatibles. Donc

$$P(A \cup B) = P(A \cap \overline{B}) + P(A \cap B) + P(\overline{A} \cap B);$$

d'autre part, $A$ est la réunion de deux événements incompatibles :

$$A = (A \cap \overline{B}) \cup (A \cap B),$$

donc $P(A) = P(A \cap \overline{B}) + P(A \cap B)$,
ce qui s'écrit aussi

$$P(A \cap \overline{B}) = P(A) - P(A \cap B);$$

de la même façon, on établirait que

$$P(\overline{A} \cap B) = P(B) - P(A \cap B);$$

reportant ces deux valeurs dans l'expression de $P(A \cup B)$, on obtient bien

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

□

5. Si $A \subset B$, alors on a

$$P(B) = P(A) + P(B - A), \quad \text{d'où} \quad P(A) \leq P(B)$$

**Définition 3.2.2** 1. *On appelle événement presque impossible tout événement $A$ de probabilité nulle.*
2. *On appelle événement presque certain tout événement $A$ de probabilité égale à 1.*

### 3.2.2 Donnée pratique d'une probabilité

Soit $(\Omega, \tau)$ un espace probabilisable fini et $A_1, A_2, \ldots, A_n$ les atomes de $\tau$.

- Tout événement étant réunion d'un nombre fini d'atomes, deux probabilités $P_1$ et $P_2$ qui sont égales pour chaque atome sont égales pour chaque événement.
- Pour connaître une probabilité , il n'est donc pas nécessaire de connaître les probabilités de tous les événements : celles des atomes suffisent ! !
  **Inversement**, soit $p_1, p_2, \ldots, p_n$, $n$ nombres réels. A quelle condition existe-t-il une probabilité $P$ (qui serait alors unique d'après la remarque qui précède) telle que
  $$(\forall i \in \{1, 2, \ldots, n\}) \quad P(A_i) = p_i?$$
  Les conditions
  (i) $(\forall i \in \{1, 2, \ldots, n\})$, $0 \leq p_i \leq 1$ ;
  (ii) $1 = P(\Omega) = \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{n} p_i$
  sont nécessaires.

**Définition 3.2.3** *Un espace de probabilité $(\Omega, \tau, P)$ est dit discret si $\tau = \mathcal{P}(\Omega)$ est la tribu discrète.*

## 3.3 Probabilité conditionnelle

**Définition 3.3.1** *Soit $(\Omega, \tau, P)$ un espace probabilisé et $B$ un événement de probabilité non nulle. L'application*

$$P_B : \tau \longrightarrow [0, 1], \quad A \longmapsto \frac{P(A \cap B)}{P(B)}$$

*est une probabilité sur $(\Omega, \tau)$ appelée probabilité conditionnelle par $B$.*

**Propriété 3.3.1** *Soit $(\Omega, \tau, P)$ un espace de probabilisé et $A$ et $B$ deux événements de probabilités non nulles. Alors on a la formule des probabilités composées :*

$$P(A \cap B) = P(B)P_B(A) = P(A)P_A(B).$$

**Démonstration.**

□

## 3.4 Événements indépendants et indépendance mutuelle

### 3.4.1 Événements indépendants

**Définition 3.4.1** *Soit $(\Omega, \tau, P)$ un espace probabilisé. deux événements $A$ et $B$ tels que*

$$P(A \cap B) = P(A)P(B)$$

*sont appelés indépendants.*

| | $a$ | $b$ | $c$ | $d$ | $e$ |
|---|---|---|---|---|---|
| $P_1$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/3 |
| $P_2$ | 1/10 | 2/10 | 3/10 | 1/10 | 3/10 |

**Exemple 3.4.1** *Soit $\Omega = \{a, b, c, d, e\}$ et $\tau = \mathcal{P}(\Omega)$. On définit deux probabilités $P_1$ et $P_2$ sur l'espace $(\Omega, \mathcal{P}(\Omega))$ par le tableau suivant :*

*et les événements*

$$a = \{a, b, c\}, \quad B = \{c, d\}, \quad A \cap B = \{c\};$$

*alors*

$$P_1(A) = \frac{3}{6}; \quad P_1(B) = \frac{2}{6}; \quad P_1(A \cap B) = \frac{1}{6}.$$

*Dans $(\Omega, \mathcal{P}(\Omega), P_1)$, $A$ et $B$ sont indépendants ($\frac{3}{6} \times \frac{2}{6} = \frac{1}{6}$).*
*D'autre part*

$$P_2(A) = \frac{6}{10}; \quad P_2(B) = \frac{4}{10}; \quad P_2(A \cap B) = \frac{3}{10}.$$

*Dans $(\Omega, \mathcal{P}(\Omega), P_2)$, $A$ et $B$ ne sont pas indépendants ($\frac{6}{10} \times \frac{4}{10} \neq \frac{3}{10}$).*

### 3.4.2 Indépendance mutuelle

**Définition 3.4.2** *Soit $(\Omega, \tau, P)$ un espace probabilisé. Des événements $A_1, A_2, \ldots, A_p$ sont dits mutuellement indépendants si, pour tout entier $k$ et pour tout sous-ensemble $\{A_{i_1}, \ldots, A_{i_k}\}$ de $\{A_1, A_2, \ldots, A_p\}$ formé d'événements distincts, on a la relation*

$$P(A_{i_1} \cap \ldots \cap \ldots, A_{i_k}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k})$$

Il résulte immédiatement de cette définition que si des événements sont mutuellement indépendants, ils sont indépendants deux à deux. **La réciproque est fausse**, comme le montre l'exemple suivant

**Exemple 3.4.2** *On tire une carte d'un jeu de 52 et on considère les événements suivants :*
- *$A$ : la carte tirée est rouge ;*
- *$B$ : la carte tirée est une majeure (cœur ou pique) ;*
- *$C$ : la carte tirée est un pique ou un carreau.*

*Alors les événements*
- *$A \cap B$ : la carte tirée est un cœur,*
- *$A \cap C$ : la carte tirée est un carreau,*
- *$B \cap C$ : la carte tirée est un pique.*

*ont chacun pour probabilité $\frac{1}{4}$, tandis que*

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$

*Les événements $A$, $B$, $C$ sont donc indépendants deux à deux.*
*Mais*

$$A \cap B \cap C = \emptyset,$$

*et donc $P(A \cap B \cap C) = 0$,*
*tandis que*

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

*Les événements $A$, $B$, $C$ ne sont donc pas mutuellement indépendants.*

## 3.5 Formule de Bayes et applications

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé et des événements $C_1, C_2, \ldots, C_n$ constituant une partition de $\Omega$, c'est-à-dire

$$(\forall i \in \{1,2,\ldots,n\},\ \forall j \in \{1,2,\ldots,n\}) \quad (i \neq j) \ \Rightarrow \ (C_i \cap C_j = \emptyset) \ \text{ et } \ \Omega = C_1 \cup C_2 \cup \ldots \cup C_n.$$

Nous supposons qu'aucun événement $C_i$ n'est presque impossible.
Soit maintenant $A$ un événement non presque impossible :

$$A = A \cap \Omega = (A \cap C_1) \cup (A \cap C_2) \cup \ldots \cup (A \cap C_n).$$

Les événements $A \cap C_i$ sont deux à deux incompatibles, et par suite :

$$P(A) = P(A \cap C_1) + P(A \cap C_2) + \ldots + P(A \cap C_n).$$

D'autre part, pour un indice $k$ donné :

$$P_A(C_k) = \frac{P(A \cap C_k)}{P(A)} = \frac{P(C_k)P_{C_k}(A)}{P(A)},$$

d'où l'on tire

$$P_A(C_k) = \frac{P(C_k)P_{C_k}(A)}{P(C_1)P_{C_1}(A) + P(C_2)P_{C_2}(A) + \ldots + P(C_n)P_{C_n}(A)}.$$

Cette formule est connue sous le nom de **formule de Bayes**.

**Exemple 3.5.1** *Un examen se compose de questions auxquelles il faut répondre par* **oui** *ou* **non**. *Si un étudiant connaît la réponse, il répond correctement ; s'il l'ignore, il tire à pile ou face la réponse qu'il inscrira.*
*Un étudiant donné connaît* $60\%$ *du programme : Quelle est la probabilité pour qu'une réponse juste soit due à ses connaissances plutôt qu'au hasard ?*
**Réponse** *: Soit les événements :* $C_1$ *: l'étudiant connaît la réponse ;*
$C_2$ *: l'étudiant ne connaît pas la réponse ;*
$A$ *: l'étudiant répond juste.*
*L'énoncé nous donne :* $P_{C_1}(A) = 1$, $P_{C_2}(A) = 0.5$, $P(C_1) = 0.6$ *et* $P(C_2) = 0.4$.
*La formule de Bayes fournit alors la probabilité cherchée :*

$$P_A(C_1) = \frac{P(C_1)P_{C_1}(A)}{P(C_1)P_{C_1}(A) + P(C_2)P_{C_2}(A)},$$

*soit*

$$P_A(C_1) = \frac{1 \times 0.6}{1 \times 0.6 + 0.5 \times 0.4} = \frac{3}{4} = 0.75$$

# Chapitre 4

# Variables aléatoires : discrètes et continues

## 4.1 Les variables aléatoires

### 4.1.1 Applications mesurables

Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisable et $X$ une application de $\Omega$ dans un ensemble $J$. Si $A$ est une partie de $J$, on notera $X^{-1}(A)$ l'ensemble des éventualitéés dont l'image appartenant à $A$.

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

**Définition 4.1.1** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisable fini. Une application $X$ de $\Omega$ dans un ensemble $J$ est dite* **mesurable** *si, et seulement si, pour tout point $x$ de $J$, $X^{-1}(\{x\})$ est un événement :*

$$\forall x \in J \quad X^{-1}(\{x\}) \in \mathcal{B}.$$

*Au lieu de "application mesurable", on dit aussi "***variable aléatoire***"*

**Remarque 4.1.1** *1. Si $J = \mathbb{R}^n$ alors*

*(a) dans le cas $n = 1$, on dit "***variable aléatoire réelle***",*

*(b) dans le cas $n > 1$, on dit "***variable aléatoire vectorielle***".*

*2. Si $\mathcal{B} = \mathcal{P}(\Omega)$, toute application de $\Omega$ dans $J$ est mesurable.*

### 4.1.2 Événements liés à une application mesurable

**Théorème 4.1.1** *Soit $(\Omega, \mathcal{B})$ un espace probabilisable et $X$ une application mesurable de $\Omega$ dans un ensemble $J$. Pour toute partie $A$ de $J$, on a $X^{-1}(A)$ est un événement.*

**Démonstration.**

1. Si $\Omega$ est fini, alors $X(\Omega) \cap A$ est fini.
   Soit donc $X(\Omega) \cap A = \{a_1, a_2, \ldots, a_n\}$ (les $a_i$ sont distincts).
   $X^{-1}(A)$ est alors la réunion d'un nombre fini d'événements

$$X^{-1}(\{a_1\}), \; X^{-1}(\{a_2\}), \; \ldots \; X^{-1}(\{a_n\}).$$

   $X^{-1}(A)$ est donc un événement
2. Si $\Omega$ n'est pas fini, on a $\Omega$ est la réunion d'atome de $\mathcal{B}$ et dans ce cas $X$ est constante sur tout atome.

$\square$

**Définition 4.1.2** *Soit $X : \Omega \longrightarrow J$ une variable aléatoire. Pour toute partie $A$ de $J$, $X^{-1}(A)$ est appelé événement. On appelle* **indicatrice d'un événement** *$A$, la fonction $\varphi_A$ définie par*

$$(\forall \omega \in \Omega) \quad \begin{cases} \varphi_A(\omega) = 1 & si \quad \omega \in A, \\ \varphi_A(\omega) = 0 & si \quad \omega \notin A, \end{cases}$$

### 4.1.3 Probabilité image

Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et $X : \Omega \to J$ une variable aléatoire. Soit l'espace probabilisable $(J, \mathcal{P}(J))$ où $\mathcal{P}$ est l'ensemble des parties mesurables de $J$. On considère l'application

$$Q : \mathcal{P}(J) \to [0,1], \quad A \mapsto P(X^{-1}(A)) \tag{4.1}$$

A toute partie $A$ de $J$ associons le nombre $Q(A)$, probabilité de l'événement $X^{-1}(A)$.

Le théorème suivant montre que $X$ permet en quelque sorte de transporter une probabilité de $(\Omega, \mathcal{B})$ vers $(J, \mathcal{P}(J))$

**Théorème 4.1.2** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et $X : \Omega \to J$ une variable aléatoire. L'application $Q$ définie en (4.1) est une probabilité sur $(J, \mathcal{P}(J))$.*

**Démonstration.** $Q$ est bien une application de $\mathcal{P}(J)$ dans $[0,1]$.

i) $X^{-1}(J) = \Omega$ donc $Q(J) = P(X^{-1}(J)) = P(\Omega) = 1$.

ii) Soit $K$ et $L$ deux parties disjointes de $J$ :
- $K \cap L = \emptyset$ et donc $X^{-1}(K) \cap X^{-1}(L) = \emptyset$,
- $X^{-1}(K \cup L) = X^{-1}(K) \cup X^{-1}(L)$ ;

par suite

$$Q(K \cup L) = P(X^{-1}(K \cup L)) = P(X^{-1}(K) \cup X^{-1}(L));$$

or $X^{-1}(K)$ et $X^{-1}(L)$ sont deux événements disjoints et donc

$$P(X^{-1}(K \cup L)) = P(X^{-1}(K)) + P(X^{-1}(L)) = Q(K) + Q(L).$$

On a donc montré que

$$(K \cap L = \emptyset) \quad \Rightarrow \quad (Q(K \cup L) = Q(K) + Q(L)).$$

$\square$

**Définition 4.1.3** *$Q$ est bien une probabilité sur $(J, \mathcal{P}(J))$ : on l'appelle la* **probabilité image** *de $P$ par $X$.*

### 4.1.4 Loi de probabilité

**Définition 4.1.4** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et $X : \Omega \to J$ une variable aléatoire vectorielle. On appelle* **loi de probabilité** *de $X$ la probabilité image de $P$ par $X$, c'est à dire*

$$Q(A) = P(X^{-1}(A)).$$

Une loi de probabilité est donc une probabilité sur l'espace probabilisable $(J^n, \mathcal{P}(J^n))$ où $J^n$ désigne une partie de $\mathbb{R}^n$ contenant $X(\Omega)$. Nous utiliserons beaucoup cette notion dans le cas $n = 1$ (variable aléatoire réelle).

**Définition 4.1.5** *On appelle variable aléatoire discrète une application $X : \Omega \to F$ où $F$ est un ensemble fini ou dénombrable.*
**Pour $x \in F$, on note de façcon concise $[X = k]$ l'événement $\{\omega \in \Omega \ / \ X(\omega) = k\}$.*
**La famille des nombres réels $(P([X = k]))_{k \in F}$ s'appelle la* **loi** *de $X$.*

### 4.1.5 Fonction de répartition

Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et $X : \Omega \to J$ une variable aléatoire. On appelle fonction de répartition de $X$ la fonction réelle de variable réelle

$$x \mapsto F_X(x) = P([X < x])$$

où l'ensemble $[X < x] = X^{-1}(]-\infty; x[)$. On a

$$F_X(x) = P(X^{-1}(]-\infty; x[))$$

**Propriété 4.1.1** *Nous donnons ici quelques propriétées concernant la fonction de répartition : Soit $X : \Omega \to J$ une variable aléatoire*

1. *Pour $x \leq \min_{y\in\Omega} X(y)$ alors $F_X(x) = P(X^{-1}(]-\infty; x[)) = P(\emptyset) = 0$ ;*
2. *Pour $x \geq \max_{y\in\Omega} X(y)$ alors $F_X(x) = P(X^{-1}(]-\infty; x[)) = P(\Omega) = 1$ ;*
3. *Une fonction de répartition est croissante ;*
4. *La probabilité pour que $X$ prenne ses valeurs dans un intervalle $[a, b[$ peut s'exprimer à l'aide de la fonction de répartition de $X$. En effet, l'événement contraire de $[a \leq X < b]$ est la réunion des deux événements disjoints :*

   $$[X \geq b], \quad \text{de probabilité} \quad 1 - F(b); \quad [X < a], \quad \text{de probabilité } F(a)$$

   *cette événement contraire a donc pour probabilité*

   $$P([X \geq b] \cup [X < a]) = 1 - F(b) + F(a),$$

   *et par suite*

   $$P([a \leq X < b]) = F(b) - F(a)$$

### 4.1.6 Quantile $q_\alpha$, Médiane et Mode

**Définition 4.1.6** *(***Quantile***)*
*Soit $X$ une variable aléatoire continue. Le nombre $q_\alpha$ tel que*

$$P([X \leq q_\alpha]) = \alpha \in [0, 1]$$

*est le* **quantile** *d'ordre $\alpha$ de la loi de probabilité de $X$.*

Ces quantile sont notés de différentes façons : par $u_\alpha$ pour la loi normale, par $t_\alpha^n$ pour la loi de Student à $n$ degrés de liberté et par $\chi_\alpha^n$ pour la loi du $\chi_n^2$.

**Définition 4.1.7** *(***Médiane***)*
*Soit $X$ une variable aléatoire continue. La valeur $q_{0.5}$ est la médiane de la variable aléatoire si*

$$P([X \leq q_{0.5}]) = 0.5.$$

*La médiane est le quantile $\alpha = 0.5$ de la variable aléatoire $X$.*

**Définition 4.1.8** *(***Mode***)*
*Soit $X$ une variable aléatoire continue de densité de probabilité $f$. Un point $x_0$ du domaine de définition $\mathcal{D}$ de $f$ tel que $f(x_0) = \max_{x\in\mathcal{D}} f(x)$ est appelé un* **mode** *de la variable aléatoire $X$.*

### 4.1.7 Systèmes de variables aléatoires réelles

**Théorème 4.1.3** *soit $(\Omega, \mathcal{B})$ un espace probabilisable et une base $\mathcal{A}$ de $\mathbb{R}^n$.*
*Une application $X$ de $\Omega$ dans $\mathbb{R}^n$ est une variable aléatoire vectorielle si, et seulement si, les $n$ applications coordonnées, $X_1, X_2, \ldots, X_n$ de $X$ dans $\mathcal{A}$ sont des variables aléatoires réelles.*

**Démonstration.** $\Rightarrow\rfloor$ Soit $S_n = (X_1, X_2, \ldots, X_n)$ une suite de $n$ variables aléatoires réelles.
Soit $\{e_1, e_2, \ldots, e_n\}$ une base de l'espace vectoriel $\mathbb{R}^n$. Posons

$$X = X_1.e_1 + X_2.e_2 + \ldots + X_n.e_n.$$

$X$ est une application de $\Omega$ dans $\mathbb{R}^n$. Est-ce une variable aléatoire vectorielle ?
Il faudrait pour cela que, quel que soit le vecteur

$$x = x_1.e_1 + x_2.e_2 + \ldots + x_n.e_n$$

de $\mathbb{R}^n$, $X^{-1}(\{x\})$ soit un événement. Or $X^{-1}(\{x\})$ est l'intersection des $n$ événements

$$[X_1 = x_1],\ [X_2 = x_2], \ldots, [X_n = x_n],$$

c'est donc un événement.
$\Leftarrow\rfloor$ soit $x$ une variable aléatoire vectorielle (à valeurs dans $\mathbb{R}^n$) et $\{e_1, e_2, \ldots, e_n\}$ une base de l'espace vectoriel $\mathbb{R}^n$.
Les coordonnées $(X_1, X_2, \ldots, X_n)$ de $X$, définies par

$$X = X_1.e_1 + X_2.e_2 + \ldots + X_n.e_n$$

sont des applications de $\Omega$ dans $\mathbb{R}$. Sont-ce des variables aléatoires réelles ?
Raisonnons pour $X_1$ ; il faudrait pour cela que, quel que soit le réel $x_1$, $X_1^{-1}(\{x_1\})$ soit un événement . Or $X_1^{-1}(\{x_1\})$ est la réunion des événements

$$X = (x_1, x_2, \ldots, x_n)$$

où $x_1, x_2, \ldots, x_n$ prennent toutes les valeurs possibles.
$X_1^{-1}(\{x_1\})$ est donc un événement et $X_1$ est bien une variable aléatoire. □

### 4.1.8 Couple indépendants de lois marginales données

La probabilité de l'intersection de deux événements ne peut se calculer à partir des probabilités de ces événements : les lois marginales d'un couple ne permettent pas de déterminer la loi conjointe.
Il en va différenement si l'on sait qu'il s'agit d'un couple indépendant. En effet, pour $(i, j) \in X(\Omega)$

$$P([X = (i, j)]) = P([X_1 = i])P([X_2 = j]).$$

**Exemple 4.1.1** *Déterminer la loi conjointe d'un couple de variables aléatoires réelles indépendantes, de même loi On prépare le tableau, avec ses marges et on écrit dans chaque case le produit des probabilités figurant dans les*

| $i$ | 0 | 1 | 2 |
|---|---|---|---|
| $P([X = i])$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |

*marges correspondantes.*

| $X_2 \backslash X_1$ | 0 | 1 | 2 | |
|---|---|---|---|---|
| 0 | $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ | $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$ | $\frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$ | $\rightarrow \ \frac{1}{2}$ |
| 1 | $\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$ | $\frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$ | $\frac{1}{3} \times \frac{1}{6} = \frac{1}{18}$ | $\rightarrow \ \frac{1}{3}$ |
| 2 | $\frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$ | $\frac{1}{3} \times \frac{1}{6} = \frac{1}{18}$ | $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$ | $\rightarrow \ \frac{1}{6}$ |
| | $\downarrow$ $\frac{1}{2}$ | $\downarrow$ $\frac{1}{3}$ | $\downarrow$ $\frac{1}{6}$ | |

### 4.1.9 Système de variables aléatoires réelles indépendantes

On peut généraliser la définition donnée pour un couple.

**Définition 4.1.9** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé. Un système de $n$ variables aléatoires réelles $X = (X_1, X_2, \ldots, X_n)$ tels que*

$$(\forall (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n \quad [X_1 = a_1],\ [X_2 = a_2],\ \ldots\ [X_n = a_n]$$

*sont des événements mutuellement indépendants est appelé système de $n$ variables aléatoires réelles indépendantes.*

Comme pour un couple, on peut se restreindre aux éléments $(a_1, a_2, \ldots, a_n)$ qui appartiennent à $X(\Omega)$.
De même que pour un couple, si $(X_1, X_2, \ldots, x_n)$ est un système indépendant, quelless que soient la parties $A_1, a_2, \ldots, \ldots, A_n$ de $\mathbb{R}$, les événements $X_1^{-1}(A_1), X_2^{-1}(A_2), \ldots, X_n^{-1}(A_n)$ sont indépendants.

## 4.2 Moments : espérance mathématique (moyenne) et variance d'une variable aléatoire réelle

**Définition 4.2.1** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé fini et $(A_1, A_2, \ldots, A_n)$ une suite formée de sous-ensembles de $\mathcal{B}$. On appelle espérance mathématique la forme linéair $E$ définie l'espace vectoriel des variables aléatoires réelles $V$ sur $(\Omega, \mathcal{B})$ par ses valeurs pour les indicatrices constituant la base : $(\varphi_{A_1}, \ldots, \varphi_{A_n})$ de la manière sivante*

$$E(\varphi_{A_i}) = P(A_i).$$

Si $X$ est une variable aléatoire réelle prenant la valeur $x_i$ sur $A_i$ ($i = 1, \ldots, n$).
On a

$$X = \sum_{i=1}^{n} x_i \varphi_{A_i}$$

et donc on a $E(X) = \sum_{i=1}^{n} x_i P(A_i)$. On peut établir que si $X(\Omega) = \{y_1, y_2, \ldots, y_r\}$ où les $y_i$ sont distincts alors l'expression de l'espérance mathématique $E$ devient

$$E(X) = \sum_{i=1}^{r} y_i P(X = y_i).$$

**Définition 4.2.2** *Une variable aléatoire réelle $X$ est dite centrée si son espérance mathématique $E(X)$ est nulle.*

**Exercice 4.2.1** *Montrer que la variable aléatoire $X - E(X)$ est centrée.*

**Exemple 4.2.1** *(Espérance d'une variable binomiale)*
*Une variable binomiale de paramètres $n$ et $p$ ($n \in \mathbb{N}^*$ et $0 < p < 1$) est une variable aléatoire telle que*

$$X(\Omega) = \{0, 1, 2, \ldots, n\} \quad et \quad P(X = i) = \mathrm{C}_n^i p^i (1-p)^{n-i}.$$

*D'après son expression, l'espérance mathématique est*

$$\begin{aligned}
E(X) &= \sum_{i=1}^{n} i \mathrm{C}_n^i p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} \frac{(n-1)!}{i!(n-i)!} p^{i-1} (1-p)^{n-i} \\
&= np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} \\
&= np(p + (1-p))^{n-1} \\
&= np
\end{aligned}$$

### 4.2.1 Espérance mathématique d'un produit

Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé fini. Un couple de variables aléatoires réelles $(X, Y)$ définies sur $(\Omega, \mathcal{B})$. Soit

$$\begin{aligned}
X(\Omega) &= \{x_1, x_2, \ldots, x_n\} \\
Y(\Omega) &= \{y_1, y_2, \ldots, y_p\}.
\end{aligned}$$

où les $x_i$ étant distincts et $y_i$ étant distincts.
On a $(X, Y)(\Omega) = \{(x_i, y_j)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}\}$ les $(x_i, y_j)$ étant distincts,

$$E(XY) = \sum_{i=1}^{n} \sum_{j=1}^{p} x_i y_j P((X, Y) = (x_i, y_j)).$$

Si on ne connaît pas les lois marginales de $X$ et $Y$, il n'est pas possible de poursuivre puisqu'on ne peut pas calculer

$$P((X, Y) = (x_i, y_j)).$$

♣ Si $X$ et $Y$ sont indépendantes, alors

$$P((X, Y) = (x_i, y_j)) = P(X = x_i) P(Y = y_j),$$

et on a

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

### 4.2.2 Moments d'ordre $k$ d'une variable aléatoire

Soit $(\Omega, \tau, P)$ un espace probabilisé et $X$ une variable aléatoire réelle.

**Définition 4.2.3** *On appelle* **moment** *d'ordre $k$ ($k \geq 0$) de $X$ et on note $M_k(X)$ l'espérance mathématique de la variable aléatoire $X^k$. On écrit*

$$M_k(X) = E[X^k].$$

**On s'intéresse plus volontiers au moment d'ordre deux par rapport à un réel $a$ :*

$$M_2(X - a) = E[(X - a)^2].$$

*En Physique, ce moment mesure la* **dispersion** *de la variable aléatoire $X$ par rapport au nombre $a$.*

**Théorème 4.2.1** *(***Changement d'origine***)*
*Soit $X$ une variable aléatoire réelle. De toutes les variables $X - a$ ($a \in \mathbb{R}$), c'est la variable centrée qui a le plus petit moment d'ordre deux.*

**Démonstration.** Considérons deux réels $a$ et $b$ et comparons les moments

$$M_2(X-a) \quad \text{et} \quad M_2(X-b).$$

On a

$$\begin{aligned}(X-a)^2 &= (X-b+(b-a))^2 \\ &= (X-b)^2 + 2(b-a)(X-b) + (b-a)^2\end{aligned}$$

d'où l'on déduit

$$E[(X-a)^2] = E[(X-b)^2] = 2(b-a)\,(E[X]-E[b]) + E[(b-a)^2];$$

or, l'espérance d'une constante est égale à cette constante. On a donc :

$$M_2(X-a) = M_2(X-b) + 2(b-a)(E[X]-b) + (b-a)^2$$

en particulier, si on choisit $b = E[X]$ :

$$M_2(X-a) = M_2(X-E[X]) + (E[X]-a)^2.$$

□

### 4.2.3 Variance et écart-type d'une variable aléatoire

**Définition 4.2.4** *On appelle* **variance** *(ou* **indicateur de dispersion***) d'une variable aléatoire réelle $X$, la quantité $\sigma^2(X) := Var(X)$ définie par*

$$Var(X) = \sigma^2(X) = E((X-E(X))^2).$$

**Propriété 4.2.1** *Soit $k$ un réel. Alors on a les propriétés suivantes :*
1. $Var(k.X) = k^2.Var(X)$,
2. $Var(X+k) = Var(X)$,
3. $Var(X) = E(X^2) - [E(X)]^2$.

**Définition 4.2.5** *On appelle l'***écart-type** *d'une variable aléatoire réelle $X$, la quantité $\sigma(X)$ définie par*

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

**Exemple 4.2.2** *Soit $X$ une variable aléatoire binomiale de paramètres $n$ et $p$ ($n \in \mathbb{N}^*$ et $0 < p < 1$). On a*

$$X(\Omega) = \{0,1,2,\ldots,n\} \quad \textit{et} \quad P(X=i) = \mathrm{C}_n^i p^i (1-p)^{n-i}.$$

*Calculons d'abord $E(X^2)$, ou mieux encore, pour un calcul plus facile calqué sur celui de l'espérance, $E(X(X-1))$ :*

$$\begin{aligned}E(X(X-1)) &= \sum_{i=0}^{n} i(i-1)\frac{n!}{i!(n-i)!}p^i(1-p)^{n-i} \\ &= np\sum_{i=0}^{n}\frac{n!}{(i-2)!(n-i)!}p^i(1-p)^{n-i} \\ &= n(n-1)p^2\sum_{i=2}^{n}\frac{(n-2)!}{(i-2)!(n-i)!}p^{i-2}(1-p)^{n-i} \\ &= n(n-1)p^2\end{aligned}$$

*et donc on a*

$$E(X^2) = E(X(X-1)) + E(X) = n(n-1)p^2 + np.$$

*En fin*

$$Var(X) = E(X^2) - [E(X)]^2 = n(n-1)p^2 + np - (np)^2 = np(1-p).$$

**Exemple 4.2.3** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé fini et $X : \Omega \to X(\Omega)$ la variable aléatoire de Bernoulli, c'est-à-dire que $X(\Omega) = \{0, 1\}$ et $P(X = 0) = q$, $P(X = 1) = p$ où $0 < q, p < 1 \quad p + q = 1$.*
*L'espérance de la variable aléatoire de Bernoulli est*

$$\begin{aligned}
E(X) &= \sum_{i=0}^{1} iP(X = i) = 0P(X = 0) + 1P(X = 1) = 0(1 - p) + 1p = p.\\
E(X^2) &= 0^2.p + 1^2.q = q = 1 - p.\\
E(\cos(\pi X)) &= \cos(0).p + \cos(\pi).q = p - q = p - (1 - p) = 2p - 1.\\
Var(X) &= E(X^2) - [E(X)]^2 = p - p^2 = p(1 - p).
\end{aligned}$$

### 4.2.4 Coefficient d'asymétrie et coefficient d'aplatissement

**Définition 4.2.6** *Soit $X$ une variable aléatoire d'espérance mathématique $\mu$ et de variance $\mathrm{Var}(X)$.*

1. *On appelle* **coefficient d'asymétrie** *(*skewness*) de la v.a. $X$, noté $\gamma_1(X)$, le coefficient donné par l'expression*

$$\gamma_1(X) = \frac{\mathbb{E}[(X - \mu)^3]}{(\mathrm{Var}(X))^3}.$$

2. *On appelle* **coefficient d'aplatissement** *(*kurtosis*) de la v.a. $X$ , noté $\gamma_2(X)$, le coefficient donné par l'expression*

$$\gamma_2(X) = \frac{\mathbb{E}[(X - \mu)^4]}{(\mathrm{Var}(X))^4}.$$

## 4.3 Loi et espérance conditionnelles

**Définition 4.3.1** *Soient $X$ et $Y$ deux variables aléatoires discrètes à valeurs respectives dans $F$ et $G$. Pour $y \in G$, on appelle* **loi conditionnelle** *de $X$ sachant $Y = y$ la famille des nombres*

$$(P_{[Y=y]}([X = x]))_{x \in X}$$

**Remarque 4.3.1** *On a*

$$P_{[Y=y]}([X = x]) = \begin{cases} \dfrac{P([X = x] \cap [Y = y])}{p([Y = y])}, \ \textit{si } P([Y = y]) > 0, \\ \\ P([X = x]), \ \textit{sinon}, \end{cases}$$

**Proposition 4.3.1** *Les variables $X$ et $Y$ sont indépendantes si, et seulement si la loi conditionnelle de $X$ sachant $Y = y$ ne dépend pas de $y \in G$. C'est-à-dire*

$$P_{[Y=y]}([X = x]) = P([X = x]).$$

## 4.4 Covariance, corrélation, vecteur-moyenne et matrice de covariance

**Définition 4.4.1** *Soit $(X, Y)$ deux variables aléatoires,*

1. *on appelle* **covariance** *entre $X$ et $Y$ la quantité définie par :*

$$\begin{aligned}
cov(X, Y) &= E((X - E(X))(Y^* - E(Y^*))),\\
&= E(XY^*) - E(X)E(Y^*).
\end{aligned}$$

2. *On dit que $X$ et $Y$ ne sont pas corrélées si cov$(X,Y)=0$ soit encore $E(XY^*)=E(X)E(Y^*)$.*
3. *On appelle* **coefficient de corrélation** *la quantité définie par :*

$$\rho(X,Y)=\frac{cov(X,Y)}{\sqrt{var(X)}\sqrt{var(Y)}}.$$

On a $-1\leq\rho(X,Y)\leq 1$. Soit $(X_1,\ldots,X_n)$ $n$ variables aléatoires de moyennes respectives $E(X_i)$.

1. On appelle **vecteur-moyenne** le vecteur de dimension $n$ dont les composantes sont les moyennes $E(X_i)$.
2. On appelle **matrice de covariance** la matrice $C$ de dimension $(n\times n)$ dont l'élément générateur est $\mathbf{C}_{ij}=$ cov$(X_i,X_j)$ pour $1\leq i\leq n$ et $1\leq j\leq n$.

### 4.4.1 Notation matricielle

Si on note :

$$\mathbf{X}=\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}$$

le vecteur aléatoire dont les $n$ composantes sont les variables aléatoires $X_k$, le vecteur-moyenne s'écrit :

$$E(\mathbf{X})=\begin{pmatrix}E(X_1)\\ \vdots\\ E(X_n)\end{pmatrix}$$

et la matrice de covariance :

$$\mathbf{C}=E((\mathbf{X}-E(\mathbf{X}))(\mathbf{X}-E(\mathbf{X}))^T).$$

Notons que les éléments diagonaux de la matrice de covariance représentent les variances respectives des $n$ variables aléatoires ; ils sont donc positifs.
Si les $n$ variables aléatoires ne sont pas corrélées, leur matrice de covariance est diagonale.

**Propriété 4.4.1** *Toute matrice de covariance est positive. Ce qui signifie que, quel que soit le vecteur $\mathrm{x}\in\mathbb{R}^n$ on a $\mathrm{x}^T\mathbf{C}\mathrm{x}\geq 0$.*

**Démonstration.** Considérons la v.a. $Y=\sum_{k=1}^N a_k(X_k-E(X_k))$. On a $E(|Y|^2)\geq 0$ car l'espérance mathématique d'une variable aléatoire positive est positif. Exprimons maintenant $E(|Y|^2)$, il vient :

$$\begin{aligned}E(|Y|^2) &= E\left(\sum_{k=1}^N a_k(X_k-E(X_k))\sum_{\ell=1}^N a_\ell^*(X_\ell-E(X_\ell))^*\right)\\ &= \sum_{k=1}^N\sum_{m=1}^N a_k E\left((X_k-E(X_k))(X_\ell-E(X_\ell))^*\right)a_m^*\\ &= \sum_{k=1}^N\sum_{m=1}^N a_k \mathrm{c}_{km}a_m^* = \mathrm{x}^T\mathbf{C}\mathrm{x}\end{aligned}$$

comme $E(|Y|^2)\geq 0$ alors $\mathrm{x}^T\mathbf{C}\mathrm{x}\geq 0$ pour tout $\mathrm{x}\in\mathbb{R}^n$. □

### 4.4.2 Suite blanche

**Définition 4.4.2** *Soit $X_1,\ldots,X_n$ $n$ variables aléatoires. On dit qu'elles forment une suite blanche si on a à la fois*

*i) var$(X_i)=\sigma^2=$ constante,*

*ii) et cov$(X_i,X_j)=0$ pour $i\neq j$.*

*Leur matrice de covaraince s'écrit donc :*

$$\mathbf{C}=\sigma^2 I_n$$

*où $I_n$ est la matrice identité de taille $n\times n$.*

# Chapitre 5

# Variable aléatoire ayant une densité

## 5.1 Densité d'une variable aléatoire et fonction de répartition

Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et $X : \Omega \to J$ une variable aléatoire.

**Définition 5.1.1** *Une fonction $f : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$ est la densité de la variable $X$ si et seulement si*

1. *la fonction $f$ est positive,*
2. $\int_{\mathcal{D}} f(t)dt = 1.$

**Définition 5.1.2** *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et $X : \Omega \to J$ une variable aléatoire réelle ayant une densité $f_X$. On appelle fonction de répartition de $X$ la fonction réelle de variable réelle définie par*

$$x \mapsto F_X(x) = P([X < x]) = \int_{-\infty}^{x} f_X(t)dt.$$

**Remarque 5.1.1** *La mesure de probabilité $dP$ et la mesure de Lebesgue $dt$ sont liées par la relation suivante : $dP(t) = f_X(t)dt$ où $f_X$ est la densité d'une variable aléatoire $X$.*

**Propriété 5.1.1** *Soit $X : \Omega \to J$ une variable aléatoire réelle ayant une densité de probabilité $f_X$, alors on a les propriétés suivantes*

1. *La densité de la variable aléatoire $X$ est donnée par :*

$$f_X(x) = \lim_{h \to 0} \frac{F_X(x+h) - F_X(x)}{h} = F'_X(x).$$

2. *Si la variable $X$ a une densité $f_X$ alors la fonction de répartition est*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

*En général, on a*

$$P(a \leq X \leq b) = \int_{a}^{b} f_X(t)dt = F_X(b) - F_X(a).$$

## 5.2 Moments d'ordre $k$ d'une variable aléatoire ayant une densité

Soit $(\Omega, \tau, P)$ un espace probabilisé et $X$ une variable aléatoire réelle ayant une densité $f_X$ de domaine de définition $\mathcal{D}$.

**Définition 5.2.1** *On appelle* **moment** *d'ordre $k$ ($k \geq 0$) de la variable aléatoire $X$, noté $M_k(X)$, l'intégrale définie par*

$$M_k(X) = \int_{\mathcal{D}} t^k f_X(t)dt.$$

Par la suite, on s'intéresse plus volontiers aux moments d'ordres 1 et 2 d'une variable aléatoire par rapport à un réel $a$ :

$$M_2(X-a) = \int_{\mathcal{D}} (t-a)^2 f_X(t)dt.$$

Et, plus particulièrement lorsque le réel $a$ est le centre de gravité de la variable aléatoire $X$.

## 5.3 Espérance mathématique et variance d'une variable aléatoire ayant densité

**Définition 5.3.1** *Soit $X$ une variable $X$ à densité de probabilité $f_X$ de domaine de définition $\mathcal{D}$, on a*

1. *l'espérance mathématique ou la moyenne de la variable aléatoire $X$, notée $\mathbb{E}(X)$, est le moment d'ordre 1 de $X$ par rapport à 0 donné par*
$$\mathbb{E}(X) = M_1(X) = \int_{\mathcal{D}} t f_X(t)dt.$$
2. *la variance de la variable aléatoire $X$, notée $\sigma^2(X) :=$ Var$(X)$, est le moment d'ordre 2 de $X$ par rapport à sa moyenne $\mathbb{E}(X)$ donnée par*
$$\sigma^2(X) = M_2(X - \mathbb{E}(X)) = \int_{\mathcal{D}} (t - \mathbb{E}(X))^2 f_X(t)dt.$$

**Propriété 5.3.1** *Soit $X$ une variable $X$ à densité de probabilité $f_X$ de domaine de définition $\mathcal{D}$, alors on a*

$$\sigma^2(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2,$$

*où* $\mathbb{E}(X^2) = \displaystyle\int_{\mathcal{D}} t^2 f_X(t)dt.$

**Démonstration.** Pour tout $t \in \mathcal{D}$, on a

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

alors $(t - \mathbb{E}(X))^2 f_X(t) = t^2 - 2\mathbb{E}(X)t + (\mathbb{E}(X))^2 f_X(t)$,
Par passage à l'intégrale, on obtient

$$\begin{aligned}
\int_{\mathcal{D}} (t - \mathbb{E}(X))^2 f_X(t)dt &= \int_{\mathcal{D}} (t^2 f_X(t) - 2\mathbb{E}(X) t f_X(t) + (\mathbb{E}(X))^2 f_X(t))dt \\
&= \int_{\mathcal{D}} t^2 f_X(t)dt - 2\mathbb{E}(X) \int_{\mathcal{D}} t f_X(t)dt + (\mathbb{E}(X))^2 \int_{\mathcal{D}} f_X(t)dt \\
&= \int_{\mathcal{D}} t^2 f_X(t)dt - 2\mathbb{E}(X)\mathbb{E}(X) + (\mathbb{E}(X))^2
\end{aligned}$$

car $\mathbb{E}(X) = \int_{\mathcal{D}} t f_X(t)dt$ et $1 = \int_{\mathcal{D}} f_X(t)dt$.
Finallement, on obtient

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

□

**Exemple 5.3.1** *On considère la variable aléatoire $X$ de densité de probabilité $f_X$ donnée par*

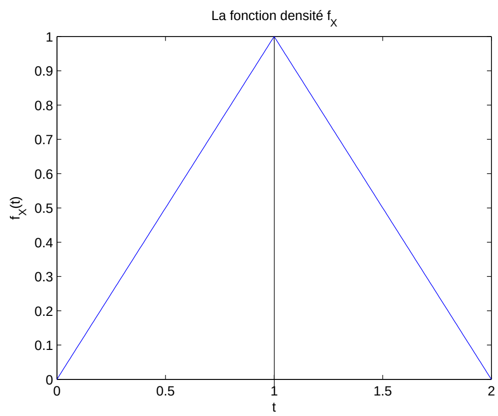
$$f_X(t)\begin{cases} t & si \quad t \in [0,1], \\ -t+2 & si \quad t \in ]1,2], \end{cases}$$

$$\mathbb{E}(X) = \int_0^2 t f_X(t)dt = \int_0^1 t^2 dt + \int_1^2 t(2-t)dt = \frac{1}{3} + \left(3 - \frac{7}{3}\right) = 1.$$

$$\mathbb{E}(X^2) = \int_0^2 t^2 f_X(t)dt = \int_0^1 t^3 dt + \int_1^2 t^2(2-t)dt = \frac{1}{4} + \left(\frac{14}{3} - \frac{15}{4}\right) = \frac{7}{6}.$$

$$Var(X) = \int_{-\infty}^{+\infty} t^2 f_X(t)dt - (\mathbb{E}(X))^2 = \frac{7}{6} - 1^2 = \frac{1}{6} = 0.1667.$$

*L' écart-type de la variable aléatoire réelle $X$ dont la densité $f_X$ est donnée par*




$$\sigma(X) = \sqrt{Var(X)} = \sqrt{\frac{1}{6}} = 0.4082.$$

## 5.4 Coefficient d'asymétrie et coefficient d'aplatissement

**Définition 5.4.1** *Soit $X$ une variable aléatoire réelle ayant une densité d'espérance $\xi = \mathbb{E}[X]$ et de variance $\zeta = \mathrm{Var}(X)$.*

1. *On appelle* **coefficient d'asymétrie** *(*skewness*) de la variable aléatoire $X$, noté $\gamma_1(X)$, le coefficient donné par l'expression*
$$\gamma_1(X) = \frac{1}{\zeta^3} \int_{\mathcal{D}} (t-\xi)^3 f_X(t)dt.$$
2. *On appelle* **coefficient d'aplatissement** *(*kurtosis*) de la variable aléatoire $X$ , noté $\gamma_2(X)$, le coefficient donné par l'expression*
$$\gamma_2(X) = \frac{1}{\zeta^4} \int_{\mathcal{D}} (t-\xi)^4 f_X(t)dt.$$

# Chapitre 6

# Lois de probabilité usuelles : discrètes et continues

## 6.1 Loi conjointe et lois marginales de variables aléatoires réelles

**Définition 6.1.1** 1. *Nous choisirons systématiquement comme base $\mathcal{A}$ la base canonique de $\mathbb{R}^n$ et nous écrirons*

$$X = (X_1, X_2, \dots, X_n)$$

2. *La loi de probabilité de $X$ est appelée la* **loi conjointe du système** $(X_1, X_2, \dots, X_n)$. *Les lois des variables aléatoires $X_i$ sont appelées les* **lois marginales du système**.

### 6.1.1 Exemple d'un couple de variables aléatoires réelles

Le cas d'un système de deux variables aléatoires réelles est particulièrement simple. Illustrons-le par un exemple. Soit $\Omega = \{a, b, c, d, e\}$ et soit $\mathcal{B}$ une tribu de parties de $\Omega$ dont les atomes sont $\{a, b\}$, $\{c\}$, $\{d\}$, $\{e\}$. On donne deux variables aléatoires réelles $X_1$ et $X_2$ par le tableau suivant :

| $\omega$ | $a$ | $b$ | $c$ | $d$ | $e$ |
|---|---|---|---|---|---|
| $X_1(\omega)$ | 0 | 0 | 1 | 1 | 2 |
| $X_2(\omega)$ | 0 | 0 | 1 | 2 | 1 |

et on considère le couple de variables aléatoires réelles $X = (X_1, X_2)$.
Donnons-nous une probabilité $P$ en posant

$$P(\{a, b\}) = \frac{1}{2}, \;\; P(\{c\}) = P(\{d\}) = P(\{e\}) = \frac{1}{6}.$$

Il est commode de présenter $X(\Omega)$ en un tableau dont chaque case représente un élément de $X(\Omega)$.

| $X_2 \backslash X_1$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | | | |
| 1 | | | |
| 2 | | | |

**Ainsi pourra-t-on présenter la loi conjointe en indiquant dans la case $(i, j)$ la probabilité**

$$P([X = (i, j)]).$$

Déterminons cette loi

D'où la loi conjointe, présentée en tableau

| Valeurs de $X$ $(i,j)$ | éventualités formant $X=(i,j)$ | Probabilité de $X=\{i,j\}$ |
|---|---|---|
| $(0,0)$ | $a,b$ | $\frac{1}{2}$ |
| $(0,1)$ | $\emptyset$ | $0$ |
| $(0,2)$ | $\emptyset$ | $0$ |
| $(1,0)$ | $\emptyset$ | $0$ |
| $(1,1)$ | $c$ | $\frac{1}{2}$ |
| $(1,2)$ | $d$ | $\frac{1}{2}$ |
| $(2,0)$ | $\emptyset$ | $0$ |
| $(2,1)$ | $e$ | $\frac{1}{6}$ |
| $(2,2)$ | $\emptyset$ | $0$ |

| $X_2$ \ $X_1$ | 0 | 1 | 2 | |
|---|---|---|---|---|
| 0 | $\frac{1}{2}$ | 0 | 0 | $\rightarrow$ total $\frac{1}{2}$ |
| 1 | 0 | $\frac{1}{6}$ | $\frac{1}{6}$ | $\rightarrow$ total $\frac{1}{3}$ |
| 2 | 0 | $\frac{1}{6}$ | 0 | $\rightarrow$ total $\frac{1}{6}$ |
| | $\downarrow$ total $\frac{1}{2}$ | $\downarrow$ total $\frac{1}{3}$ | $\downarrow$ total $\frac{1}{6}$ | |

### 6.1.2 Déterminons maintenant les lois marginales

$$\begin{aligned}
P([X_1=0]) &= P(\{a,b\}) = \frac{1}{2} \\
P([X_1=1]) &= P(\{c,d\}) = \frac{1}{3} \\
P([X_1=2]) &= P(\{e\}) = \frac{1}{6} \\
P([X_2=0]) &= P(\{a,b\}) = \frac{1}{2} \\
P([X_2=1]) &= P(\{c,e\}) = \frac{1}{3} \\
P([X_2=2]) &= P(\{d\}) = \frac{1}{6}
\end{aligned}$$

On peut remarquer que ces probabilités se trouvent sur le tableau, si on effectue, **dans les marges**, les sommes des probabilités inscrites dans une même ligne ou colonne (d'où le nom **loi marginale**).
Quoi d'étonnant à cela puisque, par exemple

$$[X_1=1]=[X=(1,0)]\cup[X=(1,1)]\cup[X=(1,2)]$$

et que les trois événements ainsi réunis sont deux à deux disjoints.

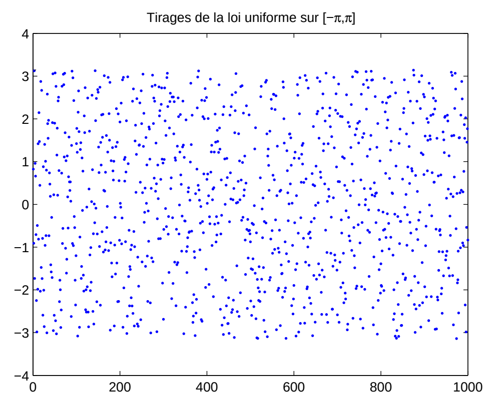
## 6.2 Loi uniforme continue sur $(a,b)$

**Définition 6.2.1** *Une variable aléatoire $X$ est dite* **uniforme** *ou* **équirépartie** *sur $(a,b)$ avec $b>a$, si sa densité de probabilité a pour expression :*

$$f_X(t)=\begin{cases} \frac{1}{b-a}, & \textit{si} \quad t\in(a,b), \\ 0, & \textit{sinon}. \end{cases}$$

On peut calculer l'espérance mathématique ou la moyenne de cette variable aléatoire, sa variance et son écart-type :

1. L'espérance mathématique :
$$E(X) = \int_a^b t f_X(t)dt = \int_a^b \frac{t}{b-a}dt = \frac{a+b}{2}.$$
2. La variance :
$$E(X^2) = \int_a^b t^2 f_X(t)dt = \int_a^b \frac{t^2}{b-a}dt = \frac{a^2+ab+b^2}{3}.$$
$$\text{var}(X) = E(X^2) - (E(X))^2 = \frac{(b-a)^2}{12}.$$
3. L'écart-type :
$$\sigma(X) = \sqrt{\text{Var}(X)} = \sqrt{\frac{(b-a)^2}{12}} = \frac{b-a}{2\sqrt{3}}.$$
4. Distribution d'un nuage de points par une loi uniforme :




## 6.3 Loi de Poisson

**Définition 6.3.1** *Soit $(\Omega, \tau, P)$ un espece probabilisé.*
*Une variable aléatoire $X$ a une loi de Poisson de paramètre $\lambda$ réel positif si les valeurs de cette variable sont les entiers positifs et si les probabilités associées sont :*
$$P(k;\lambda) = P([X=k]) = e^{-\lambda}\frac{\lambda^k}{k!}$$
*C'est-à-dire que*

*i)* $X : \Omega \longrightarrow \mathbb{N}$

*ii) Pour tout $k \geq 0$, la probabilité de l'événement $[X=k]$ est*
$$P(k;\lambda) = P([X=k]) = e^{-\lambda}\frac{\lambda^k}{k!}$$

**Propriété 6.3.1** *Soit $(\Omega, \tau, P)$ un e.p. et $X$ une variable aléatoire ayant la loi de Poisson de paramètre $\lambda$ réel positif. Alors on a*

1. $\mu = E[X] = \lambda$,
2. *Var*$(X) = \sigma^2(X) = \lambda$.

## 6.4 Loi normale $\mathcal{N}(\mu, \sigma^2)$ ou v.a. gaussienne

**Définition 6.4.1** *1. Une variable aléatoire $X$ a une loi normale réduite si elle a une densité $\varphi$ définie par :*

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

*On dit que c'est la loi normale $\mathcal{N}(0, 1)$. La fonction de répartition $\Pi$ associée à la variable $X$ est définie par*

$$\Pi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt,$$

*où $E[X] = 0$ et $\sigma(X) = 1$ est son écart-type.*

*2. Une variable aléatoire $Y$ a une loi normale non-réduite si elle a une densité $\psi$ définie par :*

$$\psi(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(t-\mu)^2/2\sigma^2}.$$

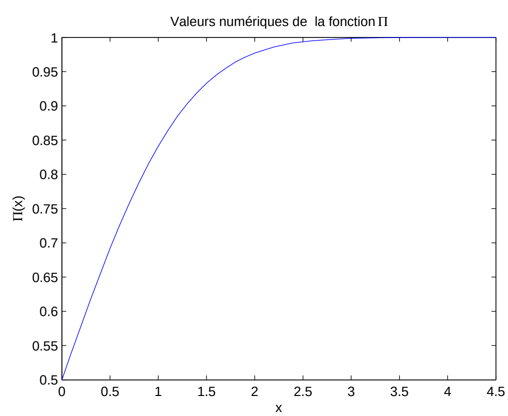
*On dit que c'est la loi normale $\mathcal{N}(\mu, \sigma^2)$. La fonction de répartition $\Pi$ associée à la variable $X$ est définie par*

$$\Pi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-(t-\mu)^2/2\sigma^2} dt,$$

*où $\mu = E[Y]$ et $\sigma$ est son écart-type.*

## 6.5 Valeurs numériques de $\Pi(x)$

| $x$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\Pi(x)$ | 0.5 | 0.54 | 0.579 | 0.618 | 0.655 | 0.692 | 0.726 | 0.758 | 0.788 | 0.816 |
| $x$ | 1.0 | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 1.6 | 1.7 | 1.8 | 1.9 |
| $\Pi(x)$ | 0.841 | 0.864 | 0.885 | 0.903 | 0.919 | 0.933 | 0.945 | 0.955 | 0.964 | 0.971 |
| $x$ | 2.0 | 2.2 | 2.4 | 2.6 | 2.8 | 3.0 | 3.3 | 3.6 | 4.0 | 4.5 |
| $\Pi(x)$ | 0.977 | 0.986 | 0.992 | 0.995 | 0.997 | 0.9986 | 0.9995 | 0.9998 | 0.99996 | 0.999997 |




**Exercice 6.5.1** *1. Montrer que $\Pi(x) + \Pi(-x) = 1$.*

*2. Soit $Y$ la loi normale $\mathcal{N}(\mu, \sigma^2)$. Montrer que $X = \frac{(Y-\mu)}{\sigma}$ est une variable aléatoire réduite.*

*3. Montrer que la fonction de densité $f$ et la fonction de répartition $F$ associées sont données par :*

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2}, \quad F(y) = \Pi\left((y-\mu)/\sigma\right).$$

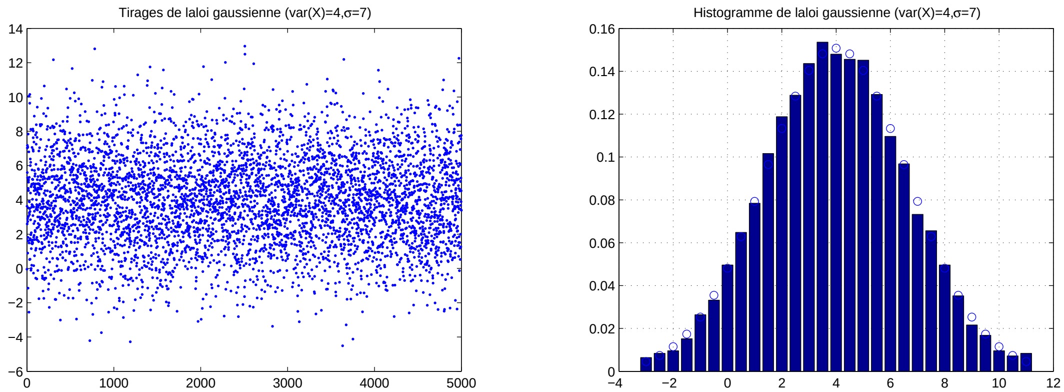



FIGURE 6.1 – Ici $\mu = \mathrm{E}(X)$ et $\sigma = \sqrt{\sigma^2}$.

## 6.6 Loi Gamma et loi Khi-deux : $\chi^2$

### 6.6.1 Fonction Gamma $\Gamma$ :

Soit $z \in \mathbb{C}$ un nombre complexe, on définit pour tout réel positif $x$, la fonction puissance $x \mapsto x^z$ par

$$x^z = e^{z \ln(x)}$$

**Lemme 6.6.1** *Soit $z$ un nombre complexe. La fonction $t \mapsto t^{z-1}e^{-t}$ est intégrable sur $]0, +\infty[$ si et seulement si* $\mathrm{Re}(z) > 0$.

**Démonstration.** La fonction $t \mapsto t^{z-1}e^{-t}$ est continue sur $]0, +\infty[$ pour nombre complexe $z$. Le lemme se démontre à l'aide des théorèmes de comparaisons du calcul intégral et des formules suivantes

$$|t^{z-1}e - t| = t^{\mathrm{Re}(z)-1}e^{-t} \sim_{t \to 0^+} t^{\mathrm{Re}(z)-1} \text{ et } |t^{z-1}e - t| = t^{\mathrm{Re}(z)-1}e^{-t} =_{t \to +\infty} o(e^{-t/2}).$$

□

**Définition 6.6.1** *On appelle la fonction* **Gamma**, *notée* $\Gamma$, *la fonction*
$\Gamma : \{z \in \mathbb{C} \;/\; \mathrm{Re}(z) > 0\} \to \mathbb{C}$ *définie par l'expression intégrale suivante :*

$$\Gamma(z) = \int_0^{+\infty} t^{z-1}e^{-t}dt.$$

*D'une façon analogue, on définit la fonction* $\Gamma_a$ *par l'expression suivante :*

$$\Gamma_\alpha(z) = \int_0^{\alpha} t^{z-1}e^{-t}dt.$$

**Exemple 6.6.1** *Pour $\alpha > 0$ on a*

$$\Gamma_a\alpha(1) = \int_0^{\alpha} t^{1-1}e^{-t}dt = \left[-e^{-t}\right]_0^{\alpha} = 1 - e^{-\alpha} \to 1$$

*lorsque $\alpha \to +\infty$. D'où* $\Gamma(1) = 1$.

**Proposition 6.6.1** *on a* $\Gamma(z+1) = z\Gamma(z)$.

**Démonstration.** On a $\Gamma_\alpha(z+1) = \int_0^\alpha t^z e^{-t} dt$. Par une intégration par partie, on a

$$\Gamma_\alpha(z+1) = \int_0^\alpha t^{z-1} e^{-t} dt = -\alpha^z e^{-\alpha} + z\Gamma_\alpha(z).$$

Ce qui implique

$$|\Gamma_\alpha(z+1) - z\Gamma_\alpha(z)| \leq \alpha^{\mathrm{Re}(z)} e^{-\alpha} \to 0$$

lorsque $\alpha \to +\infty$. D'où

$$\Gamma(z+1) = z\Gamma(z)$$

$\square$

**Corollaire 6.6.1** *Pour $n \in \mathbb{N}^*$, on a $\Gamma(n+1) = n!$.*

**Démonstration.** Pour $n \in \mathbb{N}^*$, on a $\Gamma(n+1) = n\Gamma(n)$, $\Gamma(n) = (n-1)\Gamma(n-1)$, $\ldots\ldots$ et par recurrence on obtient $\Gamma(2) = 1$. D'où

$$\Gamma(n) = n \times (n-1) \ldots 2 \times 1 = n!$$

$\square$

## 6.7 Loi Gamma

**Proposition 6.7.1** *Soit $a$ et $\theta$ des réels strictement positifs, alors la fonction réelle $f$ de variable réelle $x$ définie par*

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(a)} \theta^a x^{a-1} e^{-\theta x}, & \textit{si} \quad x > 0, \\ 0, & \qquad \textit{sinon}. \end{cases}$$

*est une densité de probabilité d'une variable aléatoire $X$.*

**Démonstration.** Pour tout $a$ et $\theta > 0$, pour $\xi = \theta x$ on a

$$\int_{-\infty}^{+\infty} \theta^a x^{a-1} e^{-\theta x} 1_{\{x>0\}} dx = \int_0^{+\infty} \theta^a x^{a-1} e^{-\theta x} dx = \int_0^{+\infty} y^{a-1} e^{-y} dy = \Gamma(a),$$

d'où le résultat. $\square$

**Définition 6.7.1** *On dit qu'une variable aléatoire $X$ suit la* **loi Gamma de paramètres** *$a$ et $\theta$ où $a$ et $\theta > 0$ et on note $X \sim \Gamma(a, \theta)$ si $X$ possède la densité de probabilité*

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(a)} \theta^a x^{a-1} e^{-\theta x}, & \textit{si} \quad x > 0, \\ 0, & \qquad \textit{sinon}. \end{cases}$$

*En particulier, pour $a = 1$ alors la variable aléatoire $X \sim \Gamma(1, \theta)$ n'est que la loi exponentielle de paramètre $\theta$.*

## 6.8 Loi Beta

**Définition 6.8.1** *Soient $z$ et $\omega$ deux variables complexes telles que $\mathrm{Re}(z) > 0$ et $\mathrm{Re}(w) > 0$. On appelle fonction* **bêta***, la fonction notée $\beta$, définie par*

$$\beta(z,w) = \int_0^1 t^{z-1}(1-t)^{w-1}dt$$

**Proposition 6.8.1** *Pour $p > 0$ et $q > 0$, on a $\beta(p,q) = \dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.*

**Démonstration.** On note $V = \{(x,y) \in \mathbb{R}^2 \ / \ x > 0,\ y > 0\}$. On a

$$\Gamma(p)\Gamma(q) = \int_V x^{p-1}y^{q-1}e^{-x-y}dxdy$$

Soit $(\rho, t) \in ]0,+\infty[\times]0,1[= U$ et soit $\Phi : U \to V$, $\Phi(\rho,t) = (x,y)$ où $x = \rho t$ et $y = \rho(1-t)$ ,
alors $U$ est la boule ouverte. Le jacobien de $\Phi$ est : $\mathrm{Jac}(\Phi) = -\rho \neq 0$.
Ce qui montre que $\Phi$ est un diféomorphisme de $U$ dans $V$. Donc,

$$\begin{aligned}
\int_V x^{p-1}y^{q-1}e^{-x-y}dxdy &= \int_U \rho^{p+q-2}t^{p-1}(1-t)^{q-1}e^{-\rho} \times |-\rho| d\rho dt \\
&= \int_0^1 t^{p-1}(1-t)^{q-1}dt \int_0^{+\infty} \rho^{p+q-1}e^{-\rho}d\rho
\end{aligned}$$

D'où

$$\Gamma(p)\Gamma(q) = \Gamma(p+q)\beta(p,q), \quad p > 0, \ q > 0$$

□

**Proposition 6.8.2** *Soient $p$ et $q$ des réels strictement positifs, alors la fonction réelle $f$ de variable réelle $x$ définie par*

$$f(x) = \begin{cases} \dfrac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1}, & \text{si} \quad 0 \leq x \leq 1, \\ 0, & \text{sinon}. \end{cases}$$

*est une densité de probabilité d'une variable aléatoire $X$.*

**Démonstration.** Pour tout $a$ et $\theta > 0$, pour $\xi = \theta x$ on a

$$\int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\beta(p,q)} \int_0^1 x^{p-1}(1-x)^{q-1}dx = \frac{1}{\beta(p,q)}\beta(p,q) = 1,$$

d'où le résultat. □

**Définition 6.8.2** *On dit qu'une variable aléatoire $X$ suit la* **loi Beta** *de paramètre $p >$ et $q > 0$ et on note $X \sim \beta(p,q)$ si $X$ possède la densité de probabilité*

$$f(x) = \begin{cases} \dfrac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1}, & \text{si} \quad 0 \leq x \leq 1, \\ 0, & \text{sinon}. \end{cases}$$

## 6.9 Loi du Khi-2 où $\chi^2$

**Définition 6.9.1** *Soient $X_1, \dots, X_n$ une suite de $n$ variables aléatoires normales centrées réduites indépendantes. On appelle loi du* **Khi-2** *à $n$ degrés de liberté, notée $\chi^2(n)$ la loi de la variables aléatoire $X = \sum_{i=1}^{n} X_i^2$.*

**Proposition 6.9.1** *Soit $X$ une variable aléatoire qui suit la loi du $\chi^2(n)$ à $n$ degrés de libertés. Alors*

1. *la loi de la variable aléatoire $\chi^2(n)$ est la loi $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ de densité $f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} 1_{]0,+\infty[}(x)$.*
2. *On a $\mathbb{E}[\chi^2(n)] = n$ et $\sigma^2(\chi^2(n)) = 2n$.*

## 6.10 Lois de Student et de Fisher-Snedecor

**Définition 6.10.1** *Soient $X$ une variable aléatoire normale centrée réduite et $Y$ une variable aléatoire de loi $\chi^2(n)$ indépendante de $X$. On appelle loi de* **Student** *à $n$ degrés de libertés, notée $t(n)$, la loi de la variable aléatoire $T_n = \frac{X}{\sqrt{\frac{Y}{n}}}$.*

**Proposition 6.10.1**

1. *la loi de la variable aléatoire $t(n)$ est la loi de densité $f(t) = \frac{\Gamma\left(\frac{(n+1)}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \frac{1}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}$.*
2. *L'espérance et la variance d'une variable aléatoire de Student sont*

$$\mathbb{E}[T_n] = 0 \quad \textit{pour tout } \; n > 1 \quad \textit{et} \quad \sigma^2(T_n) = \frac{n}{n-2} \quad \textit{pour tout } \; n > 2.$$

**Définition 6.10.2** *Soient $X$ et $Y$ deux variables aléatoires indépendantes de lois $\chi^2_n$ et $\chi^2_p$. On appelle loi de* **Fisher** *de paramétres $n$ et $p$, notée $F_{n,p}$, la loi de la variable $F = \frac{\frac{X}{n}}{\frac{Y}{p}}$.*

*L'espérance et la variance d'une variable aléatoire de loi $F_{n,p}$ sont*

$$E(F) = \frac{p}{p-2} \quad \textit{pour tout } \; p > 2,$$

$$\sigma^2(F) = \frac{2p^2(n+p-2)}{n(p-2)^2(p-4)} \quad \textit{pour tout } \; p > 4.$$

## 6.11 Quelques lois de probabilités

| Loi | Type | **Probabilité ou Densité de probabilité(ddp)** | **Moyenne** | **Variance** |
|---|---|---|---|---|
| Bernoulli | D | $P([X=0]) = 1-p$ et $P([X=1]) = p$ | $p$ | $p(1-p)$ |
| Uniforme | D | $P([X=i]) = \frac{1}{n}$ pour $1 \leq i \leq n$ | $\frac{n+1}{2}$ | $\frac{n^2-1}{12}$ |
| Binomiale | D | $P([X=i]) = C_n^i p^i (1-p)^{n-i}$ pour $1 \leq i \leq n$ | $np$ | $np(1-p)$ |
| Géométrique | D | $P([X=i]) = p(1-p)^{i-1}$ pour $i = 1, 2, \ldots$ | $\frac{1}{p}$ | $\frac{1-p}{p^2}$ |
| Pascal | D | $P([X=i]) = C_{i-1}^{n-1} p^n (1-p)^{i-n}$ | $\frac{n}{p}$ | $\frac{n(1-p)}{p^2}$ |
| Poisson | D | $P([X=i]) = \frac{\lambda^i}{i!} e^{-\lambda}$ pour $\lambda > 0$ et $i = 1, 2, \ldots$ | $\lambda$ | $\lambda$ |
| Uniforme | C | $f(x) = \frac{1}{b-a} \mathbb{I}_{[a,b]}(x)$ | $\frac{b+a}{2}$ | $\frac{(b-a)^2}{12}$ |
| Gauss | C | $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ pour $x \in \mathbb{R}$ | $\mu$ | $\sigma^2$ |
| Cauchy | C | $f(x) = \frac{a}{\pi(a^2+x^2)}$ pour $x \in \mathbb{R}$ | non défini | non défini |
| Gamma | C | $f(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}$ pour $x > 0$ | $\frac{k}{\lambda}$ | $\frac{k}{\lambda^2}$ |
| Exponentielle | C | $f(x) = \frac{1}{a} x^{k-1} e^{-\frac{x}{a}}$ pour $x > 0$ et $a > 0$ | $a$ | $a^2$ |
| Rayleigh | C | $f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$ pour $x > 0$ | $\sigma\sqrt{\frac{\pi}{2}}$ | $\sigma^2(1-\frac{\pi}{2})$ |
| Laplace | C | $f(x) = \frac{a}{2} e^{-a\|x\|}$ | $0$ | $\frac{2}{a^2}$ |
| $\chi^2$ | C | $f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \mathbb{I}_{]0,+\infty[}(x)$ | $n$ | $2n$ |
| Student | C | $f(x) = \frac{\frac{n+1}{2}}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)\left(1+\frac{x^2}{n}\right)^{\frac{n+1}{2}}}$ | $0$ | $\frac{n}{n-2}$ pour $n > 2$ |
| Weibull | C | $f(x) = \beta x^{\beta-1} e^{-x^\beta}$ | $\Gamma(1+\frac{1}{\beta})$ | $\Gamma(1+\frac{2}{\beta}) - (E(X))^2$ |

TABLE 6.1 – Type : $D \equiv$ loi discrète ; $C \equiv$ loi continue