Calcul de probabilités et Statistique descriptive

Polycopié de cours et exercices de travaux dirigés

Issa Cherif GERALDO

Département de Mathématiques, Faculté des Sciences, Université de Lomé

Octobre 2023

TABLE DES MATIÈRES

Première partie

Calcul de probabilités

CHAPITRE 1

ENSEMBLES FINIS ET ANALYSE COMBINATOIRE

Résumé – Ce chapitre rappelle quelques propriétés des ensembles finis ainsi que les formules classiques d'analyse combinatoire généralement présentées en classe de Terminale et qui seront très utiles dans la suite du cours. L'objectif est de savoir résoudre les problèmes classiques d'analyse combinatoire en utilisant la formule adaptée.

Sommaire

1.1 ENSEMBLES FINIS

1.1.1 Généralités

Définition 1.1

- Un ensemble Ω est une collection d'objets appelés *éléments* de l'ensemble. On écrit $x \in \Omega$ si x est élément de Ω et $x \notin \Omega$ sinon.
- On admet l'existence de l'*ensemble vide* noté $\emptyset$ et qui ne contient aucun élément.

Définition 1.2

- On dit qu'une ensemble Ω est *fini* s'il est vide ou s'il peut être mis en relation avec un ensemble de numéros de la forme $\{1, \ldots, n\}$ $(n \in \mathbb{N}^*)$ de sorte qu'à chaque élément de E corresponde un unique numéro et qu'à chaque numéro corresponde un unique élément de Ω.

- Dans ce cas, le nombre n est appelé le *cardinal* de Ω noté $\text{Card}(\Omega)$ avec la convention $\text{Card}(\emptyset) = 0$.
- Un ensemble non fini est dit *infini*.

Exemple 1.3 L'ensemble des lettres minuscules $\Omega = \{a, b, c, d\}$ est fini de cardinal 4 alors que l'ensemble $\mathbb{N}$ des entiers naturels et l'ensemble $\mathbb{R}$ des nombres réels sont des ensembles infinis.

Définition 1.4 Soit Ω un ensemble. Un ensemble A est un *sous-ensemble* (ou une *partie*) de Ω si tout élément de A est nécessairement élément de Ω. On note $A \subset \Omega$. L'ensemble des parties de Ω est noté $\mathscr{P}(\Omega)$. En d'autres termes,

$$A \in \mathscr{P}(\Omega) \text{ si et seulement si } A \subset \Omega.$$

Par convention, l'ensemble vide $\emptyset$ est inclus dans tous les ensembles.

Exemple 1.5 Si $\Omega = \{a, b\}$ alors $\mathscr{P}(\Omega) = \Big\{\emptyset, \{a\}, \{b\}, \{a, b\}\Big\}$.

1.1.2 Opérations sur les ensembles finis et calcul de cardinaux

Définition 1.6 Soient Ω un ensemble et A, B deux parties de Ω.

- L'*intersection* de A et B, notée $A \cap B$, est l'ensemble des éléments appartenant à A et à B.
- Les ensembles A et B sont dits *disjoints* s'ils n'ont aucun élément en commun c'est-à-dire si $A \cap B = \emptyset$.
- La *réunion* (ou l'*union*) de A et B, notée $A \cup B$, est l'ensemble des éléments appartenant à A ou à B ou en d'autres termes, l'ensemble des éléments appartenant à l'un au moins des deux ensembles.
- Le *complémentaire* de A dans Ω, noté $\overline{A}$, est l'ensemble des éléments de Ω n'appartenant pas à A.
- La *différence* de A et B, notée $A \setminus B$, est l'ensemble des éléments appartenant à A et n'appartenant pas à B.
$$A \setminus B = A \cap \overline{B}.$$
- La *différence symétrique* de A et B, notée $A\Delta B$, est l'ensemble des éléments appartenant à une et une seule des deux parties A et B (c'est-à-dire que $A\Delta B$ contient les éléments de A n'appartenant pas à B et les éléments de B n'appartenant pas à A).
$$A\Delta B = (A \cup B) \setminus (A \cap B) = (A \cap \overline{B}) \cup (B \cap \overline{A}).$$

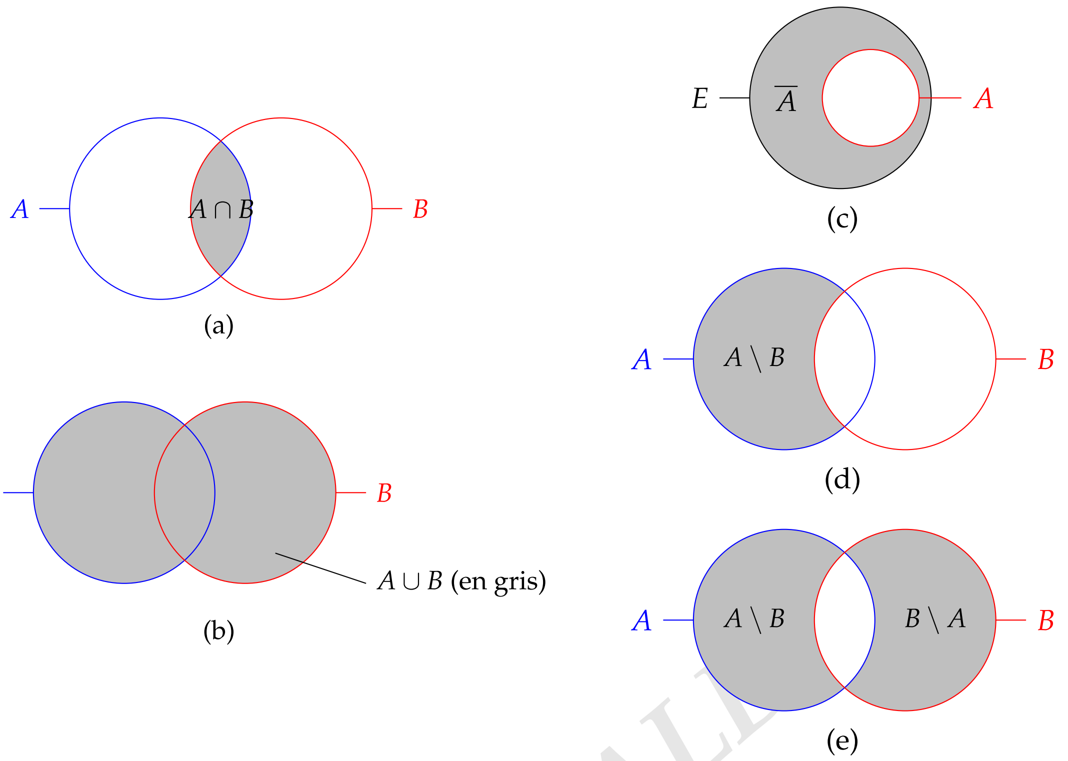


FIGURE 1.1 – Illustration des différentes opérations d'ensembles (diagrammes de Venn). Les parties en grises représentent respectivement : (a) $A \cap B$; (b) $A \cup B$; (c) $\overline{A}$; (d) $A \setminus B$; (e) $A \Delta B$.

Illustrons les différentes définitions à l'aide d'un diagramme d'Euler-Venn [1] ou d'Euler-Venn [2].

Remarque 1.7 On peut étendre les notions d'intersection et de réunion à une famille constituée de plus de deux ensembles. Soient $A_1, A_2, \ldots, A_n$ des ensembles.

- La réunion de $A_1, A_2, \ldots, A_n$ notée $A_1 \cup A_2 \cup \cdots \cup A_n$ ou encore $\bigcup_{i=1}^{n} A_i$ est l'ensemble des éléments appartenant à l'un au moins des A_i.
- L'intersection de $A_1, A_2, \ldots, A_n$ notée $A_1 \cap A_2 \cap \cdots \cap A_n$ ou encore $\bigcap_{i=1}^{n} A_i$ est l'ensemble des éléments appartenant à tous les A_i.

Le théorème suivant donne une propriété très utile dans la manipulation des ensembles.

Théorème 1.8 (Lois de De Morgan [3]) *Soient Ω un ensemble et A, B deux parties de Ω. Alors :*

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \qquad et \qquad \overline{A \cap B} = \overline{A} \cup \overline{B}. \tag{1.1}$$

Donnons à présent quelques règles de calcul des cardinaux d'ensembles obtenus suite à des opérations sur des ensembles finis de cardinaux connus.

1. John Venn (1834-1923), mathématicien et logicien britannique
2. Leonhard Euler (1707-1783), mathématicien et physicien suisse

Théorème 1.9 *Soit Ω un ensemble fini et soient A et B des parties de Ω. Alors, on a :*

1) $\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B)$

2) $\mathrm{Card}(A \setminus B) = \mathrm{Card}(A) - \mathrm{Card}(A \cap B)$

3) $\mathrm{Card}(A \Delta B) = \mathrm{Card}(A \setminus B) + \mathrm{Card}(B \setminus A)$

4) $\mathrm{Card}(\overline{A}) = \mathrm{Card}(\Omega) - \mathrm{Card}(A)$.

Application 1.10 On a relevé les systèmes d'exploitation installés sur l'ensemble des ordinateurs (ensemble Ω) des membres d'un laboratoire. On a observé que le système d'exploitation Windows est installé sur 65 ordinateurs et le système Linux sur 25 ordinateurs. On a aussi observé que 10 ordinateurs sont en *dual-boot* (ils ont les deux systèmes à la fois). On désigne par W et L respectivement les ensembles d'ordinateurs ayant Windows et Linux. Traduire les hypothèses de l'énoncé et déterminer le nombre d'ordinateurs :

1) ayant au moins l'un des deux systèmes;

2) n'ayant pas Linux;

3) n'ayant aucun des deux systèmes (ayant un autre système, Mac OS par exemple);

4) ayant Windows mais pas Linux;

5) ayant un seul des deux systèmes.

Définition 1.11 Soient A et B deux ensembles. On appelle *produit cartésien* de deux ensembles A et B, noté $A \times B$, l'ensemble des couples (a, b) tels que $a \in A$ et $b \in B$ c'est-à-dire

$$A \times B = \{(a, b) \mid a \in A \text{ et } b \in B\}.$$

Lorsque $A = B$, le produit cartésien $A \times B$ est noté A^2 .

Exemple 1.12 Considérons le lancer de deux dés différents dont les faces sont numérotées de 1 à 6. Le résultat d'un lancer est de la forme (x_1, x_2) où x_1 (resp. x_2) représente le chiffre obtenu sur la face supérieure du premier (resp. second) dé. Comme $x_1 \in \{1, \ldots, 6\}$ et $x_2 \in \{1, \ldots, 6\}$, l'ensemble des résultats est donc le produit cartésien $\{1, \ldots, 6\}^2$.

On peut généraliser le produit à cartésien à un nombre fini quelconque d'ensembles.

Définition 1.13 Soient $n \in \mathbb{N}^*$ et $E_1, \ldots, E_n$ des ensembles. Pour tous $x_1 \in E_1, \ldots, x_n \in E_n$, l'élément $(x_1, \ldots, x_n)$ est un $n-$uplet et on note $E_1 \times \cdots \times E_n$ l'ensemble de tous ces $n-$uplets c'est-à-dire

$$E_1 \times \cdots \times E_n = \{(x_1, \ldots, x_n) \mid x_1 \in E_1 \text{ et } \ldots \text{ et } x_n \in E_n\}.$$

Théorème 1.14 *Soient n un entier naturel non nul et $E_1, \ldots, E_n$ des ensembles finis. Alors,*

$$\mathrm{Card}(E_1 \times E_2 \times \cdots \times E_n) = \mathrm{Card}(E_1) \times \mathrm{Card}(E_2) \times \cdots \times \mathrm{Card}(E_n). \tag{1.2}$$

Exemple 1.15 Considérons les numéros de téléphone formés de 8 chiffres dont chaque chiffre peut prendre les valeurs dans l'ensemble $E = \{0, 1, \ldots, 9\}$. Un numéro de téléphone peut être considéré comme un $8-$uplet $(x_1, \ldots, x_8)$ où x_i ($i = 1, \ldots, 8$) représente le $i^{\text{ème}}$ chiffre du numéro. L'ensemble des numéros est donc le produit cartésien $E^8 = \{0, 1, 2, \ldots, 9\}^8$ et le nombre total de numéros est

$$\text{Card}(E^8) = (\text{Card}(E))^8 = 10^8.$$

1.2 Formules classiques d'analyse combinatoire

1.2.1 Généralités

L'analyse combinatoire est une branche des mathématiques qui étudie comment compter les dispositions que l'on peut former à l'aide des éléments d'un ensemble fini. Une *disposition* est un ensemble d'éléments pris parmi les éléments étudiés. Deux critères importants sont utilisés pour distinguer les dispositions : l'*ordre* et la *répétition*.

Définition 1.16

1) Une disposition est dite *ordonnée* lorsque l'ordre des éléments qui la composent est important. Elle est dite *non-ordonnée* sinon. Une disposition ordonnée contenant n éléments est encore appelée $n-$*uplet*.
2) Une disposition est dite *sans répétition* lorsqu'un élément y apparaît au plus une fois. Elle est dite *avec (possibilité de) répétition* sinon.
3) Une disposition non-ordonnée est appelée une *combinaison* et une disposition ordonnée est appelée un *arrangement*.

Exemple 1.17

1) Dispositions ordonnées (arrangements) avec répétition : caractères d'une plaque minéralogique, un numéro de téléphone, un mot de passe, un code, les numéros issus d'un tirage de loterie, ...
2) Dispositions ordonnées (arrangements) sans répétition : le classement d'une compétition, les membres composant un comité (les postes étant deux à distincts et le cumul de postes étant interdit), ...
3) Dispositions non-ordonnées (combinaisons) sans répétition : la liste des équipes qualifiées pour la phase finale de la coupe du monde, la liste des étudiants assidus au cours, ...
4) Dispositions non-ordonnées (combinaisons) avec répétition : distribution d'objets à des personnes (chacun pouvant se servir plusieurs fois), ...

Remarque 1.18 Dans la suite du cours, les dispositions ordonnées seront notées avec des parenthèses et celles non-ordonnées sans répétition seront notées avec des accolades. Les dispositions non ordonnées avec répétition seront notées avec des crochets et ne doivent pas être confondues avec les intervalles. Par exemple, soit $\Omega = \{a, b, c, d\}$. Alors

(a) (d, a, b, c, a, b, a, d) est une disposition ordonnée de 8 éléments avec répétition ;

(b) $[a, a, b, c, a, b, a, d]$ est une disposition non-ordonnée de 8 éléments avec répétition ;

(c) (d, a, b) est une disposition ordonnée de 3 éléments sans répétition ;

(d) $\{a, b\}$ est une disposition non-ordonnée de 2 éléments sans répétition.

1.2.2 Principe multiplicatif

Il permet de compter le nombre de résultats d'une expérience pouvant se décomposer en k étapes *successives* ($k \in \mathbb{N}^*$), la $i-$ème étape ayant n_i choix possibles ($i = 1, \ldots, k$). Pour la première étape, il y a n_1 choix possibles, et pour chaque choix de l'étape 1, il y a n_2 choix possibles pour l'étape 2, et ainsi de suite. Le nombre total de résultats possibles de l'expérience globale est donc

$$n = \prod_{i=1}^{k} n_i = n_1 \times n_2 \times \cdots \times n_k.$$

Application 1.19 Une assemblée élit son bureau composé d'un président, d'un trésorier et d'un secrétaire. Le cumul de poste étant interdit, combien de bureaux peut-on former :

1) avec huit candidats (en supposant que chaque candidat est éligible à tous les postes) ?

2) sachant qu'il y a deux candidats pour le poste de président, trois pour le poste de trésorier et trois pour le poste de secrétaire ?

1.2.3 Principe d'addition

Il permet de compter le nombre de dispositions lorsque l'on raisonne par cas.

Proposition 1.20 (Principe d'addition) *Soient n un entier naturel tel que $n \geqslant 2$ et A_1, A_2, ..., A_n des ensembles finis* deux à deux disjoints *c'est-à-dire que pour tous $i, j \in \{1, \ldots, n\}$, on a $A_i \cap A_j = \emptyset$ dès que $i \neq j$. Alors, on a l'égalité*

$$\text{Card}(A_1 \cup A_2 \cup \cdots A_n) = \text{Card}(A_1) + \text{Card}(A_2) + \cdots + \text{Card}(A_n)$$

que l'on peut réécrire plus simplement

$$\text{Card}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \text{Card}(A_i).$$

Application 1.21 Soit $\Omega = \{A, B, C, D, E\}$. Combien de mots de 3 lettres constitués uniquement de consonnes ou uniquement de voyelles peut-on former avec les lettres de Ω ?

1.2.4 Arrangements avec répétition

Proposition 1.22 *Soit Ω un ensemble de cardinal n et soit $p \in \mathbb{N}$. Le nombre d'arrangements avec répétition de p éléments parmi n est*

$$\underbrace{n \times n \times \cdots \times n}_{p \text{ fois}} = n^p.$$

Exemple 1.23 Soit $\Omega = \{1, 2, 3, 4, 5\}$. Combien de nombres de trois chiffres peut-on former avec les éléments de Ω ? Un nombre de trois chiffres est de la forme $(x_1, x_2, x_3) \in \Omega^3$. C'est donc un arrangement avec répétition de 3 éléments parmi 5. Le nombre de possibilités est donc

$$N = 5 \times 5 \times 5 = 5^3 = 125.$$

1.2.5 Arrangements sans répétition - Permutations

Proposition 1.24 *Le nombre d'arrangements sans répétition de p éléments parmi n ($p \leqslant n$) est*

$$A_n^p = \underbrace{n \times (n-1) \times \cdots \times (n-p+1)}_{p \text{ facteurs}}. \tag{1.3}$$

Application 1.25 Une compétition réunit 16 équipes. Un podium est un triplet (x_1, x_2, x_3) avec x_1, x_2 et x_3 désignant respectivement le vainqueur, le finaliste malheureux et l'équipe classée troisième. Combien y a-t-il de podiums possibles ?

Définition 1.26 (Permutation) Lorsque $p = n$, un arrangement de p éléments parmi n est appelé une *permutation*. Une permutation de n éléments distincts est donc un réarrangement ordonné, sans répétition de ces n éléments.

Exemple 1.27 Les permutations possibles de trois éléments A, B et C sont ABC, ACB, BAC, BCA, CAB et CBA.

Proposition 1.28 *Soit $n \in \mathbb{N}^*$. Le nombre de permutations possibles de n éléments distincts vaut :*

$$A_n^n = \underbrace{n \times (n-1) \times \cdots \times 1}_{n \text{ facteurs}}$$

et est noté $n!$ (lire « factorielle n »). Par convention, $0! = 1$.

Exemple 1.29 De combien de façons peut-on faire asseoir 5 personnes dans une rangée de 5 chaises ? Il s'agit du nombre de permutations de 5 personnes. Le nombre de possibilités est donc

$$N = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120.$$

1.2.6 Combinaisons sans répétition

1.2.6.1 Définition et exemple

Proposition 1.30 *Soit Ω un ensemble de cardinal n et soit $p \in \mathbb{N}^*$ tel que $p \leqslant n$. Le nombre de combinaisons sans répétition de p éléments parmi n (ou le nombre de parties de Ω contenant p éléments) est*

$$C_n^p = \frac{n!}{p!(n-p)!}. \tag{1.4}$$

Exemple 1.31 Soit $\Omega = \{a, b, c, d\}$ un ensemble à 4 éléments. Le nombre de parties de Ω contenant 2 éléments est aussi le nombre de combinaisons sans répétition de 2 éléments parmi 4. Il vaut

$$C_4^2 = \frac{4!}{2!2!} = 6.$$

Si on essaie de compter les parties de E ayant 2 éléments, on obtient :

$$\{a,b\}, \quad \{a,c\}, \quad \{a,d\}, \quad \{b,c\}, \quad \{b,d\} \quad \text{et} \quad \{c,d\}.$$

Elles sont bien au nombre de 6.

Application 1.32 On demande de pronostiquer les deux finalistes de la CM. Combien y a-t-il de possibilités ?

1.2.6.2 Propriétés

Proposition 1.33 *Soient n et p deux entiers naturels tels que $p \leqslant n$. On a :*
(1) $C_n^0 = 1$, (2) $C_n^1 = n$, (3) $C_n^n = 1$, (4) $C_n^p = C_n^{n-p}$.

Théorème 1.34 (Formule du binôme de Newton) *Soient $n \in \mathbb{N}$ et $x, y \in \mathbb{R}$. On a :*

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k} = \sum_{k=0}^{n} C_n^k x^{n-k} y^k. \tag{1.5}$$

1.2.6.3 Cardinal de l'ensemble des parties

Proposition 1.35 *Soit Ω un ensemble fini tel que* $\mathrm{Card}(\Omega) = n$. *Le nombre total de parties de Ω est :*

$$\mathrm{Card}(\mathcal{P}(\Omega)) = 2^n.$$

Preuve : On peut faire la preuve par deux méthodes.

- **Méthode 1**. Le nombre total de parties de Ω est le nombre de parties à 0 élément + le nombre de parties à 1 élément + le nombre de parties à 2 éléments + $\cdots$ + le nombre de parties à n éléments. Ainsi :

$$\begin{aligned}\mathrm{Card}(\mathscr{P}(\Omega)) &= C_n^0 + C_n^1 + \cdots + C_n^n = \sum_{k=0}^{n} C_n^k = \sum_{k=0}^{n} C_n^k (1)^k (1)^{n-k} \\ &= (1+1)^n \quad \text{d'après la formule du binôme de Newton} \\ &= 2^n.\end{aligned}$$

- **Méthode 2**. Soient x_1, ..., x_n les éléments de Ω. La preuve peut aussi se faire en remarquant qu'une partie A de E peut être représentée par le $n-$uplet

$$(\chi_1, \ldots, \chi_n) \in \{0,1\}^n,$$

où pour tout $i = 1, \ldots, n$, χ_i vaut 1 si $x_i \in A$ et 0 sinon. Ainsi,

$$\text{Card}(\mathscr{P}(\Omega)) = \text{Card}\left(\{0,1\}^n\right) = 2^n.$$

□

Exemple 1.36 Si $\Omega = \{a, b, c\}$ alors $\mathscr{P}(\Omega)$ a $2^3 = 8$ parties. On peut les énumérer :

- l'ensemble vide : $\emptyset$,
- trois (03) singletons : $\{a\}$, $\{b\}$, $\{c\}$,
- trois (03) parties ayant deux éléments : $\{a, b\}$, $\{a, c\}$, $\{b, c\}$,
- Ω tout entier : $\{a, b, c\}$.

1.2.7 Combinaisons avec répétition

Proposition 1.37 *Soit Ω un ensemble de cardinal n et soit p un entier naturel non nul. Le nombre de combinaisons avec répétition de p éléments parmi n est noté K_n^p et égal à*

$$K_n^p = C_{n+p-1}^p. \tag{1.6}$$

Soient $n \in \mathbb{N}^*$ et $E = \{e_1, \ldots, e_n\}$. Une combinaison de p éléments parmi n avec possibilité de répétition est une disposition de la forme $[y_1, \ldots, y_p]$ où chaque y_i $(i = 1, \ldots, p)$ est un élément de E. Elle peut être représentée par un $n-$uplet de la forme $(x_1, \ldots, x_n)$ où x_i $(i = 1, \ldots, n)$ représente le nombre de fois que l'élément e_i apparaît dans la disposition $[y_1, \ldots, y_p]$. Les entiers x_i sont alors liés par la relation $x_1 + \cdots + x_n = p$.

Théorème 1.38 *Soient $n \in \mathbb{N}^*$ et $p \in \mathbb{N}$. Le nombre de $n-$uplets $(x_1, \ldots, x_n) \in \mathbb{N}^n$ qui vérifient la relation*

$$x_1 + \cdots + x_n = p \tag{1.7}$$

est égal à $K_n^p = C_{n+p-1}^p$. En d'autres termes,

$$\text{Card}\left((x_1, \ldots, x_n) \in \mathbb{N}^n \mid x_1 + \cdots + x_n = p\right) = K_n^p = C_{n+p-1}^p. \tag{1.8}$$

Application 1.39 De combien de façons peut-on distribuer 5 chemises identiques dans 2 tiroirs ?

1.3 Exercices

Exercice 1.1 On dispose de cinq tiroirs notés A, B, C, D et E dans lesquels on doit ranger à sa guise 3 objets discernables numérotés 1 à 3. On suppose que chaque tiroir est suffisamment grand pour contenir tous les objets.

1) Combien de rangements peut-on faire ?
2) Combien de rangements peut-on faire si l'on ne veut pas ranger deux objets dans un même tiroir ?

Exercice 1.2 On doit répartir 6 enveloppes contenant chacune une certaine somme d'argent entre 6 personnes. On ne s'intéresse pas à la forme des enveloppes mais seulement à leurs contenus respectifs et on suppose que chaque personne ne reçoit pas plus d'une enveloppe. De combien de façons peut-on faire la répartition

1) si on suppose que les contenus sont tous différents ?
2) si on suppose que trois des enveloppes contiennent le même montant (50 000 F) et les trois autres contiennent respectivement 10 000 F, 20 000 F et 30 000 F ?
3) si on suppose que les six enveloppes contiennent le même montant ?

Exercice 1.3 On suppose que les numéros de téléphone à 8 chiffres peuvent comporter à chaque position n'importe quel chiffre de 0 à 9. Combien de numéros de téléphone peut-on former :

1) au total ?
2) si les chiffres sont deux à deux distincts ?
3) si les numéros commencent par 22 ?
4) si l'on ne doit utiliser que les chiffres 0, 2 et 3 ?
5) si les numéros doivent contenir exactement deux chiffres ?
6) si les numéros contiennent exactement deux fois le chiffre 5 ?
7) si les numéros contiennent exactement deux fois le chiffre 0, deux fois le chiffre 3 et quatre fois le chiffre 2 ?
8) si les chiffres pairs et impairs doivent alterner ?
9) la somme des chiffres du numéro est égale à 9 ?
10) la somme des chiffres du numéro est égale à 10 ?

Exercice 1.4 Quatre personnes se présentent à un entretien après lequel le recruteur a la possibilité soit de ne choisir personne, soit d'en choisir une, deux, trois ou quatre. Quel est le nombre total de possibilités ?

Exercice 1.5 Le tiercé est une forme de pari qui consiste à donner les trois premiers chevaux ayant franchi la ligne d'arrivée. Quinze (15) chevaux sont au départ d'un grand prix. Combien y a-t-il de tiercés possibles :

1) au total ?
2) dans lesquels les 3 chevaux de tête sont prédits dans l'ordre ?

3) dans lesquels les 3 chevaux de tête sont prédits dans l'ordre ou le désordre ?

4) dans lesquels les 3 chevaux de tête sont prédits dans le désordre ?

Exercice 1.6 Lors d'un recrutement pour 4 postes identiques, 6 femmes et 8 hommes se présentent. Combien de recrutements distincts sont possibles

1) au total ?

2) sachant que l'on embauche 2 hommes et 2 femmes ?

Exercice 1.7 Un facteur dispose de p enveloppes à répartir dans n boîtes aux lettres. On suppose qu'une boîte peut contenir plusieurs enveloppes. De combien de façons peut se faire cette répartition :

1) si les enveloppes sont distinctes ?

2) si les enveloppes sont identiques ?

Exercice 1.8 Une multinationale décide de lancer un dentifrice pour chien. Le nom de ce nouveau produit doit comporter 3 lettres de l'alphabet. Déterminer les nombres :

1) N_1 de noms que l'on peut fabriquer

2) N_2 de noms contenant trois lettres distinctes

3) N_3 de mots commençant et finissant par une consonne

4) N_4 de mots contenant au moins une voyelle

5) N_5 de mots contenant deux consonnes et une voyelle.

Exercice 1.9

1) Montrer que le nombre N de façons de répartir n personnes en trois groupes de n_1, n_2 et n_3 personnes respectivement tels que $n_1 + n_2 + n_3 = n$ est donné par

$$N = \frac{n!}{n_1!n_2!n_3!}.$$

2) Dans un commissariat, huit (08) agents de police sont répartis aléatoirement de telle sorte qu'il y ait 3 agents de patrouille, 2 agents de garde au commissariat et 3 agents de réserve. Quel est le nombre total de répartitions possibles ?

CHAPITRE 2

PHÉNOMÈNES ALÉATOIRES ET MODÈLE PROBABILISTE

Résumé – Ce chapitre introduit la notion de probabilité et présente le concept classique d'équiprobabilité lié à un ensemble fini. L'objectif de ce chapitre est de savoir calculer des probabilités dans un cadre d'équiprobabilité sur un ensemble fini.

Sommaire

2.1 NOTION D'EXPÉRIENCE ALÉATOIRE

Les phénomènes dans lesquels apparaît souvent l'effet du hasard sont d'un grand intérêt dans plusieurs domaines très différents. A titre d'exemples, on peut citer les jeux de hasard, les banques et la finance (arrivée des clients à un guichet, évaluation des capacités d'un individu pris au hasard à rembourser un crédit avant de le lui octroyer, gestion de portefeuille par les compagnies boursières), la biologie (où le hasard est omniprésent et considéré comme source de la diversité biologique des individus et de la variabilité des caractères), la médecine (où on évalue les « chances » de succès de divers traitements), l'électronique (durée de vie des composantes électroniques), l'informatique (théorie des graphes, filtrage de spams, théorie des files d'attente et des réseaux), l'assurance (évaluation des risques de sinistres, primes payées par les assurés), etc. . .

Définition 2.1

- Une expérience *aléatoire* est une expérience dont il est impossible de prévoir le résultat, c'est-à-dire, qui répétée dans des conditions identiques, peut donner des résultats différents.

- L'ensemble des résultats possibles, souvent noté Ω, est appelé *univers des possibles* (ou tout simplement *univers*) ou encore *espace fondamental*.

Exemple 2.2 Voici quelques exemples d'expériences aléatoires et les univers qu'on peut leur associer :

TABLE 2.1 – Exemples d'expériences aléatoires et d'univers associés

Expérience	**Univers des possibles**
Jet d'un dé à six faces	$\Omega = \{1, 2, 3, 4, 5, 6\}$
Jet de deux dés	$\Omega = \{1, \ldots, 6\}^2 = \{(i, j) \in \mathbb{N}^2 \mid 1 \leqslant i, j \leqslant 6\}$
Lancer d'une pièce	$\Omega = \{\text{pile}, \text{face}\}$
Durée de fonctionnement d'un appareil	$\Omega = \mathbb{R}_+$

Bien que le résultat précis d'une expérience aléatoire soit imprévisible, l'observation et l'intuition amènent à penser que ces phénomènes obéissent à certaines lois. La théorie des probabilités vise à fournir un *modèle mathématique* pour décrire ces phénomènes. Un *modèle* associé à une expérience aléatoire contient trois objets mathématiques essentiels :

- l'espace fondamental Ω,
- une classe $\mathcal{A}$ de sous-ensembles de Ω appelés *événements*,
- une *probabilité* $\mathbb{P}$ sur l'espace $(\Omega, \mathcal{A})$, c'est-à-dire l'affectation d'un poids à chaque événement traduisant la « chance » de réalisation dudit événement.

2.2 Evénements

2.2.1 Notion d'événement

La théorie des probabilités utilise le langage des ensembles pour modéliser une expérience aléatoire. Dans toute la suite, nous considérerons une expérience aléatoire d'univers Ω.

Définition 2.3

- Un *événement* A est une propriété dont on peut dire si elle est vraie ou non, une fois l'expérience accomplie. On dira que l'événement A est *réalisé* ou non suivant que la propriété est vraie ou fausse une fois l'expérience accomplie.
- Un événement A est aussi identifié à l'ensemble, encore noté A, de tous les éléments de Ω pour lesquels ladite propriété est vraie. Si A est l'ensemble associé à un événement et si le résultat de l'expérience $\omega \in A$ alors ledit événement est dit *réalisé*.

Exemple 2.4 On jette un dé. Soit A l'événement « le résultat obtenu est pair ». L'univers des possibles est $\Omega = \{1, 2, 3, 4, 5, 6\}$. L'événement A correspond à l'ensemble $A = \{2, 4, 6\}$. Si l'expérience donne comme résultat 2 alors A est réalisé mais si le résultat est 5 alors A n'est pas réalisé.

2.2.2 Langage des événements

Avec le mode de représentation mathématique des événements introduit ci-dessus, les opérations sur les événements se traduisent par des opérations ensemblistes. Le tableau 2.2 donne la correspondance entre les deux langages.

TABLE 2.2 – Tableau de correspondance entre le langage des événements et celui des ensembles

Vocabulaire probabiliste	*Vocabulaire ensembliste*	*Notation*
Evénement certain	Ensemble entier	Ω
Evénement impossible	Ensemble vide	$\varnothing$
Evénement élémentaire	Singleton	$\{\omega\}$
Evénement contraire de A	Complémentaire de A	$\overline{A}$
A et B	Intersection de A et B	$A \cap B$
A ou B (ou non exclusif)	Réunion de A et B	$A \cup B$
A et B sont incompatibles	A et B sont disjoints	$A \cap B = \varnothing$

Exemple 2.5 On reprend l'exemple du jet d'un dé à six faces numérotées de 1 à 6.

1) L'événement A : « Obtenir le chiffre 7 » est impossible.
2) L'événement B : « Obtenir un chiffre inférieur ou égal à 7 » est certain.
3) Les événements C : « Obtenir un chiffre pair » et D : « Obtenir un chiffre impair » sont contraires et incompatibles.

Remarque 2.6 Les opérations sur les événements peuvent bien sûr faire intervenir plus de deux événements. Ainsi, si $A_1, \ldots, A_n$ sont des événements,

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$$

est l'ensemble des éléments qui sont dans l'un au moins des A_i. C'est donc l'événement

« réalisation de l'un au moins des A_i $(1 \leqslant i \leqslant n)$ ».

De même :

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$$

est l'ensemble des éléments qui sont dans tous les A_i. C'est donc l'événement

« réalisation de tous les A_i $(1 \leqslant i \leqslant n)$ ».

Application 2.7 Trois boules sont tirées successivement d'une urne contenant des boules blanches et des boules rouges. Pour tout $i \in \{1, 2, 3\}$, on définit l'événement :

$$A_i = \text{« la boule n}^{\text{o}}\ i \text{ est blanche ».}$$

Exprimer à l'aide des événements A_1, A_2 et A_3 les événements suivants :

$$\begin{aligned}
B &= \text{« toutes les boules tirées sont blanches »,}\\
C &= \text{« une boule au moins est blanche »,}\\
D &= \text{« la troisième boule est blanche »,}\\
E &= \text{« seule la troisième est blanche »,}\\
F &= \text{« aucune boule n'est blanche »,}\\
G &= \text{« au moins une des boules n'est pas blanche ».}
\end{aligned}$$

Définition 2.8 (Système complet d'événements) Une famille d'événements A_1, A_2, ..., A_n forme un *système complet d'événements* si les parties A_1, A_2, ..., A_n vérifient les conditions suivantes :

1) chaque A_i est non vide ;

2) Les parties A_i sont deux à deux incompatibles :

$$\forall i, j \in \{1, \dots, n\}, i \neq j \implies A_i \cap A_j = \emptyset;$$

3) La réunion des A_i donne Ω i.e. $\bigcup_{i=1}^{n} A_i = \Omega$.

Exemple 2.9 Soit A un événement différent de l'événement impossible et de l'événement certain Ω. Les événements A et $\overline{A}$ forment un système complet d'événements.

Exemple 2.10 Reprenons l'exemple du lancer d'un dé à six faces dont l'univers des possibles est $\Omega = \{1, 2, 3, 4, 5, 6\}$. Alors,

- les événements $A = \{1, 3, 5\}$ (le résultat est impair) et $\overline{A} = \{2, 4, 6\}$ (le résultat est pair) forment un système complet d'événements ;
- les événements $C = \{1, 2\}$ (le résultat est inférieur ou égal à 2), $D = \{3\}$ (le résultat est égal à 3) et $E = \{4, 5, 6\}$ (le résultat est strictement supérieur à 3) forment aussi un système complet d'événements.

2.3 Espaces probabilisés

2.3.1 Définition de la probabilité

Nous commençons par introduire les notions de tribu et d'espace probabilisable, préalables à la définition d'une probabilité.

Définition 2.11 On appelle *tribu* sur Ω, tout sous-ensemble $\mathcal{A}$ de $\mathscr{P}(\Omega)$ vérifiant les trois propriétés suivantes :

1) $\Omega \in \mathcal{A}$;
2) si $A \in \mathcal{A}$ alors $\overline{A} \in \mathcal{A}$ (stabilité par passage au complémentaire) ;
3) si $(A_i)_{i\in\mathbb{N}}$ est une suite d'événements de $\mathcal{A}$, alors $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (stabilité par l'union dénombrable).

Le couple $(\Omega, \mathcal{A})$ est alors appelé *espace probabilisable*.

Exemple 2.12 Pour tout ensemble Ω,

1) la famille $\mathscr{P}(\Omega)$ des parties de Ω, est une tribu ;
2) la famille $\{\emptyset, \Omega\}$ est une tribu appelée *tribu triviale* sur Ω.

Remarque 2.13 Dans les exemples et exercices donnés dans la suite du cours, et sauf mention contraire, nous supposerons $\mathcal{A} = \mathscr{P}(\Omega)$.

Définition 2.14 (Probabilité) On appelle *probabilité* sur un espace probabilisable $(\Omega, \mathcal{A})$, toute application $\mathbb{P}$ définie de $\mathcal{A}$ dans $[0,1]$ telle que :

1) $\mathbb{P}(\Omega) = 1$;
2) pour toute suite $(A_i)_{i\in\mathbb{N}}$ d'événements deux à deux incompatibles :

$$\mathbb{P}\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} \mathbb{P}(A_i). \qquad (\sigma - \text{additivité de } \mathbb{P})$$

Le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ est appelé *espace probabilisé* ou *espace de probabilité*. Le nombre $\mathbb{P}(A)$ est appelé la *probabilité* de l'événement A.

2.3.2 Propriétés

Théorème 2.15 *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Les propriétés suivantes sont satisfaites :*

1) Pour tout $A \in \mathcal{A}$, on a : $\mathbb{P}(A) \in [0,1]$.

2) $\mathbb{P}(\emptyset) = 0$.

3) Pour tous événements $A_1, \ldots, A_n$ deux à deux incompatibles, *on a :*

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

En particulier, pour deux événements incompatibles *A et B, on a :*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

4) *Si A et B sont des événements* quelconques, *on a*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \tag{2.1}$$

5) *Pour tout $A \in \mathcal{A}$, on a : $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.*

6) *Pour deux événements A et B,*

$$\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) \tag{2.2}$$

$$\mathbb{P}(A \Delta B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A). \tag{2.3}$$

7) *Si $A, B \in \mathcal{A}$ et si $A \subset B$, alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.*

La formule suivante, appelée **formule de Poincaré** permet de calculer la probabilité de la réunion de n événements ($n \geqslant 2$). Elle est une généralisation de la formule (2.1).

Proposition 2.16 (Formule de Poincaré) *Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $n \geqslant 2$ et $A_1, A_2, \ldots, A_n$ une suite d'événements de $\mathcal{A}$. On a :*

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mathbb{P}\left(A_i\right) - \sum_{1 \leqslant i < j \leqslant n} \mathbb{P}\left(A_i \cap A_j\right) + \sum_{1 \leqslant i < j < k \leqslant n} \mathbb{P}\left(A_i \cap A_j \cap A_k\right) - \cdots\cdots + (-1)^{n-1} \mathbb{P}(A_1 \cap \cdots \cap A_n), \tag{2.4}$$

ou encore

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{k=1}^{n} \Big((-1)^{k-1} \sum_{1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n} \mathbb{P}\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right)\Big). \tag{2.5}$$

L'application des formules équivalentes (2.4) et (2.5) est aisée. Elle s'inspire de la méthodologie utilisée pour énumérer tous les sous-ensembles d'un ensemble donné. On commence par les ensembles pris individuellement, puis on prend les intersections deux à deux, puis les intersections trois à trois et ainsi de suite jusqu'à prendre l'intersection des n ensembles. On commence par affecter le signe "+" aux probabilités des ensembles pris un à un puis on alterne les signes chaque fois qu'on passe d'un nombre d'ensembles au nombre suivant. Par exemple, pour trois événements A, B et C, on aura :

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C). \tag{2.6}$$

2.4 Notion d'équiprobabilité ou de probabilité uniforme

La théorie des probabilités ne dit pas quelle loi de probabilité mettre sur un ensemble Ω parmi toutes les nombreuses lois possibles. Nous présentons ici l'équiprobabilité ou probabilité uniforme mais évidemment il existe bien d'autres façons de définir une loi de probabilité, chacune ayant des avantages et des limites. Quelque soit la loi choisie, il faut garder à l'esprit qu'un modèle probabiliste n'est qu'une représentation de la réalité et ses hypothèses doivent être mises à l'épreuve des faits.

Lorsque Ω est fini et si les conditions de l'expérience le permettent, on considère *a priori* que les événements élémentaires ont la même probabilité : on parle d'*équiprobabilité* ou de *probabilité uniforme*.

Définition 2.17 Soit $\Omega = \{\omega_1, \ldots, \omega_n\}$ un univers fini. L'équiprobabilité ou probabilité uniforme sur Ω est définie par :

$$\mathbb{P}(\{\omega_i\}) = \frac{1}{n}, \quad \forall i = 1, \ldots, n. \tag{2.7}$$

Exemple 2.18 Si l'expérience aléatoire consiste à jeter un dé parfait (non truqué) à six faces numérotées de 1 à 6, on peut supposer que toutes les faces ont la même probabilité d'apparaître qui égale à $1/6$.

Le calcul des probabilités n'est donc plus qu'une affaire de dénombrement, d'où la célèbre formule suivante.

Proposition 2.19 *Soient $n \in \mathbb{N}^*$, $\Omega = \{\omega_1, \ldots, \omega_n\}$ un univers fini et $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Si $\mathbb{P}$ représente l'équiprobabilité (ou probabilité uniforme) sur Ω, alors pour tout événement $A \in \mathcal{A}$,*

$$\mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = \frac{\textit{Nombre de cas favorables à } A}{\textit{Nombre de cas possibles}}. \tag{2.8}$$

Application 2.20 Dans une urne contenant dix boules indiscernables au toucher numérotées de 0 à 9, on tire successivement trois boules avec remise et on relève dans l'ordre les trois chiffres obtenus. Décrire le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ associé à cette expérience et calculer la probabilité de chacun des évènements suivants :

1) A : « Les trois chiffres obtenus sont identiques ».
2) B : « Les trois chiffres obtenus sont deux à deux distincts ».
3) C : « On n'obtient pas le chiffre 4 ».
4) D : « On obtient le chiffre 4 au moins une fois ».
5) E : « Le second chiffre est impair ».
6) F : « L'un des deux premiers chiffres au moins est pair ».

2.5 EXERCICES

Exercice 2.1 Dans un groupe de 100 personnes, 30 aiment le cinéma (événement C), 40 aiment le théâtre (événement T) et 35 n'aiment aucun des deux loisirs. On sélectionne une personne au hasard. Quelle est la probabilité que cette personne

1) aime au moins un des deux loisirs ?
2) aime les deux loisirs ?
3) aime le cinéma mais pas le théâtre ?
4) aime un seul des deux loisirs ?

Exercice 2.2 On procède à deux jets consécutifs d'un dé parfaitement équilibré dont les faces sont numérotées de 1 à 6 et on relève à chaque fois le numéro obtenu sur la face supérieure du dé. On considère ensuite les deux événements : A « la somme des numéros obtenus est impaire » ; B « le numéro 1 est obtenu au moins une fois ». Calculer les probabilités des événements : A, B, $A \cap B$, $A \cup B$, $A \cap \overline{B}$.

Exercice 2.3 On lance n fois un dé régulier.

1) Calculer la probabilité p_n de l'événement : « On obtient au moins une fois la face six ».
2) Comment choisir n pour que p_n soit supérieur ou égal à 0.95 ?

Exercice 2.4 On relève les dates d'anniversaire dans un groupe de 30 personnes (jours numérotés de 1 à 365). On suppose que les jours de naissance sont équiprobables. Quelle est la probabilité que deux personnes au moins aient le même anniversaire ?

Exercice 2.5 Une entreprise a ses sites de productions répartis sur trois pays : A, B et C. Cette entreprise a 4 sites de production dans le pays A, 6 dans le pays B et 9 dans le pays C. On cherche à sélectionner trois sites pour y installer des centres de recherche.

1) De combien de façons peut-on choisir ces trois sites ?
2) Quelle est la probabilité que chaque pays dispose d'un centre de recherche ?
3) Quelle est la probabilité qu'un seul pays dispose de tous les centres de recherche ?

Exercice 2.6 On effectue cinq lancers d'un dé cubique parfaitement équilibré dont les faces sont numérotées de 1 à 6. On relève les numéros obtenus sur les faces supérieures du dé. Déterminer la probabilité des événements suivants :

1) A : « on n'a pas obtenu la face 6 ».
2) B : « on a obtenu exactement une seule fois la face 6 ».
3) C : « on a obtenu au moins deux fois la face 6 ».
4) D : « on a obtenu 1 pour la première fois au cinquième lancer ».
5) E : « on a obtenu exactement deux fois la face 6 ».
6) F : « on a obtenu exactement deux fois la face 6 et une fois la face 4 ».

Exercice 2.7 Trois personnes vont au cinéma. Il passe trois films différents. Chaque personne choisit son film au hasard sans tenir compte du choix des autres. Quelle est la probabilité que :

1) les trois personnes aient vu les trois films différents ?
2) les trois personnes aient vu le même film ?
3) deux personnes exactement aient vu le même film ?
4) au moins deux personnes aient vu le même film ?

Exercice 2.8 Dix-huit (18) personnes se sont présentées à une collecte de sang. Parmi celles-ci, on a noté : 11 personnes du groupe O, 4 personnes du groupe A, 2 personnes du groupe B et 1 personne du groupe AB. A l'issue de la collecte, on prélève au hasard 3 flacons parmi les 18 flacons obtenus. Calculer la probabilité des événements suivants :

1) E_1 = « les 3 flacons appartiennent au même groupe »
2) E_2 = « parmi les 3 flacons prélevés, il y a au moins un flacon contenant du sang du groupe A »
3) E_3 = « les sangs des 3 flacons appartiennent à des groupes différents »

Exercice 2.9 On compose au hasard un numéro de téléphone à 8 chiffres. Quelle est la probabilité pour que :

1) tous les chiffres du numéro soient distincts ?
2) le produit des chiffres soit divisible par 2 ?
3) les chiffres forment une suite strictement croissante ?
4) les chiffres forment une suite croissante au sens large ?

Exercice 2.10 On range au hasard 10 livres notés $L_1, \ldots, L_{10}$ sur une étagère. Quelle est la probabilité p pour que 3 livres donnés (par exemple L_1, L_4 et L_{10}) soient rangés côte à côte ?

CHAPITRE 3

PROBABILITÉ CONDITIONNELLE ET INDÉPENDANCE

Résumé – Ce chapitre présente les concepts de probabilité conditionnelle et d'indépendance. Deux importantes formules, le théorème des probabilités totales et la formule de Bayes, y sont présentées. L'objectif est de savoir calculer des probabilités dans un contexte de probabilités conditionnelles.

Sommaire

3.1 INTRODUCTION ET DÉFINITION

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ l'espace probabilisé associé à une expérience aléatoire. Nous avons vu dans le chapitre précédent que la probabilité d'un événement $A \in \mathcal{A}$ pouvait être considérée comme une quantification *a priori* (c'est-à-dire avant la réalisation de l'expérience) des « chances » de réalisation de A. Si l'on dispose d'une information supplémentaire sur la réalisation d'un autre événement B, il est logique de penser que cette nouvelle information pourrait modifier les « chances » de voir A se réaliser. Illustrons cela à l'aide l'exemple suivant :

Exemple 3.1 Dans un groupe de 100 personnes, il y a 40 femmes dont 10 sont au chômage et 60 hommes dont 10 sont chômage.

1) Si on tire une personne au hasard dans la population, la probabilité qu'elle soit au chômage est

$$p_1 = \frac{20}{100} = 0.2$$

car il y a au total 20 personnes (10 hommes et 10 femmes) au chômage.

2) Maintenant, si on suppose que la personne tirée est une femme, la probabilité qu'elle soit au chômage est
$$p_2 = \frac{10}{40} = 0.25.$$

La notion de probabilité conditionnelle abordée dans ce chapitre va permettre, dans une certaine mesure, de rendre compte de l'information apportée par la réalisation d'un événement donné sur la réalisation (éventuelle) d'un autre événement.

Théorème – Définition 3.2 *Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et B un événement de probabilité $\mathbb{P}(B) > 0$. Alors la fonction d'ensembles $\mathbb{P}(\cdot|B)$, définie pour tout événement $A \in \mathcal{A}$ par*
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{3.1}$$
est une probabilité sur $(\Omega, \mathcal{A})$. On l'appelle probabilité conditionnelle sachant *B.*

Exemple 3.3 On lance deux fois un dé parfaitement équilibré dont les faces sont numérotées de 1 à 6. Quelle est la probabilité d'obtenir la face 6 au premier lancer sachant que la somme des chiffres obtenus vaut 8 ?

- L'univers Ω est l'ensemble des couples (x_1, x_2) où chaque $x_i \in \{1, \ldots, 6\}$. On a $\mathrm{Card}(\Omega) = 6^2 = 36$.
- Soit les événements A : « Obtenir la face 6 au premier lancer » et B : « la somme des chiffres obtenus vaut 8 ». Il s'agit de calculer $\mathbb{P}(A|B)$.
- $A = \{(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$, $B = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$ et $A \cap B = \{(6,2)\}$.
- On a :
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/36}{5/36} = \frac{1}{5}.$$

Remarque 3.4 La probabilité conditionnelle étant avant tout une probabilité, elle vérifie donc les propriétés données par le théorème 2.15 (page 25). Si $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace probabilisé et B est un événement tel que $\mathbb{P}(B) > 0$, alors, pour tous événements A, A_1 et A_2, on a :

1) $\mathbb{P}(A|B) \in [0, 1]$.
2) $\mathbb{P}(\varnothing|B) = 0$, $\mathbb{P}(\Omega|B) = 1$.
3) $\mathbb{P}(A_1 \cup A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) - \mathbb{P}(A_1 \cap A_2|B)$.
4) $\mathbb{P}(\overline{A}|B) = 1 - \mathbb{P}(A|B)$.
5) $\mathbb{P}((A_1 \setminus A_2)|B) = \mathbb{P}(A_1|B) - \mathbb{P}(A_1 \cap A_2|B)$.
6) $\mathbb{P}(A_1 \Delta A_2|B) = \mathbb{P}((A_1 \setminus A_2)|B) + \mathbb{P}((A_2 \setminus A_1)|B)$.

3.2 NOTION D'INDÉPENDANCE

3.2.1 Indépendance de deux événements

Définition 3.5 Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Deux événements A et B sont *indépendants* pour la probabilité $\mathbb{P}$ s'ils vérifient la propriété :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B). \tag{3.2}$$

Application 3.6 On lance deux fois de suite un dé parfaitement équilibré et on note les deux chiffres obtenus. On considère les événements A : « le premier chiffre est pair » et B : « le second chiffre est supérieur ou égal à 5 ». Montrons que les événements A et B sont indépendants.

- L'univers Ω est l'ensemble $\{1, \ldots, 6\}^2$. On a $\text{Card}(\Omega) = 6^2 = 36$.
- $A = \{2, 4, 6\} \times \{1, \ldots, 6\}$, $\text{Card}(A) = 3 \times 6 = 18$ et $\mathbb{P}(A) = 18/36 = 1/2$.
- $B = \{1, \ldots, 6\} \times \{5, 6\}$, $\text{Card}(B) = 6 \times 2 = 12$ et $\mathbb{P}(B) = 12/36 = 1/3$.
- $A \cap B = \{2, 4, 6\} \times \{5, 6\}$, $\text{Card}(A \cap B) = 3 \times 2 = 6$ et $\mathbb{P}(A \cap B) = 6/36 = 1/6$.
- On a $\mathbb{P}(A) \times \mathbb{P}(B) = \mathbb{P}(A \cap B) = 1/6$ d'où on conclut que A et B sont indépendants.

Remarque 3.7 Si deux événements de probabilités non nulles A et B sont indépendants, alors

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

et on a de même $\mathbb{P}(B|A) = \mathbb{P}(B)$. La probabilité de réalisation de l'un n'est donc pas modifiée par une information concernant la réalisation de l'autre.

3.2.2 Propriétés

Proposition 3.8 *Si A et B sont deux évènements indépendants alors les événements A et $\overline{B}$, $\overline{A}$ et B, $\overline{A}$ et $\overline{B}$ le sont aussi.*

Preuve : Il suffit d'appliquer la définition de l'indépendance. Supposons que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. On a :

$$\begin{aligned} \mathbb{P}(A \cap \overline{B}) &= \mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(A)(1 - \mathbb{P}(B)) \\ &= \mathbb{P}(A)\mathbb{P}(\overline{B}). \end{aligned}$$

$$\begin{aligned} \mathbb{P}(\overline{A} \cap B) &= \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(B)(1 - \mathbb{P}(A)) \\ &= \mathbb{P}(B)\mathbb{P}(\overline{A}). \end{aligned}$$

$$
\begin{aligned}
\mathbb{P}(\overline{A} \cap \overline{B}) &= \mathbb{P}(\overline{A \cup B}) = 1 - \mathbb{P}(A \cup B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B) \\
&= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B) = (1 - \mathbb{P}(A))(1 - \mathbb{P}(B)) \\
&= \mathbb{P}(\overline{A})\mathbb{P}(\overline{B}).
\end{aligned}
$$

□

3.2.3 Incompatibilité et indépendance

Il ne faut pas confondre les notions d'incompatibilité et d'indépendance. Deux événements A et B sont incompatibles si $A \cap B = \emptyset$ c'est-à-dire que A et B ne peuvent pas se réaliser au même moment. Par contre A et B sont indépendants si la réalisation de l'un n'influence pas la réalisation de l'autre.

Exemple 3.9 On lance un dé parfaitement équilibré à six faces numérotées de 1 à 6. On considère les événements A : « le chiffre obtenu est pair » et B : « le chiffre obtenu est 5 ». L'ensemble fondamental associé à cette expérience est $\Omega = \{1, \ldots, 6\}$. On peut vérifier que $A \cap B = \emptyset$ et

$$\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2} \quad \text{et} \quad \mathbb{P}(B) = \frac{1}{6}.$$

Les événements A et B sont incompatibles mais ne sont pas indépendants puisque $0 = \mathbb{P}(A \cap B) \neq \mathbb{P}(A) \times \mathbb{P}(B)$.

De façon intuitive, on peut même imaginer que deux événements incompatibles A et B ne peuvent pas être indépendants puisque par définition de l'incompatibilité, la réalisation de A par exemple empêche celle de B. Ceci est formalisé par la proposition suivante.

Proposition 3.10 *Si A et B sont deux événements incompatibles de probabilités non nulles, alors A et B ne sont pas indépendants.*

Preuve : Par définition de l'incompatibilité, on a $\mathbb{P}(A \cap B) = 0$. Or $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$ d'où $\mathbb{P}(A)\mathbb{P}(B) > 0$. On conclut que $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)$. □

3.2.4 Indépendance mutuelle

Définition 3.11 Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Les événements $A_1, A_2, \ldots, A_n$ sont dits *mutuellement indépendants* si pour tout sous-ensemble d'indices $I \subset \{1, \ldots, n\}$ ayant au moins deux éléments, on a :

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i).$$

Remarque 3.12 L'indépendance mutuelle entraîne l'indépendance deux à deux mais la réciproque est fausse. En d'autres termes, l'indépendance mutuelle est beaucoup plus forte que l'indépendance deux à deux. Par exemple, on dira que trois événements A_1, A_2 et A_3 sont mutuellement indépendants si les quatre égalités suivantes sont vérifiées :

$$\left\{\begin{array}{rcl} \mathbb{P}(A_1 \cap A_2) & = & \mathbb{P}(A_1)\mathbb{P}(A_2) \\ \mathbb{P}(A_1 \cap A_3) & = & \mathbb{P}(A_1)\mathbb{P}(A_3) \\ \mathbb{P}(A_2 \cap A_3) & = & \mathbb{P}(A_2)\mathbb{P}(A_3) \\ \mathbb{P}(A_1 \cap A_2 \cap A_3) & = & \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3). \end{array}\right.$$

Exemple 3.13 On lance deux fois un dé équilibré. Soient les événements A_1 = « le 1er nombre obtenu est pair », A_2 = « le 2ème nombre obtenu est impair », A_3 = « la somme des deux nombres est paire ». On vérifie que $\text{Card}(\Omega) = 36$, $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = 18/36 = 1/2$, $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = 1/4$ et $\mathbb{P}(A_1 \cap A_2 \cap A_3) = 0$. Les trois événements sont deux à deux indépendants mais ne sont pas mutuellement indépendants.

3.3 FORMULES IMPORTANTES

3.3.1 Théorème des probabilités totales

Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé et A un événement de probabilité non nulle. On considère un système complet d'événements $B_1, \ldots, B_n$ représentant par exemple l'ensemble de toutes les causes possibles de réalisation de A. On a :

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^{n} B_i \right) = \bigcup_{i=1}^{n} (A \cap B_i).$$

La première égalité est justifiée par le fait que $A \subset \Omega$, la seconde par le fait que le système (B_i) est un système complet d'événements et la dernière, par la distributivité de l'intersection par rapport à la réunion . Comme les B_i sont deux à deux disjoints et que $A \cap B_i \subset B_i$, alors les $A \cap B_i$ sont aussi deux à deux disjoints d'où

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=1}^{n} (A \cap B_i) \right) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i). \tag{3.3}$$

Finalement, en tenant compte de la relation $\mathbb{P}(A|B_i) = \mathbb{P}(A \cap B_i)/\mathbb{P}(B_i)$, on en déduit que $\mathbb{P}(A \cap B_i) = \mathbb{P}(A|B_i)\mathbb{P}(B_i)$. On obtient alors la très importante formule suivante :

Théorème 3.14 (Formule des probabilités totales) *Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, A un événement et $B_1, \ldots, B_n$ un système complet d'événements tel que $\mathbb{P}(B_i) > 0$ pour chaque $i = 1, \ldots, n$. On a :*

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i). \tag{3.4}$$

Application 3.15 Une compagnie d'assurance assure un nombre égal de conducteurs masculins et féminins. On estime la probabilité d'avoir un accident dans l'année à $\alpha = 0.3$ pour un conducteur (masculin) et $\beta = 0.1$ pour une conductrice. La compagnie sélectionne un assuré au hasard. On souhaite déterminer la probabilité pour que la personne sélectionnée ait un accident pendant l'année. On considère les événements M : « La personne sélectionnée est de sexe masculin », F : « La personne sélectionnée est de sexe féminin » et A : « La personne sélectionnée a un accident ». Calculer $\mathbb{P}(A)$.

3.3.2 Formule de Bayes

On conserve les notations de la sous-section précédente. Considérons une des causes susceptibles de réaliser l'événement A, la cause B_k par exemple. On suppose l'événement A réalisé et on cherche à déterminer la probabilité que la cause B_k ait provoqué la réalisation de A c'est-à-dire $\mathbb{P}(B_k|A)$. Pour tout $k \in \{1, \ldots, n\}$, on a :

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B_k)}{P(A)} = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)}.$$

En utilisant la formule (3.4), on obtient le théorème suivant qui porte le nom de *formule de Bayes* [1].

Théorème 3.16 (Formule de Bayes) *Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, A un événement et $B_1, \ldots, B_n$ un système complet d'événements tel que $\mathbb{P}(B_i) > 0$ pour chaque $i = 1, \ldots, n$. Si $\mathbb{P}(A) > 0$, alors on a :*

$$\forall k \in \{1, \ldots, n\}, \quad \mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}. \tag{3.5}$$

Exemple 3.17 Lorsque le système complet est constitué d'un événement B et son contraire $\overline{B}$ tous deux de probabilité non nulle, la formule (3.5) se simplifie en :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\overline{B})\mathbb{P}(\overline{B})}.$$

Application 3.18 Dans une usine trois machines M_1, M_2, M_3 fabriquent des boulons de même type. La machine M_1 sort en moyenne 0.3 % de boulons défectueux, M_2 0.8 % et M_3 1 %. On mélange 1000 boulons dans une caisse, 500 provenant de M_1, 350 de M_2 et 150 de M_3. On tire un boulon au hasard dans la caisse et on constate qu'il est défectueux. Quelle est la probabilité qu'il ait été fabriqué par la machine M_1 ?
N.B. *On pourra considérer les événements M_i : « le boulon tiré provient de la machine i » ($i = 1, 2, 3$) et D : « le boulon tiré est défectueux ».*

1. Thomas Bayes ((1701 ou 1702) - 1761), mathématicien et pasteur britannique

3.4 EXERCICES

Exercice 3.1 Parmi les vols quittant une ville V_1 pour une ville V_2, 89.5% partent à l'heure (événement D) et arrivent à destination à l'heure (événement A), 3.5% partent à l'heure et arrivent en retard, 1.5% partent en retard et arrivent à l'heure, et 5.5% partent en retard et arrivent en retard.

1) Traduire les données de l'énoncé et expliquer pourquoi elles sont cohérentes.
2) Quelle est la probabilité
   (a) qu'un vol parte à l'heure ?
   (b) qu'un vol arrive en retard ?
   (c) qu'un vol arrive à l'heure sachant qu'il est parti à l'heure ?
   (d) qu'un vol arrivé en retard soit parti en retard ?

Exercice 3.2 Un test médical pour une maladie M donne deux résultats possibles : $+$ et $-$. Les probabilités sont données par : $P(+ \cap M) = 0.009$, $P(+ \cap \overline{M}) = 0.099$, $P(- \cap M) = 0.001$ et $P(- \cap \overline{M}) = 0.891$.

1) Calculer $P(M)$, $P(\overline{M})$, $P(+|M)$, $P(-|\overline{M})$, $P(+|\overline{M})$ et $P(-|M)$. Les interpréter.
2) Quelle est la probabilité d'avoir la maladie sachant que le test est positif ? Ce test vous semble-t-il fiable ?

Exercice 3.3 Vous contactez deux amis A et B afin de former un groupe de travail de mathématiques. La probabilité que l'ami A accepte est de 0.2 et celle que l'ami B accepte est de 0.6. De plus, sachant que l'ami A a accepté, l'ami B acceptera avec une probabilité de 0.7. Calculer la probabilité que :

1) vos deux amis acceptent ;
2) l'ami A accepte sachant que B a accepté ;
3) seul l'ami A accepte ;
4) l'un au moins de vos deux amis accepte ;
5) aucun de vos deux amis n'accepte.

Exercice 3.4 Soient A et B deux événements tels que $P(A) = 1/5$ et $P(A \cup B) = 1/2$. Calculer $P(B)$ dans chacun des cas suivants :

1) A et B sont incompatibles.
2) A et B sont indépendants.

Exercice 3.5 On jette deux dés parfaitement équilibrés dont les faces sont numérotées de 1 à 6. Soit A l'événement « la somme des chiffres indiqués est impaire » et soit B l'événement « l'un des dés indique le chiffre 1 ». Les événements A et B sont-ils indépendants ?

Exercice 3.6 Trois personnes A, B et C tirent sur un oiseau. On suppose que les tirs sont indépendants les uns des autres et que les probabilités de succès sont respectivement 70%, 50% et 90%. Calculer la probabilité que l'oiseau soit touché

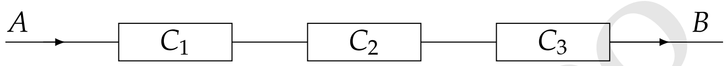
1) en calculant d'abord la probabilité de l'événement contraire ;

2) en utilisant l'expression de $P(A \cup B \cup C)$.

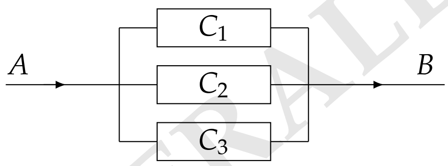
Exercice 3.7 D'après le cours, on sait que l'indépendance mutuelle de trois événements se définit à l'aide de quatre (4) égalités. Combien d'égalités faut-il pour définir l'indépendance mutuelle de n événements ?

Exercice 3.8 On se propose de calculer la probabilité π de fonctionnement des montages électriques ci-après à l'aide des probabilités de fonctionnement de leurs composants. Dans chaque montage, les composants ont des fonctionnements mutuellement indépendants mais ils ont une même probabilité p de fonctionner. On dit que le circuit fonctionne si le courant peut passer du point A au point B et pour tout $i = 1, \ldots, 5$, on note C_i = « le composant C_i fonctionne ».

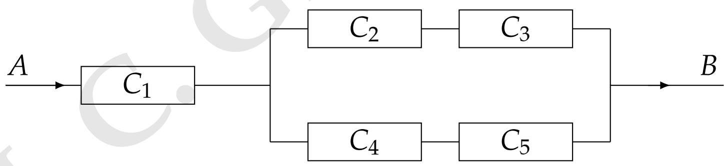
1) Montage en série (nécessite le fonctionnement simultané des composants) :



2) Montage en parallèle (nécessite le fonctionnement d'un composant au moins) :



3) Montage mixte :



Exercice 3.9 Dans une association, deux bénévoles A et B collent des timbres sur des enveloppes. Le travail est réparti entre les deux bénévoles de telle sorte que A traite 40% des enveloppes (et donc B en traite 60%). Sous l'effet de la fatigue, A omet de mettre un timbre dans 3% des cas et B dans 2% des cas. On choisit une enveloppe au hasard. Calculer la probabilité :

1) qu'elle ne porte pas de timbre.

2) qu'elle ait été traitée par A sachant qu'elle ne porte pas de timbre.

Exercice 3.10 On considère une population composée de 48% d'hommes et de 52% de femmes. La probabilité qu'un homme soit daltonien est 0.05, la probabilité qu'une femme soit daltonienne est 0.0025. On choisit une personne au hasard dans la population. Quelle est la probabilité qu'elle soit daltonienne ?

Exercice 3.11 Une usine s'adresse à trois fournisseurs A_1, A_2 et A_3 pour l'approvisionnement d'un composant électronique. On suppose que 30% des composants sont achetés à

A_1, 34% à A_2 et 36% à A_3. Un contrôle de conformité effectué sur un échantillon aléatoire de composants électroniques a révélé que le taux de composants défectueux est de 4% pour A_1, 2% pour A_2 et 7% chez A_3. On choisit un composant au hasard et on désigne par D l'événement : « le composant est défectueux » et par A_i : « le composant choisi provient du fournisseur A_i », $i \in \{1, 2, 3\}$.

1) Calculer la probabilité que le composant soit défectueux.
2) On suppose que le composant choisi est défectueux. Calculer la probabilité qu'il provienne du fournisseur A_1.

Exercice 3.12 Pour se rendre à l'université, un étudiant a le choix entre 4 itinéraires : A, B, C et D. La probabilité qu'il a de choisir l'itinéraire A (resp. B, C) est $\frac{1}{3}$ (resp. $\frac{1}{4}$, $\frac{1}{12}$). La probabilité d'arriver en retard en empruntant A (resp. B, C) est $\frac{1}{20}$ (resp. $\frac{1}{10}$, $\frac{1}{5}$). En empruntant D, il n'est jamais en retard. Calculer la probabilité pour que :

1) l'étudiant choisisse l'itinéraire D ?
2) l'étudiant arrive en retard ?
3) l'étudiant ait emprunté l'itinéraire C sachant qu'il est arrivé en retard ?

Exercice 3.13 On dispose de deux urnes contenant des boules indiscernables au toucher. L'urne 1 contient trois boules blanches et une boule noire. L'urne 2 contient une boule blanche et deux boules noires. On lance un dé non truqué. Si le dé donne un numéro inférieur ou égal à 2, on tire une boule dans l'urne 1. Sinon on tire une boule dans l'urne 2. On considère les événements B : « la boule tirée est blanche » et U_i : « la boule est tirée dans l'urne i », $i \in \{1,2\}$.

1) Calculer les probabilités $\mathbb{P}(U_1)$, $\mathbb{P}(U_2)$, $\mathbb{P}(B|U_1)$ et $\mathbb{P}(B|U_2)$.
2) Calculer la probabilité de tirer une boule blanche.
3) On suppose que la boule tirée est blanche. Calculer le probabilité qu'elle provienne de l'urne 1.

Exercice 3.14 On dispose de trois cartes dont l'une a deux faces vertes, la seconde a deux faces rouges et la troisième a une face verte et une face rouge. On choisit une carte au hasard et on choisit une face au hasard. Si le côté choisi est vert, quelle est la probabilité que l'autre côté soit aussi vert ? L'on pourrait répondre intuitivement $1/2$. Montrer que la bonne réponse est $2/3$.

Exercice 3.15 La proportion des pièces défectueuses dans un lot est de 0.05. Le contrôle de fabrication des pièces est tel que :

- Si la pièce est bonne, elle est acceptée avec une probabilité de 0.96.
- Si la pièce est mauvaise, elle est refusée avec probabilité 0.98.

On choisit une pièce au hasard et la contrôle. Probabilité pour :

1) qu'il y ait une erreur de contrôle ?
2) qu'une pièce acceptée soit mauvaise ?

Exercice 3.16 Un train se compose de 10 wagons citernes contenant un produit dangereux. Chacun des wagons peut avec une probabilité 0.1 (et indépendamment des autres) avoir un défaut. Avant le départ, les wagons sont examinés par un contrôleur. On admet que ce dernier peut déceler le défaut (s'il y en a un) d'un wagon donné avec une probabilité égale à 0.7. Un seul défaut suffit pour que le train soit retardé. Trouver les probabilités des événements suivants :

1) Le train est retardé.

2) Le train part avec au moins un wagon défectueux.

Chapitre 4

Généralités sur les variables aléatoires

Résumé – Ce chapitre présente la notion de variable aléatoire (v.a.) et ses caractéristiques fonctionnelles et numériques (densité, fonction de répartition, espérance, variance). L'objectif est de savoir faire des calculs relatifs à ces caractéristiques.

Sommaire

4.1 Généralités

Dans ce chapitre, on considère une expérience aléatoire dont l'espace fondamental est noté Ω. Une **variable aléatoire** X est une grandeur numérique inconnue dont la valeur dépend exclusivement du résultat de ladite expérience (d'où la terminologie « variable »). Avant de réaliser l'expérience, on connaît l'ensemble $X(\Omega)$ des valeurs possibles pour X mais on ne sait pas encore quelle valeur X va effectivement prendre lorsqu'on réalise l'expérience. Illustrons cela à l'aide d'un exemple.

Exemple 4.1 On jette 3 fois une pièce de monnaie. L'espace fondamental est

$$\Omega = \{\mathsf{P},\mathsf{F}\}^3 = \Big\{(\mathsf{P},\mathsf{P},\mathsf{P}),(\mathsf{P},\mathsf{P},\mathsf{F}),(\mathsf{P},\mathsf{F},\mathsf{P}),(\mathsf{P},\mathsf{F},\mathsf{F}), (\mathsf{F},\mathsf{P},\mathsf{P}),(\mathsf{F},\mathsf{P},\mathsf{F}),(\mathsf{F},\mathsf{F},\mathsf{P}),(\mathsf{F},\mathsf{F},\mathsf{F})\Big\},$$

où P et F désignent respectivement « Pile » et « Face ». On désigne par X le nombre de fois que P apparaît. L'ensemble des valeurs prises par X est

$$X(\Omega) = \{0,1,2,3\}.$$

Si l'expérience donne comme résultat $\omega = (\mathsf{P}, \mathsf{P}, \mathsf{F})$ par exemple, alors $X(\omega) = 2$.

Pour tout sous-ensemble $A \subset X(\Omega)$, l'ensemble

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

est appelé l'**image inverse** de A. Par exemple, en reprenant l'expérience de l'exemple 4.1, on a

$$X^{-1}(\{2\}) = \{(\mathsf{P}, \mathsf{P}, \mathsf{F}), (\mathsf{P}, \mathsf{F}, \mathsf{P}), (\mathsf{F}, \mathsf{P}, \mathsf{P})\}$$

qui correspond à l'événement « on obtient 2 fois Pile ».

Voici maintenant la définition générale d'une variable aléatoire.

Définition 4.2 Soient $(\Omega, \mathcal{A})$ et $(E, \mathcal{E})$ deux espaces probabilisables.

- Une *variable aléatoire* (v.a.) X à valeurs dans E est une application de Ω dans E telle que pour tout $A \in \mathcal{E}$, l'on a $X^{-1}(A) \in \mathcal{A}$, où l'ensemble

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

  est l'image inverse de A par l'application X.
- Lorsque $E = \mathbb{R}$, on parle de v.a. réelle (v.a.r.) et lorsque $E = \mathbb{R}^n$, on parle de vecteur aléatoire.

Théorème – Définition 4.3 *Soit X une v.a. réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un espace probabilisable $(E, \mathcal{E})$. Alors, l'application $\mathbb{P}_X$ définie par :*

$$\mathbb{P}_X : \mathcal{E} \longrightarrow [0,1], \quad A \longmapsto \mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)) \tag{4.1}$$

définit une probabilité sur $(E, \mathcal{E})$. Pour tous $A \subset X(\Omega)$ et $x \in X(\Omega)$, on note :

$$\mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}\left(\{\omega \in \Omega \mid X(\omega) \in A\}\right)$$
$$\mathbb{P}(X = x) = \mathbb{P}(X^{-1}(\{x\})) = \mathbb{P}\left(\{\omega \in \Omega \mid X(\omega) = x\}\right).$$

Exemple 4.1 (suite) On considère l'exemple 4.1. Calculons $\mathbb{P}(X = 2)$. L'événement $\{X = 2\}$ correspond à l'ensemble

$$X^{-1}(\{2\}) = \{(\mathsf{P}, \mathsf{P}, \mathsf{F}), (\mathsf{P}, \mathsf{F}, \mathsf{P}), (\mathsf{F}, \mathsf{P}, \mathsf{P})\}$$

d'où

$$\mathbb{P}(X = 2) = \frac{\mathrm{Card}(X = 2)}{\mathrm{Card}(\Omega)} = \frac{3}{8}.$$

4.2 Variables aléatoires discrètes et continues

Définition 4.4 Une v.a. X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ est dite

- *discrète* si l'ensemble $X(\Omega)$ des valeurs prises par X est un ensemble fini ou dénombrable (en bijection avec $\mathbb{N}$) de la forme $X(\Omega) = \{x_1, x_2, \ldots\}$.
- *continue* si $X(\Omega)$ est un intervalle ou une réunion d'intervalles.

Exemple 4.5 Comme exemples de v.a. discrètes, on peut citer :

- le nombre de fois que "Pile" apparaît dans 5 lancers d'une pièce ($X(\Omega) = \{0, \ldots, 5\}$),
- le nombre de lancers nécessaires pour obtenir "Face" pour la première fois ($X(\Omega) = \mathbb{N}^*$),
- le nombre d'accidents survenant dans une zone donnée ($X(\Omega) = \mathbb{N}$),
- etc. . .

Exemple 4.6 Comme exemples de v.a. continues on peut citer :

- la durée de vie d'un téléphone portable ($X(\Omega) = \mathbb{R}_+^*$),
- le temps passé dans une file d'attente devant un guichet ($X(\Omega) = \mathbb{R}_+^*$),
- le poids d'un individu ($X(\Omega) = \mathbb{R}_+^*$),
- la température ($X(\Omega) = \mathbb{R}$),
- etc. . .

4.2.1 Variables aléatoires discrètes

Définition 4.7 On appelle *loi de probabilité* de la v.a. discrète X, la donnée des probabilités $\mathbb{P}(X = x_k)$ pour toutes les valeurs $x_k \in X(\Omega)$ prises par X.

La loi de probabilité d'une v.a. discrète X à valeurs dans $\{x_1, x_2, \ldots\}$ est *bien définie* si et seulement elle vérifie les deux relations :

$$\begin{cases} \forall k \geqslant 1, \quad \mathbb{P}(X = x_k) \geqslant 0 \\ \\ \displaystyle\sum_{k \geqslant 1} \mathbb{P}(X = x_k) = 1. \end{cases} \tag{4.2}$$

Si X prend les valeurs $x_1, \ldots, x_n$ avec probabilités respectives p_1, …, p_n, la loi de X peut être résumée à l'aide d'un tableau comme suit :

x	x_1	x_2	$\cdots$	x_n
$\mathbb{P}(X = x)$	p_1	p_2	$\cdots$	p_n

Définition 4.8 On appelle *fonction de répartition* (f.r.) de la v.a. discrète X, la fonction définie de $\mathbb{R}$ dans $[0,1]$ par :

$$F_X(x) = \mathbb{P}(X \leqslant x). \tag{4.3}$$

Proposition 4.9 *La fonction de répartition F d'une v.a.r. X vérifie les propriétés suivantes :*

1) $\forall x \in \mathbb{R}, 0 \leqslant F(x) \leqslant 1.$

2) F est une fonction croissante (au sens large), continue à droite en tout point $x \in \mathbb{R}$.

3) $\lim\limits_{x \to -\infty} F(x) = 0$ *et* $\lim\limits_{x \to +\infty} F(x) = 1.$

Remarque 4.10

1) L'ensemble de définition de la f.r. est $\mathbb{R}$ et non $X(\Omega)$.
2) En particulier, pour une *v.a. discrète*, la fonction de répartition est une *fonction en escalier, constante par morceaux et continue à droite*.

La proposition suivante est très utile pour calculer la fonction de répartition d'une v.a. discrète.

Proposition 4.11 *Soit X une variable aléatoire discrète prenant un ensemble fini de valeurs $\{x_1, \ldots, x_n\}$ avec probabilités respectives $p_1, \ldots, p_n$ telles que*

$$\sum_{i=1}^{n} p_i = 1.$$

La fonction de répartition de X est obtenue par la formule :

$$F(x) = \begin{cases} 0 & si & x < x_1 \\ p_1 & si & x_1 \leqslant x < x_2 \\ p_1 + p_2 & si & x_2 \leqslant x < x_3 \\ \vdots & \vdots & \vdots \\ \sum_{j=1}^{i} p_j & si & x_i \leqslant x < x_{i+1} \\ \vdots & \vdots & \vdots \\ \sum_{j=1}^{n-1} p_j & si & x_{n-1} \leqslant x < x_n \\ 1 & si & x \geqslant x_n. \end{cases} \tag{4.4}$$

Exemple 4.1 (suite) On lance trois fois une pièce parfaitement équilibrée et on désigne par X le nombre de fois que "Pile" apparaît. Déterminons la loi de probabilité et la f.r. de X.

- On rappelle que l'ensemble des valeurs de X est $X(\Omega) = \{0, 1, 2, 3\}$.
- On a :
  - $X^{-1}(\{0\}) = \{(\mathsf{F}, \mathsf{F}, \mathsf{F})\}$ donc $\mathbb{P}(X = 0) = \frac{\text{Card}(X=0)}{\text{Card}(\Omega)} = \frac{1}{8}$.

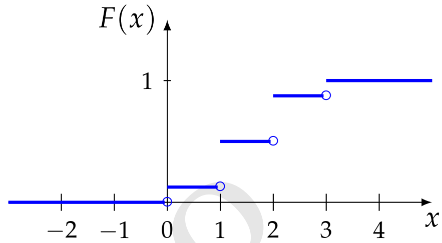
- $X^{-1}(\{1\}) = \{(\mathsf{P},\mathsf{F},\mathsf{F}),(\mathsf{F},\mathsf{P},\mathsf{F}),(\mathsf{F},\mathsf{F},\mathsf{P})\}$ donc $\mathbb{P}(X=1) = \frac{\text{Card}(X=1)}{\text{Card}(\Omega)} = \frac{3}{8}$.
- $X^{-1}(\{2\}) = \{(\mathsf{P},\mathsf{P},\mathsf{F}),(\mathsf{P},\mathsf{F},\mathsf{P}),(\mathsf{F},\mathsf{P},\mathsf{P})\}$ donc $\mathbb{P}(X=2) = \frac{\text{Card}(X=2)}{\text{Card}(\Omega)} = \frac{3}{8}$.
- $X^{-1}(\{3\}) = \{(\mathsf{P},\mathsf{P},\mathsf{P})\}$ donc $\mathbb{P}(X=3) = \frac{\text{Card}(X=3)}{\text{Card}(\Omega)} = \frac{1}{8}$.

- La loi de X est donc donnée par le tableau suivant :

x_k	0	1	2	3
$\mathbb{P}(X = x_k)$	1/8	3/8	3/8	1/8

- Fonction de répartition de X :

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 1/8 & \text{si} \quad 0 \leqslant x < 1 \\ 1/2 & \text{si} \quad 1 \leqslant x < 2 \\ 7/8 & \text{si} \quad 2 \leqslant x < 3 \\ 1 & \text{si} \quad x \geqslant 3. \end{cases}$$



4.2.2 Variables aléatoires continues

Définition 4.12 Une v.a. X est dite *absolument continue* (ou à densité) s'il existe une fonction f définie sur $\mathbb{R}$, **positive**, **continue sauf peut-être en un nombre fini de points**, telle que

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1 \tag{4.5}$$

et pour tout intervalle $B \subset \mathbb{R}$, on a :

$$\mathbb{P}(X \in B) = \int_B f(x)\,\mathrm{d}x. \tag{4.6}$$

La fonction f est appelée la **densité** de la loi de X.

Proposition 4.13 *Soit X une v.a. continue de densité f. Alors :*

1) *La f.r. de X est définie par :*

$$\forall x \in \mathbb{R}, \quad F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t. \tag{4.7}$$

2) *La densité de probabilité est la dérivée de la f.r. c'est-à-dire $F'(x) = f(x)$ en tout point x où f est continue.*

3) *Pour tous réels a et b tels que $a \leqslant b$, on a :*

$$\mathbb{P}(a < X \leqslant b) = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x. \tag{4.8}$$

4) *La probabilité d'un point est toujours nulle : $\forall x \in \mathbb{R},\ \mathbb{P}(X = x) = 0$.*

Définition 4.14 (Quantiles) Soit X une v.a. à densité dont la fonction de répartition est notée F. Pour tout $\alpha \in]0,1[$, on appelle **quantile** d'ordre α, le réel q_α tel que

$$F(q_\alpha) = \alpha \qquad \text{ou encore} \qquad \mathbb{P}(X \leqslant q_\alpha) = \alpha. \tag{4.9}$$

La fonction quantile est l'inverse de la fonction de répartition.

Remarque 4.15 Certains quantiles portent des noms particuliers. Ainsi,

- le quantile d'ordre $\alpha = 1/2$ est appelé la *médiane* ;
- les quantiles d'ordre $1/4$ et $3/4$ sont appelés respectivement *premier quartile* et *troisième quartile* ;
- les quantiles d'ordre $k/10$ pour $k = 1, \ldots, 9$ sont appelés *déciles*.

Application 4.16 Soit f la fonction définie de $\mathbb{R}$ dans $\mathbb{R}$ par :

$$f(x) = \frac{1}{2}\mathbb{I}_{[1,3]}(x) = \begin{cases} 1/2 & \text{si} \quad x \in [1,3] \\ 0 & \text{sinon,} \end{cases}$$

où $\mathbb{I}_{[1,3]}$ est la fonction indicatrice de $[1,3]$ ($\mathbb{I}_{[1,3]}(x) = 1$ si $x \in [1,3]$ et $\mathbb{I}_{[1,3]}(x) = 0$ sinon).

1) Vérifier que f est bien une densité de probabilité.
2) Soit X la v.a. dont f est la densité de probabilité. Déterminer la f.r. F de X et la représenter.
3) Calculer $\mathbb{P}(2 \leqslant X \leqslant 3)$ et $\mathbb{P}(X > 2)$.
4) Déterminer le quantile d'ordre 0.75 de la loi de X.

4.3 Moments d'une variable aléatoire

Les moments d'une variable aléatoire sont des caractéristiques numériques qui donnent des renseignements sur la loi de cette variable aléatoire (sans pour autant la déterminer complètement). Les plus couramment utilisés sont l'**espérance mathématique** (ou **moyenne**) et la **variance**.

4.3.1 Espérance mathématique

Définition 4.17 Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une v.a. réelle définie sur Ω. L'*espérance mathématique* (ou tout simplement *espérance*) ou encore *moyenne* de X, si elle existe, est notée $\mathbb{E}(X)$ et définie comme suit :

1) Si X est discrète et prend les valeurs $x_1, x_2, \ldots$ avec probabilités respectives $p_1, p_2, \ldots$,

$$\mathbb{E}(X) = \sum_{k \geqslant 1} x_k p_k \tag{4.10}$$

2) Si X est continue de densité f,

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x. \tag{4.11}$$

Définition 4.18 Une v.a.r. X est dite *centrée* si $\mathbb{E}(X) = 0$.

Proposition 4.19

1) L'espérance d'une v.a. réelle est une constante réelle.

2) L'espérance mathématique est linéaire c'est-à-dire, si a désigne une constante réelle et si X, $X_1, \ldots, X_n$ désignent des v.a., on a :

   (a) $\mathbb{E}(X + a) = \mathbb{E}(X) + a$ *;*

   (b) $\mathbb{E}(aX) = a\mathbb{E}(X)$ *;*

   (c) $\mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)$.

3) L'espérance conserve le signe et l'ordre c'est-à-dire, si X et Y sont des variables aléatoires, alors

   (a) si $X \geqslant 0$, alors $\mathbb{E}(X) \geqslant 0$;

   (b) si $X \leqslant Y$, alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

Dans la pratique, il arrive qu'on ait besoin de calculer $\mathbb{E}(h(X))$ où $h(X)$ est donnée en fonction de X. Le résultat suivant permet de résoudre un tel problème.

Théorème 4.20 *Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Alors pour toute fonction $h : \mathbb{R} \longrightarrow \mathbb{R}$, l'espérance mathématique de $h(X) = h \circ X$, si elle existe, est donnée par la règle suivante :*

1) Si X est discrète et prend les valeurs x_1, x_2, ... avec probabilités respectives p_1, p_2, ...,

$$\mathbb{E}(h(X)) = \sum_{k \geqslant 1} h(x_k) p_k \tag{4.12}$$

2) Si X est continue de densité f,

$$\mathbb{E}(h(X)) = \int_{-\infty}^{+\infty} h(x) f(x)\,\mathrm{d}x. \tag{4.13}$$

Exemple 4.21

1) Si X est discrète et prend les valeurs $x_1, x_2, \ldots, x_n$ avec probabilités respectives p_1, p_2, ..., p_n,

$$\mathbb{E}(X^2) = \sum_{k=1}^{n} (x_k)^2 p_k \tag{4.14}$$

2) Si X est continue de densité f,

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x. \tag{4.15}$$

4.3.2 Variance

Définition 4.22 Soit X une v.a. telle que $\mathbb{E}(X^2)$ existe. La *variance* de X est définie par :

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] \tag{4.16}$$

et le réel positif $\sigma(X) = \sqrt{\mathbb{V}\mathrm{ar}(X)}$ est appelé *écart-type* de X.

Dans la pratique, on utilise plutôt le théorème suivant :

Théorème 4.23 (König-Huygens) *Soit X une v.a.r. Si $\mathbb{V}\mathrm{ar}(X)$ existe, alors*

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2. \tag{4.17}$$

Proposition 4.24 *Soit X une v.a.r telle que $\mathbb{V}\mathrm{ar}(X)$ existe. Alors :*

1) $\mathbb{V}\mathrm{ar}(X) \geqslant 0$.

2) Si a et b sont des constantes alors $\mathbb{V}\mathrm{ar}(aX + b) = a^2\mathbb{V}\mathrm{ar}(X)$.

Remarque 4.25 La variance peut-être utilisée comme un indicateur de dispersion. Plus elle est élevée, plus les valeurs sont dispersées autour de la valeur moyenne.

Application 4.26 Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$ pour la v.a. X dont la loi est définie comme suit :

x_k	0	1	2	3
$\mathbb{P}(X = x_k)$	1/8	3/8	3/8	1/8

Application 4.27 On considère la v.a. continue X de l'exemple 4.16. Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.

4.4 Loi d'une fonction d'une variable aléatoire $Y = h(X)$

Etant donné une variable aléatoire X dont la loi est connue, on peut avoir besoin de déterminer la loi de la v.a. $Y = h(X)$ pour une certaine fonction $h : \mathbb{R} \longrightarrow \mathbb{R}$.

4.4.1 Cas d'une v.a. discrète

Pour trouver la loi de Y, on identifie $Y(\Omega)$ et on détermine $\mathbb{P}(Y = y) = \mathbb{P}(h(X) = y)$ pour tout $y \in Y(\Omega)$.

Exemple 4.28 Soit X une v.a.r. dont la loi est définie par le tableau suivant :

k	-2	-1	0	2
$\mathbb{P}(X=k)$	$1/15$	$2/15$	$8/15$	$4/15$

Déterminons la loi de $Y = X^2$.

On a : $Y(\Omega) = \{0, 1, 4\}$ et

$$\begin{aligned}
\mathbb{P}(Y=0) &= \mathbb{P}(X^2=0) = \mathbb{P}(X=0) = 8/15 \\
\mathbb{P}(Y=1) &= \mathbb{P}(X^2=1) = \mathbb{P}(X=-1 \text{ ou } X=1) = \mathbb{P}(X=-1) + \mathbb{P}(X=1) \\
&= 2/15 + 0 = 2/15 \\
\mathbb{P}(Y=4) &= \mathbb{P}(X^2=4) = \mathbb{P}(X=-2 \text{ ou } X=2) = \mathbb{P}(X=-2) + \mathbb{P}(X=2) \\
&= 1/15 + 4/15 = 5/15
\end{aligned}$$

La loi de Y peut être résumée dans le tableau suivant :

k	0	1	4
$\mathbb{P}(Y=k)$	$8/15$	$2/15$	$5/15$

4.4.2 Cas d'une v.a. continue

Soient F_X la f.r. de X et f_X sa densité. Pour trouver la loi de $Y = h(X)$, on peut procéder comme suit :

- déterminer la fonction de répartition F_Y de Y définie, pour tout $y \in \mathbb{R}$, par
$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(h(X) \leqslant y);$$
- déterminer la densité f_Y de Y en dérivant F_Y.

Exemple 4.29 Soient U la v.a. dont la densité est définie sur $\mathbb{R}$ par

$$f_U(u) = \begin{cases} 1 & \text{si} \quad u \in [0,1] \\ 0 & \text{sinon} \end{cases}$$

et soit $\lambda \in \mathbb{R}_+^*$. Déterminons la densité de la loi de la v.a.

$$X = -\frac{\ln U}{\lambda}.$$

Soient F_X et F_Y les f.r. respectives de U et X et f_X la densité de la loi de X. Pour tout réel x, on a

$$\begin{aligned}
F_X(x) = \mathbb{P}(X \leqslant x) &= \mathbb{P}\left(-\frac{\ln U}{\lambda} \leqslant x\right) = \mathbb{P}\left(-\ln U \leqslant \lambda x\right) \\
&= \mathbb{P}\left(\ln U \geqslant -\lambda x\right) \\
&= \mathbb{P}\left(U \geqslant \mathrm{e}^{-\lambda x}\right) \\
&= 1 - F_U(\mathrm{e}^{-\lambda x}).
\end{aligned}$$

Après dérivation, on obtient

$$\begin{aligned} f_X(x) = \frac{\mathrm{d}F_U}{\mathrm{d}x} &= \lambda \mathrm{e}^{-\lambda x} F'_U(\mathrm{e}^{-\lambda x}) = \lambda \mathrm{e}^{-\lambda x} f_U(\mathrm{e}^{-\lambda x}) \\ &= \begin{cases} \lambda \mathrm{e}^{-\lambda x} & \text{si} \quad \mathrm{e}^{-\lambda x} \in [0, 1] \\ 0 & \text{sinon.} \end{cases} \end{aligned}$$

Or, $0 \leqslant \mathrm{e}^{-\lambda x} \leqslant 1 \iff -\infty < -\lambda x \leqslant 0 \iff 0 \leqslant x < +\infty$. Ainsi, la v.a. X suit la loi de densité

$$f_X(x) = \lambda \mathrm{e}^{-\lambda x} \mathbb{I}_{\mathbb{R}_+}(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & \text{si} \quad x \in [0, +\infty[\\ 0 & \text{sinon.} \end{cases}$$

4.5 EXERCICES

Exercice 4.1 On considère un dé cubique truqué dont les faces sont numérotés de 1 à 6 et on note X la variable aléatoire donnée par le numéro de la face supérieure du dé. On suppose que le dé est truqué de sorte que la probabilité d'obtenir une face est proportionnelle au numéro inscrit sur cette face c'est-à-dire qu'il existe un réel a tel que $\mathbb{P}(X = k) = ka$ pour tout $k \in \{1, \ldots, 6\}$. Déterminer la loi de X et calculer $\mathbb{E}(X)$.

Exercice 4.2 On lance simultanément deux dés bien équilibrés dont les faces sont numérotées de 1 à 6. On note X la valeur absolue de la différence des nombres portés sur les faces supérieures.

1) Quelle est la loi de probabilité de X ?
2) Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.

Exercice 4.3 On lance simultanément deux dés bien équilibrés dont les faces sont numérotées de 1 à 6. On note X le maximum des chiffres obtenus sur les faces supérieures. Déterminer la loi de probabilité de X.

Exercice 4.4 Une urne contient 3 jetons blancs (1 carré et 2 ronds) et 4 jetons noirs (3 carrés et 1 rond). On tire au hasard et simultanément 3 jetons de l'urne. Soit X la v.a. égale au nombre de jetons blancs obtenus lors d'un tirage. Déterminer la loi de X et calculer $\mathbb{E}(X)$.

Exercice 4.5 Une urne contient une boule qui porte le numéro 0, deux qui portent le numéro 1, trois qui portent le numéro 2 et quatre qui portent le numéro 3. On extrait simultanément deux boules de cette urne et on désigne par X la somme des numéros obtenus.

1) Déterminer la loi de X.
2) Calculer son espérance et sa variance.

Exercice 4.6 Quatre personnes A, B, C et D se présentent à un entretien après lequel le recruteur a la possibilité soit de ne choisir personne, soit d'en choisir une, deux, trois ou quatre. On suppose que pour chaque personne, les chances d'être recruté ou non sont les mêmes. Soient Ω l'univers associé à cette expérience et X la variable aléatoire égale au nombre des personnes recrutées. Déterminer la loi de X et calculer $\mathbb{E}(X)$.

Exercice 4.7 Dans un rayon de magasin, il y a deux produits A et B (par exemple une table et une chaise). La probabilité pour qu'un client achète le produit A est 0.3. La probabilité pour qu'il achète le produit B quand il a acheté le produit A est 0.8 et la probabilité qu'il achète le produit B quand il n'a pas acheté le produit A est 0.1. Les produits A et B sont respectivement vendus à 10 000 F et 4 000 F. Soit X la v.a. égale à la dépense du client. Déterminer la loi de X puis calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.

Exercice 4.8 Dans une loterie, il y a 100 billets de 500 F et 2 lots : l'un de 5 000 F et l'autre de 10 000 F. Les numéros des deux billets gagnants sont tirés au sort indépendamment l'un de l'autre (le même billet peut donc gagner les deux lots). Soit X la v.a. représentant la somme d'argent gagnée (ou perdue) par une personne qui achète un seul billet. On représentera les pertes par des nombres négatifs (par exemple si la personne ne gagne aucun lot, elle perd 500 F, on aura alors $X = -500$). Déterminer la loi de X et calculer $\mathbb{E}(X)$.

Exercice 4.9 Soit X une v.a.r prenant les valeurs -4, -2, 0, 2 et 4 avec les mêmes probabilités. Déterminer la loi de $|X|$ et ses caractéristiques numériques (espérance mathématique, variance et écart-type).

Exercice 4.10 Soit f la fonction de $\mathbb{R}$ vers $\mathbb{R}$ définie pour tout $x \in \mathbb{R}$ par :

$$f(x) = \begin{cases} kx & \text{si } x \in [0, 3] \\ 0 & \text{sinon.} \end{cases}$$

1) Déterminer k pour que f soit la densité de probabilité d'une variable aléatoire X.
2) Déterminer la fonction de répartition de X.
3) Calculer $\mathbb{P}(X > 1)$, $\mathbb{P}(X < 2)$ et $\mathbb{P}(1 < X < 2)$.
4) Calculer l'espérance et la variance de X.
5) Déterminer le quantile d'ordre 0.25 de la loi de X.

Exercice 4.11 Soit f la fonction définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} kx(1-x) & \text{si } \quad 0 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

1) Déterminer k pour que f soit une densité de probabilité.
2) Soit X la v.a.r. dont la densité de probabilité est f. Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.
3) Déterminer la fonction de répartition $F(x)$ de X.
4) Calculer : $\mathbb{P}(X \leqslant \frac{1}{2})$ et $\mathbb{P}(\frac{1}{4} \leqslant X \leqslant \frac{1}{2})$.

Exercice 4.12 Soit f la fonction définie de $\mathbb{R}$ dans $\mathbb{R}$ par $f(x) = \lambda e^{-|x|}$ où λ est un réel.

1) Déterminer λ pour que f soit la densité de probabilité d'une v.a. X (on dit que X suit la loi de Laplace[1] standard).
2) Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.

1. Pierre-Simon de Laplace (1749-1827), mathématicien, astronome, physicien et homme politique français.

CHAPITRE 5

LOIS DISCRÈTES ET CONTINUES USUELLES

Résumé – Ce chapitre présente les lois de probabilité discrètes et continues souvent utilisées dans la pratique. L'objectif est de savoir faire des calculs en utilisant les lois de probabilité usuelles.

Sommaire

Dans tout le chapitre, on utilise le symbole $\rightsquigarrow$ pour exprimer le fait qu'une v.a. suive une certaine loi de probabilité.

5.1 LOIS DISCRÈTES USUELLES

5.1.1 Loi uniforme discrète

Pour tout $n \in \mathbb{N}^*$, on note $[\![1,n]\!] = \{1,\ldots,n\}$.

Définition 5.1 Soit $n \in \mathbb{N}^*$. Une variable aléatoire X suit une loi uniforme discrète sur $[\![1,n]\!]$ si X prend ses valeurs dans l'ensemble $[\![1,n]\!] = \{1,\ldots,n\}$ et si

$$\forall k \in \{1,\ldots,n\}, \quad \mathbb{P}(X=k) = \frac{1}{n}. \tag{5.1}$$

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.2 *Si X suit une loi uniforme discrète sur $[1, n]$, alors*

$$\mathbb{E}(X) = \frac{n+1}{2} \quad et \quad \mathbb{V}\mathrm{ar}(X) = \frac{n^2-1}{12}. \tag{5.2}$$

Remarque 5.3 La loi uniforme discrète intervient dans de nombreux domaines comme les jeux de pile ou face ou les jeux de dés (avec une pièce ou un dé parfaitement équilibré(e)), les jeux de cartes, etc...

5.1.2 Loi de Bernoulli

Définition 5.4 Soit $p \in [0, 1]$. Une variable aléatoire X suit une loi de Bernoulli de paramètre p et on note $X \rightsquigarrow \mathcal{B}(1, p)$ si la v.a. X ne peut prendre que deux valeurs 1 et 0 avec probabilités respectives p et $q = 1 - p$ c'est-à-dire $X(\Omega) = \{0, 1\}$ et

$$\mathbb{P}(X = 1) = p \quad \text{et} \quad \mathbb{P}(X = 0) = q = 1 - p. \tag{5.3}$$

L'espérance mathématique et la variance sont données par la proposition suivante.

Proposition 5.5 *Si $X \rightsquigarrow \mathcal{B}(1, p)$, alors*

$$\mathbb{E}(X) = p \quad et \quad \mathbb{V}\mathrm{ar}(X) = p(1-p) = pq. \tag{5.4}$$

Remarque 5.6 La loi de Bernoulli $\mathcal{B}(1, p)$ est souvent utilisée pour modéliser une *épreuve de Bernoulli* c'est-à-dire une expérience aléatoire dont le résultat ne peut prendre que deux valeurs appelées, par convention, *succès* et *échec*. Dans ce cas, la v.a. X associée prend la valeur 1 pour le succès et la valeur 0 pour l'échec avec probabilités respectives p et $q = 1 - p$.

5.1.3 Loi binomiale

Considérons une expérience aléatoire n'ayant que deux résultats possibles : le succès (avec probabilité $p \in [0, 1]$) et l'échec (avec probabilité $q = 1 - p$). Supposons que l'on répète cette expérience n fois de façon indépendante et soit X le nombre de succès obtenus au cours des n répétitions de l'expérience. Il est clair que la v.a. X prend ses valeurs dans l'ensemble $X(\Omega) = \{0, 1, \ldots, n\}$. Par définition, la v.a. X suit une **loi binomiale** de paramètres n et $p \in [0, 1]$, notée $\mathcal{B}(n, p)$ et on a :

$$\forall k \in \{0, 1, \ldots, n\}, \quad \mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k} \tag{5.5}$$

où

$$C_n^k = \frac{n!}{k!(n-k)!}.$$

La quantité $\mathbb{P}(X = k)$ représente la probabilité d'obtenir k succès au cours des n répétitions.

Application 5.7 Vous avez un devoir de Mathématiques à rendre et vous contactez 10 de vos amis les plus sincères pour vous aider. Chacun, indépendamment des autres a une probabilité de 0.2 d'accepter de vous aider. A la fin de tous vos 10 coups de fil, la v.a. X du nombre total d'amis qui acceptent vous aider suit une loi binomiale $\mathcal{B}(10, 0.2)$. En effet, il s'agit de $n = 10$ répétitions indépendantes de l'épreuve de Bernoulli « appeler un ami et demander son aide » (le *succès* étant observé lorsque l'ami accepte et l'*échec* sinon). Quel est la probabilité que vous :

1) ne trouviez aucun ami pour vous aider ?

2) trouviez au moins un ami pour vous aider ?

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.8 *Si* $X \rightsquigarrow \mathcal{B}(n, p)$, *alors*

$$\mathbb{E}(X) = np \quad \textit{et} \quad \mathbb{V}\mathrm{ar}(X) = np(1-p) = npq. \tag{5.6}$$

5.1.4 Loi géométrique

Considérons une expérience aléatoire n'ayant que deux résultats possibles : le succès (avec probabilité $p \in [0, 1]$) et l'échec (avec probabilité $q = 1 - p$). Répétons cette expérience plusieurs fois de façon indépendante et désignons par X le **nombre de répétitions nécessaires pour obtenir le premier succès**. Il est évident que la v.a. X prend ses valeurs dans l'ensemble $X(\Omega) = \mathbb{N}^*$. Par définition, la v.a. X suit une **loi géométrique** de paramètre $p \in [0, 1]$, notée $\mathcal{G}(p)$ et on a :

$$\forall k \in \mathbb{N}^*, \quad \mathbb{P}(X = k) = (1-p)^{k-1} p. \tag{5.7}$$

Cette quantité représente la probabilité d'obtenir $k - 1$ échecs avant d'obtenir le succès à la $k^{\text{ème}}$ répétition.

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.9 *Si X suit une loi* $\mathcal{G}(p)$, *alors*

$$\mathbb{E}(X) = \frac{1}{p} \quad \textit{et} \quad \mathbb{V}\mathrm{ar}(X) = \frac{1-p}{p^2}. \tag{5.8}$$

5.1.5 Loi de Poisson

Définition 5.10 La **loi de Poisson** de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$, est la loi d'une v.a. discrète prenant ses valeurs dans $\mathbb{N}$ avec les probabilités :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = \frac{\mathrm{e}^{-\lambda} \lambda^k}{k!}. \tag{5.9}$$

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.11 *Si $X \rightsquigarrow \mathcal{P}(\lambda)$, alors*

$$\mathbb{E}(X) = \mathbb{V}\mathrm{ar}(X) = \lambda. \tag{5.10}$$

Remarque 5.12 La loi de Poisson est utilisée pour décrire les « événements rares » comme le nombre d'appels reçus par un standard téléphonique pendant une période donnée, le nombre de suicides par an dans un pays donné, le nombre de pièces défectueuses dans une livraison importante (la production étant de bonne qualité), le nombre d'accidents dans un atelier, etc. . . .

5.1.6 Loi hypergéométrique

On considère une population de N individus dont M vérifient une certaine propriété. On choisit n individus au hasard dans cette population et on s'intéresse au nombre X d'individus dans l'échantillon vérifiant la propriété. La probabilité d'avoir exactement k individus vérifiant la propriété est :

$$\mathbb{P}(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}. \tag{5.11}$$

Les valeurs extrêmes de k sont :

$$k_{\min} = \max(0, n - N + M) \qquad \text{et} \qquad k_{\max} = \min(n, M).$$

La variable aléatoire X ainsi définie suit une **loi hypergéométrique** dépendant de trois paramètres N, n et p notée $\mathcal{H}(N, n, p)$, où $p = M/N$ désigne la proportion d'individus vérifiant la propriété dans la population entière.

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.13 *Si X suit une loi hypergéométrique $\mathcal{H}(N, n, p)$, alors*

$$\mathbb{E}(X) = np \qquad \textit{et} \qquad \mathbb{V}\mathrm{ar}(X) = \frac{N-n}{N-1} np(1-p). \tag{5.12}$$

5.2 LOIS CONTINUES USUELLES

5.2.1 Loi uniforme continue $\mathcal{U}[a, b]$

Définition 5.14 Soient a et b deux réels tels que $a < b$. Une variable aléatoire X à valeurs dans $[a, b]$ suit une loi uniforme continue sur $[a, b]$, notée $\mathcal{U}[a, b]$, si X est une v.a. continue et admet pour densité de probabilité, la fonction (voir Figure 5.1) définie par :

$$f(x) = \frac{1}{b-a} \mathbb{I}_{[a,b]}(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon.} \end{cases} \tag{5.13}$$

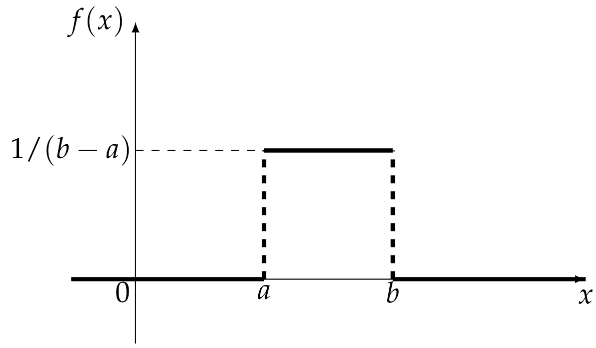


FIGURE 5.1 – Densité de la loi uniforme $\mathcal{U}[a,b]$.

Remarque 5.15 La loi uniforme définie par sa densité

$$f(x) = \begin{cases} 1 & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases} \tag{5.14}$$

est appelée *loi uniforme standard*.

La loi uniforme sur $[0,1]$ et la loi uniforme sur $[a,b]$ sont liées par les relations suivantes.

Proposition 5.16
1) Si $X \rightsquigarrow \mathcal{U}[a,b]$, alors $U = \frac{X-a}{b-a} \rightsquigarrow \mathcal{U}[0,1]$.
2) Si $U \rightsquigarrow \mathcal{U}[0,1]$, alors $X = (b-a)U + a \rightsquigarrow \mathcal{U}[a,b]$.

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.17 *Si X est une variable aléatoire de loi uniforme $\mathcal{U}[a,b]$, alors*

$$\mathbb{E}(X) = \frac{a+b}{2} \qquad et \qquad \mathbb{V}\mathrm{ar}(X) = \frac{(b-a)^2}{12}. \tag{5.15}$$

La fonction de répartition (f.r.) de la loi uniforme $\mathcal{U}[a,b]$ est donnée par la proposition suivante.

Proposition 5.18 *Si $X \rightsquigarrow \mathcal{U}[a,b]$, alors la f.r. est définie par :*

$$F(x) = \begin{cases} 0 & si \quad x \leqslant a \\ \dfrac{x-a}{b-a} & si \quad x \in [a,b] \\ 1 & si \quad x \geqslant b. \end{cases} \tag{5.16}$$

5.2.2 Loi exponentielle $\mathcal{E}(\lambda)$

Définition 5.19 Une variable aléatoire X à valeurs dans $[0,+\infty[$ suit une **loi exponentielle** de paramètre $\lambda > 0$, notée $\mathcal{E}(\lambda)$, si X est une v.a. continue et admet pour densité de probabilité la fonction (voir Figure 5.2) :

$$f(x) = \lambda \mathrm{e}^{-\lambda x} \mathbb{I}_{\mathbb{R}_+}(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0. \end{cases} \tag{5.17}$$

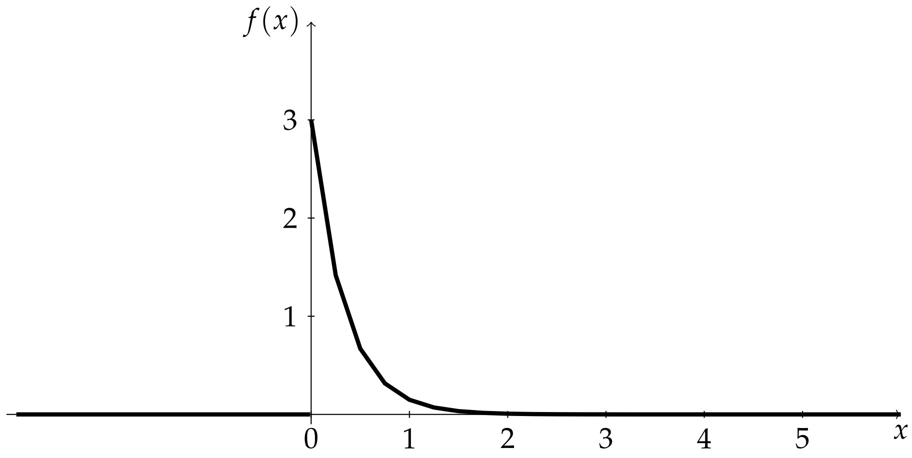


FIGURE 5.2 – Densité de la loi exponentielle $\mathcal{E}(\lambda)$ pour $\lambda = 3$.

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.20 *Si X est une variable aléatoire de loi exponentielle $\mathcal{E}(\lambda)$, alors*

$$\mathbb{E}(X) = \frac{1}{\lambda} \qquad et \qquad \mathbb{V}\mathrm{ar}(X) = \frac{1}{\lambda^2}. \tag{5.18}$$

La f.r. de la loi exponentielle est donnée par la proposition suivante.

Proposition 5.21 *Si $X \rightsquigarrow \mathcal{E}(\lambda)$, alors la f.r. est définie par :*

$$F(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x} & si \quad x \geqslant 0 \\ 0 & si \quad x < 0. \end{cases} \tag{5.19}$$

5.2.3 Loi normale $\mathcal{N}(m, \sigma)$

5.2.3.1 Définition et propriétés

Définition 5.22 Une v.a.r. X, prenant ses valeurs dans $\mathbb{R}$, suit la loi **normale** (ou **loi de Gauss**) de paramètres $m \in \mathbb{R}$ et $\sigma > 0$, notée $\mathcal{N}(m, \sigma)$, si sa densité de probabilité est donnée par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \, \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}. \tag{5.20}$$

Définition 5.23 La loi $\mathcal{N}(0, 1)$ est appelée **loi normale centrée réduite** ou **loi normale standard**. Sa densité (voir Figure 5.3) est souvent notée φ :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-x^2/2}. \tag{5.21}$$

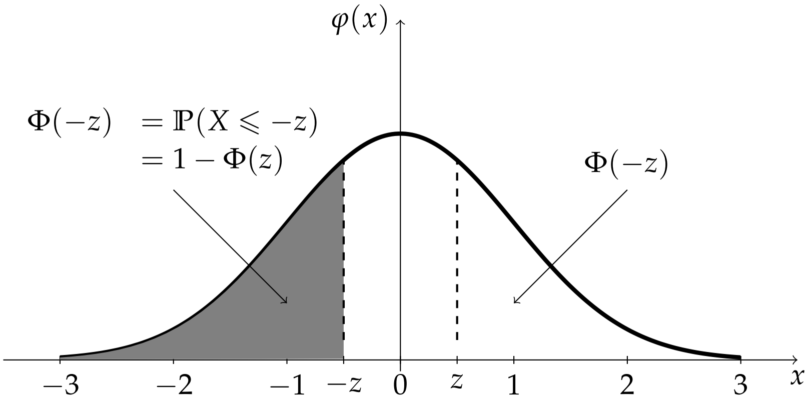


FIGURE 5.3 – Densité de la loi $\mathcal{N}(0,1)$. La zone coloriée correspond à $\Phi(-z)$.

Remarque 5.24 La courbe représentant la densité de probabilité d'une variable normale (ou gaussienne) $\mathcal{N}(m,\sigma)$ a un axe de symétrie vertical pour $x = m$ et du fait de sa forme, elle est souvent appelée « courbe en cloche ».

L'espérance mathématique et la variance sont données par la proposition suivante :

Proposition 5.25 *Si* $X \rightsquigarrow \mathcal{N}(m,\sigma)$, *alors* $\mathbb{E}(X) = m$ *et* $\mathbb{V}\text{ar}(X) = \sigma^2$.

5.2.3.2 Fonction de répartition de la loi $\mathcal{N}(m,\sigma)$

Définition 5.26 La fonction de répartition (f.r.) de la loi $\mathcal{N}(m,\sigma)$ est définie par :

$$F(x) = \mathbb{P}(X \leqslant x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

pour tout $x \in \mathbb{R}$. En particulier, la f.r. de la loi $\mathcal{N}(0,1)$, généralement notée Φ, est définie par

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{t^2}{2}} dt.$$

pour tout $z \in \mathbb{R}$.

La f.r. de la loi $\mathcal{N}(m,\sigma)$ n'a pas d'expression analytique exacte. Pour contourner ce problème, on utilise la proposition suivante qui montre que les valeurs de la f.r. d'une loi normale quelconque peuvent être obtenues à partir de celles de la loi normale centrée réduite.

Proposition 5.27 *Soit X une v.a. suivant la loi* $\mathcal{N}(m,\sigma)$. *Alors, la v.a. Z définie par*

$$Z = \frac{X - m}{\sigma} \tag{5.22}$$

et appelée variable centrée réduite associée *à X, suit la loi* $\mathcal{N}(0,1)$.

Cette propriété permet de ramener tous les calculs liés à la fonction de répartition d'une loi normale quelconque à celle de la loi normale centrée réduite. En effet, supposons que X suive une loi normale $\mathcal{N}(m, \sigma)$ et que nous voulions calculer $F(x) = \mathbb{P}(X < x)$. Soit

$$Z = \frac{X - m}{\sigma}$$

la variable centrée réduite associée à X. On a :

$$\mathbb{P}(X < x) = \mathbb{P}\left(\frac{X - m}{\sigma} < \frac{x - m}{\sigma}\right) = \mathbb{P}\left(Z < \frac{x - m}{\sigma}\right) = \Phi\left(\frac{x - m}{\sigma}\right). \tag{5.23}$$

Il ne reste donc plus qu'à savoir calculer la f.r. de la loi $\mathcal{N}(0, 1)$. Dans ce cas, on ne dispose pas de formule analytique exacte et l'on a plutôt recours à une **table statistique** (voir page 118).

5.2.3.3 Fonction de répartition de la loi $\mathcal{N}(0, 1)$

La table de la fonction de répartition de la loi $\mathcal{N}(0, 1)$ est donnée dans l'annexe A (voir page 118).

Exemple 5.28 Pour lire $\Phi(1.96)$, on fait la décomposition $1.96 = 1.90 + 0.06$. On repère, dans la première colonne, la ligne où se trouve 1.90 et, dans la première ligne, la colonne où se trouve 0.06 puis on lit la valeur se trouvant à l'intersection de la ligne correspondant à 1.90 et de la colonne correspondant à 0.06. On trouve $\Phi(1.96) = 0.975$.

On peut remarquer que seules les valeurs $\Phi(z) = \mathbb{P}(X < z)$ avec u positif y figurent. On montre que la fonction de répartition de la loi $\mathcal{N}(0, 1)$ vérifie la proposition suivante.

Proposition 5.29 *La f.r. de la loi $\mathcal{N}(0, 1)$ vérifie la propriété suivante :*

$$\forall z \in \mathbb{R}, \quad \Phi(-z) = 1 - \Phi(z) \tag{5.24}$$

et en particulier $\Phi(0) = 1/2$.

En résumé,

- pour $z \geqslant 0$, la valeur de $\Phi(z)$ est donnée par la table.
- pour $z < 0$, on a $-z > 0$ et $\Phi(z) = 1 - \Phi(-z)$.

5.2.3.4 Calcul de $\mathbb{P}(a < X < b)$ pour la loi $\mathcal{N}(0, 1)$

Pour tous a et $b \in \mathbb{R}$ tels que $a < b$, on a :

$$\mathbb{P}(X < b) = \Phi(b) \tag{5.25}$$
$$\mathbb{P}(X > a) = 1 - \Phi(a) \tag{5.26}$$
$$\mathbb{P}(a < X < b) = \Phi(b) - \Phi(a). \tag{5.27}$$

5.2.3.5 Calcul de $\mathbb{P}(a < X < b)$ pour la loi $\mathcal{N}(m, \sigma)$

Pour une v.a. $X \rightsquigarrow \mathcal{N}(m, \sigma)$, on sait que

$$Z = \frac{X - m}{\sigma} \rightsquigarrow \mathcal{N}(0, 1).$$

Pour tous a et $b \in \mathbb{R}$ tels que $a \leqslant b$, on a :

$$\begin{aligned}
\mathbb{P}(X < b) &= \mathbb{P}\left(\frac{X - m}{\sigma} < \frac{b - m}{\sigma}\right) \\
&= \mathbb{P}\left(Z < \frac{b - m}{\sigma}\right) \\
&= \Phi\left(\frac{b - m}{\sigma}\right),
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X > a) &= \mathbb{P}\left(\frac{X - m}{\sigma} > \frac{a - m}{\sigma}\right) \\
&= \mathbb{P}\left(Z > \frac{a - m}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{a - m}{\sigma}\right)
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{P}(a < X < b) &= \mathbb{P}\left(\frac{a - m}{\sigma} < \frac{X - m}{\sigma} < \frac{b - m}{\sigma}\right) \\
&= \mathbb{P}\left(\frac{a - m}{\sigma} < Z < \frac{b - m}{\sigma}\right) \\
&= \Phi\left(\frac{b - m}{\sigma}\right) - \Phi\left(\frac{a - m}{\sigma}\right).
\end{aligned}$$

Exemple 5.30 Soit X une variable suivant la loi normale $\mathcal{N}(3, 2)$. Calculons les probabilités suivantes : $\mathbb{P}(X < 4)$, $\mathbb{P}(X < -1)$, $\mathbb{P}(X > 1)$, $\mathbb{P}(-3 < X < 4)$. On sait que

$$Z = \frac{X - 3}{2} \rightsquigarrow \mathcal{N}(0, 1).$$

On a :

$$\begin{aligned}
\mathbb{P}(X < 4) &= \mathbb{P}\left(\frac{X - 3}{2} < \frac{4 - 3}{2}\right) \\
&= \mathbb{P}(Z < 1/2) \\
&= \Phi(0.5) \\
&\simeq 0.6915.
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X < -1) &= \mathbb{P}\left(\frac{X - 3}{2} < \frac{-1 - 3}{2}\right) \\
&= \mathbb{P}(Z < -2) \\
&= \Phi(-2) \\
&= 1 - \Phi(2) \\
&\simeq 0.0227.
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X > 1) &= \mathbb{P}\left(\frac{X-3}{2} > \frac{1-3}{2}\right) \\
&= \mathbb{P}(Z > -1) \\
&= 1 - \mathbb{P}(Z < -1) \\
&= 1 - (1 - \Phi(1)) \\
&= \Phi(1) \\
&\simeq 0.8413.
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(-3 < X < 4) &= \mathbb{P}\left(\frac{-3-3}{2} < \frac{X-3}{2} < \frac{4-3}{2}\right) \\
&= \mathbb{P}(-3 < Z < 1/2) \\
&= \Phi(1/2) - \Phi(-3) \\
&= \Phi(1/2) - 1 + \Phi(3) \\
&\simeq 0.6901.
\end{aligned}$$

5.2.3.6 Quantiles

Les **quantiles** de la loi normale $\mathcal{N}(0,1)$ jouent un grand rôle en probabilité et en statistique (tests d'hypothèses, construction d'intervalles de confiance, ...).

On rappelle ci-après la définition des quantiles.

Définition 5.31 Soit X une v.a. à densité dont la fonction de répartition est notée F. Pour tout $\alpha \in]0,1[$, on appelle **quantile** d'ordre α, le réel q_α tel que

$$F(q_\alpha) = \alpha \qquad \text{ou encore} \qquad \mathbb{P}(X \leqslant q_\alpha) = \alpha.$$

La fonction quantile est l'inverse de la fonction de répartition.

Les quantiles usuels de la loi normale centrée réduite sont donnés sous la forme d'une table disponible en annexe (voir page 118). Ils vérifient la propriété fondamentale suivante :

Proposition 5.32 *Soit $\alpha \in]0,1[$. Notons z_α le quantile d'ordre α de la loi normale centrée réduite. Alors, on a :*

$$z_{1-\alpha} = -z_\alpha.$$

Exemple 5.33 On a : $z_{0.975} = 1.96$ et $z_{0.025} = z_{1-0.975} = -z_{0.975} = -1.96$.

5.3 Exercices

Exercice 5.1 Vous vous livrez chaque matin (du lundi au vendredi inclus) au petit jeu suivant : ayant jeté une pièce de monnaie, vous vous rendez en cours si le résultat obtenu est « face » ; dans le cas contraire, vous restez au lit. Soit X le nombre de jours où vous avez « gagné » le droit au repos.

1) Quelle est la loi de X ? Justifiez votre réponse.
2) Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.

Exercice 5.2 On jette 10 pièces de monnaies truquées de telle sorte que pour chacune d'elles, la probabilité d'obtenir "pile" soit 0.3. Soit X la v.a. égale au nombre de "piles" obtenus au cours de ce lancer.

1) Loi de X, $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$?
2) Probabilité d'obtenir 3 "piles" ? moins de 3 "piles" ?
3) Probabilité que l'on ait obtenu plus de 3 piles sachant que l'on a obtenu au plus 5 "piles" ?

Exercice 5.3 On considère deux types d'avions A et B ayant respectivement 4 et 2 moteurs. Les moteurs sont supposés indépendants les uns des autres, et ils ont une même probabilité p $(0 < p < 1)$ de tomber en panne. Chaque avion arrive à destination si moins de la moitié de ses moteurs tombe en panne (c'est-à-dire que l'avion A arrive à destination si au plus un des moteurs tombe en panne et l'avion B arrive à destination si aucun des moteurs ne tombe en panne). Soient X et Y respectivement le nombre de moteurs tombant en panne pour A et B.

1) Quelles sont les lois respectives de X et Y ?
2) Déterminer en fonction de p les probabilités p_A (resp. p_B) pour que l'avion A (resp. B) arrive à destination.
3) Quel avion choisirez-vous ? (on discutera en fonction de la valeur de p).

Exercice 5.4 Le nombre d'ordinateurs vendus chaque jour dans un magasin spécialisé suit une loi de Poisson de paramètre 4. Calculer la probabilité que dans une journée :

1) on ne vende aucun ordinateur,
2) on vende 4 ordinateurs,
3) on vende au moins un ordinateur,
4) on vende au plus 3 ordinateurs,
5) le nombre d'ordinateurs vendus est compris entre 2 et 6.

Exercice 5.5 Soit X une v.a. de loi exponentielle $\mathcal{E}(\lambda)$ avec $\lambda > 0$.

1) En utilisant la f.r. de X, vérifier que $\mathbb{P}(X > h + t | X > h) = \mathbb{P}(X > t)$. On dit que la loi exponentielle est « sans mémoire ».
2) On suppose que la durée de vie d'une clé USB suit une loi exponentielle de paramètre $\lambda = 1/24$. L'unité de temps est le mois. Quelle est la probabilité que :

(a) la clé USB fonctionne au moins 8 mois ?

(b) la clé fonctionne encore 4 mois ou plus sachant qu'elle fonctionne toujours au bout de 8 mois ?

Exercice 5.6 Une usine produit des aspirateurs dont le prix de revient est de 100 UM (unités monétaires) et le prix de vente est de 150 UM l'unité. Chaque aspirateur vendu est retourné à l'usine et intégralement remboursé en cas de panne pendant les deux premières années après la vente. La durée de bon fonctionnement d'un aspirateur neuf, exprimée en années, est une variable aléatoire X de loi exponentielle de paramètre $\lambda = 0.1$.

1) Calculer la probabilité p qu'un aspirateur vendu tombe en panne pendant les deux premières années (donner une valeur approchée de p avec 3 chiffres après la virgule).

2) Vous décidez d'acheter simultanément deux aspirateurs dont les pannes sont supposées indépendantes. On désigne par Y le nombre d'aspirateurs qui tomberont en panne pendant les deux premières années.

   (a) Quelle loi de probabilité de Y ?

   (b) Calculer les probabilités : (i) qu'aucun aspirateur ne tombe en panne pendant ces deux premières années ; (ii) qu'au moins un aspirateur tombe en panne pendant ces deux premières années.

   (c) Déterminer en fonction de Y, l'expression du bénéfice B réalisé par l'usine sur ces deux ventes.

Exercice 5.7

1) Quelle est la probabilité d'obtenir un double (deux mêmes nombres) lorsqu'on lance une fois une paire de dés ?

2) On lance de façon répétée une paire de dés jusqu'à la première obtention d'un double. Soit X la variable aléatoire égale au nombre de lancers ainsi réalisés. Donner les valeurs de $\mathbb{P}(X = 1)$, $\mathbb{P}(X = 2)$ et plus généralement $\mathbb{P}(X = k)$ où $k \geqslant 1$ est un entier quelconque. Comment s'appelle la loi de X ?

3) Soit $n \geqslant 1$ un nombre entier naturel. Calculer la probabilité de n'obtenir aucun double lors des n premiers lancers. En déduire $\mathbb{P}(X > n)$.

4) Pour $k > n$, vérifier que $\mathbb{P}(X = k|X > n) = \mathbb{P}(X = k - n)$. En donner une interprétation.

Exercice 5.8 Soit X une v.a. suivant la loi normale centrée réduite. Calculer $\mathbb{P}(-2 < X < -1)$, $\mathbb{P}(|X| < 3)$, $\mathbb{P}(X < 2|X > 1.65)$.

Exercice 5.9 Soit X une v.a.r suivant la loi $\mathcal{N}(8, 4)$. Calculer $\mathbb{P}(X < 7.5)$, $\mathbb{P}(X > 8.5)$, $\mathbb{P}(6.5 < X < 10)$ et $\mathbb{P}(X > 6 \mid X > 5)$.

Exercice 5.10 Une machine remplit des sachets de pure water de 50 cl. On suppose que la quantité effective d'eau dans un sachet est une variable aléatoire X de loi normale $\mathcal{N}(50, 4)$.

1) Calculer $\mathbb{P}(X < 48)$, $\mathbb{P}(48 < X < 52)$, $\mathbb{P}(X > 52)$, $\mathbb{P}(X < 53)$ et $\mathbb{P}(X < 52 \mid X > 48)$.

2) Déterminer a tel que $\mathbb{P}(X > a) = 0.99$

Exercice 5.11 Pour se rendre à son travail un ouvrier a le choix entre deux chemins 1 et 2. Pour le chemin 1, la durée du parcours exprimée en minutes est une variable aléatoire normale X_1 de loi $\mathcal{N}(27, 5)$ alors que pour le chemin 2 la durée du parcours exprimée en minutes est une variable aléatoire normale X_2 de loi $\mathcal{N}(30, 2)$. A votre avis, quel chemin cet ouvrier doit-il choisir pour ne pas arriver en retard s'il dispose de 30 minutes ?

Exercice 5.12 On suppose que la taille X des hommes suit une loi normale d'espérance 1.75 m et d'écart-type 0.07 m. On mesure la taille chez un individu.

1) Calculer la probabilité que sa taille soit
    (a) inférieure à 1.54 m
    (b) supérieure à 1.85 m
    (c) comprise entre 1.54 m et 1.85 m.
2) On mesure la taille chez 2000 individus. On assimile la situation à un tirage avec remise et on désigne par X le nombre d'individus dont la taille est supérieure à 1.85 m. Déterminer la loi de X et donner le nombre moyen d'individus dont la taille est supérieure à 1.85m.

Exercice 5.13 Dans un pays dont la population est composée de femmes à 55%, la taille (en mètres) des adultes de sexe masculin (resp. féminin) est distribuée selon une loi normale de moyenne 1.78 (resp. 1.72) et d'écart-type 0.14 (resp. 0.1). On tire une personne au hasard dans la population. Quelle est la probabilité

1) qu'elle mesure moins de 1.76m sachant c'est une femme ?
2) qu'elle mesure moins de 1.76m sachant c'est un homme ?
3) que ce soit une personne mesurant moins de 1.76m ?
4) que ce soit une femme sachant qu'elle mesure moins de 1.76m ?
5) que ce soit un homme mesurant moins de 1.76m ?

CHAPITRE 6

APPROXIMATIONS USUELLES DES LOIS DE PROBABILITÉ

Résumé – Ce chapitre présente les approximations usuelles des lois de probabilité discrètes et continues souvent utilisées dans la pratique. L'objectif est de savoir faire des calculs de probabilité en utilisant lesdites approximations.

Sommaire

6.1 APPROXIMATIONS USUELLES

6.1.1 Approximation binomiale de la loi hypergéométrique

Quand N tend vers l'infini, on peut approximer la loi hypergéométrique $\mathcal{H}(N, n, p)$ par la loi binomiale $\mathcal{B}(n, p)$. Dans la pratique, ce résultat s'applique dès que $n/N < 10\%$ c'est-à-dire dès que la population est 10 fois plus grande que l'échantillon. C'est pour cela qu'en statistique, lorsque la taille N de la population est très grande, on peut assimiler un tirage sans remise (loi hypergéométrique) à une tirage avec remise (loi binomiale).

6.1.2 Approximation de la loi binomiale par une loi de Poisson

Quand $n \to +\infty$ et $p \to 0$, on peut approximer la loi binomiale $\mathcal{B}(n, p)$ par une loi de Poisson $\mathcal{P}(np)$. En pratique, cela est possible pour des valeurs très faibles de la probabilité p ($p < 0.1$) et des valeurs très grandes de n ($n > 50$). Une illustration en est donnée par la figure 6.1.

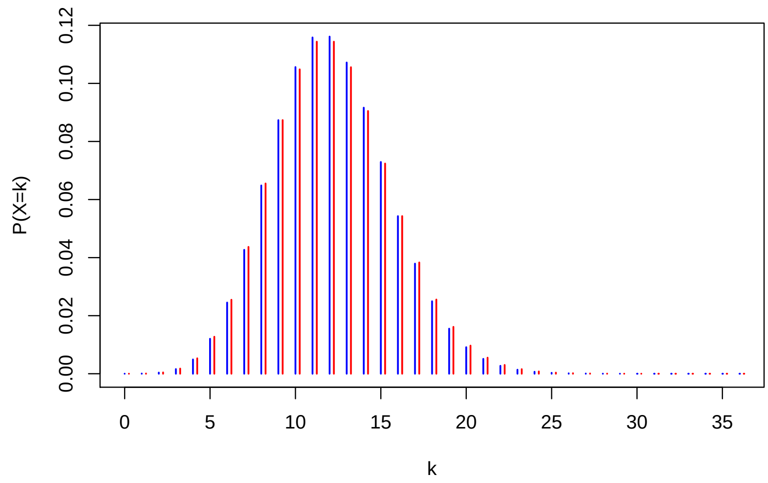


FIGURE 6.1 – Approximation de la loi $\mathcal{B}(400; 0,03)$ (en bleu) par la loi $\mathcal{P}(12)$ (en rouge). Tous les bâtons ne sont pas affichés (pour la loi de Poisson, il y en a une infinité). On se limite aux valeurs de k pour lesquelles la probabilité est assez éloignée de zéro pour être visible sur le graphique.

6.1.3 Approximation de la loi binomiale par une loi normale

Lorsque $n \to +\infty$, on peut approximer la loi binomiale $\mathcal{B}(n, p)$ par la loi normale $\mathcal{N}(np, \sqrt{npq})$ où $q = 1 - p$. En pratique, les deux conditions généralement requises sont $np > 5$ et $n(1 - p) > 5$.

Comme il s'agit de l'approximation d'une loi discrète par une loi continue, on procède à ce que l'on appelle **correction de continuité** de la façon suivante : Si $X \sim \mathcal{B}(n, p)$ et si $k \in \{1, \ldots, n\}$ alors

$$\begin{aligned}
\mathbb{P}(X = k) &\simeq \mathbb{P}(k - 0.5 < X < k + 0.5) \\
&\simeq \mathbb{P}\left(\frac{k - 0.5 - np}{\sqrt{npq}} < Z < \frac{k + 0.5 - np}{\sqrt{npq}}\right) \\
&\simeq \Phi\left(\frac{k + 0.5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k - 0.5 - np}{\sqrt{npq}}\right) \qquad (6.1)
\end{aligned}$$

où Z est la v.a. centrée réduite de loi $\mathcal{N}(0, 1)$ associée à X et Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$; de même,

$$\begin{aligned}
\mathbb{P}(a \leqslant X \leqslant b) &\simeq \mathbb{P}(a - 0.5 \leqslant X \leqslant b + 0.5) \\
&\simeq \mathbb{P}\left(\frac{a - 0.5 - np}{\sqrt{npq}} \leqslant U \leqslant \frac{b + 0.5 - np}{\sqrt{npq}}\right) \\
&\simeq \Phi\left(\frac{b + 0.5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - 0.5 - np}{\sqrt{npq}}\right). \qquad (6.2)
\end{aligned}$$

Si on n'appliquait pas la correction de continuité, on trouverait systématiquement $\mathbb{P}(X = k) = 0$ puisque la loi normale est une loi continue.

Exemple 6.1 Pour $X \sim \mathcal{B}(100, 0.3)$, $np = 30$ et $nq = 70$, la valeur exacte de $\mathbb{P}(X = 30)$ est

$$\mathbb{P}(X = 30) = C_{100}^{30} 0.3^{30} 0.7^{70} \approx 0.08678.$$

La formule d'approximation avec une loi $\mathcal{N}(np, \sqrt{npq}) = \mathcal{N}(30, \sqrt{21})$ donne le résultat :

$$\mathbb{P}(X = 30) \approx \Phi\left(\frac{30 + 0.5 - 30}{\sqrt{21}}\right) - \Phi\left(\frac{30 - 0.5 - 30}{\sqrt{21}}\right) \approx 0.08688.$$

L'erreur d'approximation est très faible.

6.1.4 Approximation de la loi de Poisson par une loi normale

Lorsque $\lambda \to +\infty$, on peut approximer la loi de Poisson $\mathcal{P}(\lambda)$ par la loi normale $\mathcal{N}(\lambda, \sqrt{\lambda})$. En pratique, cette approximation est satisfaisante dès que $\lambda > 18$. Ici aussi, comme il s'agit d'une approximation d'une loi discrète par une loi continue, il faudra appliquer la correction de continuité.

6.2 Théorème de la limite centrale

6.2.1 Indépendance de variables aléatoires

Comme pour les événements, on définit la notion de variables aléatoires mutuellement indépendantes comme suit.

Définition 6.2 Soit $X_1, \ldots, X_n$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que c'est une suite de *variables aléatoires mutuellement indépendantes* si pour toute suite finie extraite $X_{i_1}, \ldots, X_{i_k}$ et toute suite finie $B_{i_1}, \ldots, B_{i_k}$ d'intervalles, on a :

$$\mathbb{P}(X_{i_1} \in B_{i_1}, \ldots, X_{i_k} \in B_{i_k}) = \mathbb{P}(X_{i_1} \in B_{i_1}) \times \cdots \times \mathbb{P}(X_{i_k} \in B_{i_k}).$$

En particulier si les variables $X_1, \ldots, X_n$ sont mutuellement indépendantes et ont la même loi, on parle de suite de variables aléatoires *indépendantes identiquement distribuées* (i.i.d).

L'objectif de ce cours n'est pas de démontrer l'indépendance mutuelle de v.a. Dans toute la suite, on supposera que des v.a. sont mutuellement indépendantes si rien ne laisse penser que certaines v.a. agissent sur d'autres. Une suite de v.a. mutuellement indépendantes sera simplement dite indépendante.

6.2.2 Théorème de la limite centrale

Le théorème de la limite centrale (encore appelé théorème limite central) permet d'approcher la loi de la somme de variables aléatoires indépendantes par la loi normale sous certaines hypothèses.

Théorème 6.3 *Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires i.i.d. dont la loi commune admet une espérance mathématique $m = \mathbb{E}(X_1)$ et une variance finie $\sigma^2 = \mathbb{V}\mathrm{ar}(X_1) < +\infty$. Pour tout $n \geqslant 1$, posons*

$$S_n = \sum_{i=1}^{n} X_i = X_1 + \cdots + X_n \quad et \quad \overline{X}_n = \frac{S_n}{n} = \sum_{i=1}^{n} X_i.$$

Quand $n \to +\infty$, la v.a. Z_n définie par

$$Z_n = \frac{\overline{X}_n - m}{\sigma/\sqrt{n}} = \frac{S_n - nm}{\sigma\sqrt{n}}$$

suit approximativement la loi $\mathcal{N}(0,1)$.

Le théorème 6.3 joue un rôle très important en probabilités et en statistique du fait que pour n « grand » (en général, $n \geqslant 30$), on peut écrire

$$S_n \rightsquigarrow \mathcal{N}(nm, \sigma\sqrt{n}) \qquad \text{et} \qquad \overline{X}_n \rightsquigarrow \mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right).$$

6.3 EXERCICES

Exercice 6.1 On jette 360 fois un dé équilibré dont les faces sont numérotées de 1 à 6 et on désigne par X le nombre d'apparitions du numéro 1.

1) Quelle est la loi exacte de X ?
2) En approximant la loi de X par une loi normale, déterminer la probabilité que le nombre d'apparitions du numéro 1 soit compris entre 48 et 72.

Exercice 6.2 Une compagnie d'assurance doit rembourser $n = 100$ personnes. L'unité monétaire est notée kFCFA. On suppose que ces remboursements sont des v.a. $X_1, \ldots, X_n$ indépendantes et identiquement distribuées d'espérance 50 kFCFA et d'écart type 40 kFCFA.

1) Calculer la probabilité que $\bar{X}_n$ soit comprise entre 45 kFCFA et 55 kFCFA.
2) Déterminer le montant a tel que $\mathbb{P}\left(\overline{X}_n < a\right) = 0.95$
3) Déterminer le montant b tel que $\mathbb{P}\left(\overline{X}_n > b\right) = 0.8$

Exercice 6.3 En 2020, une compagnie d'assurances assure n personnes contre un certain risque. On note X_i le remboursement annuel total qu'aura à verser la compagnie au $i-$ème client et on suppose que les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes identiquement distribuées suivant une loi d'espérance m et d'écart type σ.

1) Soit a la prime annuelle payée par chacun des n clients. La compagnie est perdante si le remboursement total dépasse la prime totale encaissée. Déterminer la valeur minimale de a pour que la probabilité que la compagnie soit perdante soit inférieure à 1%.
2) On donne $n = 100$, $m = 400$ et $\sigma = 150$. Déterminer b tel que $\mathbb{P}(\overline{X}_n > b) = 0.99$.

On rappelle que $\Phi^{-1}(0.99) = 2.326$

Exercice 6.4 Un ascenseur peut contenir une charge maximale de 600 kg. Le poids de chaque personne est modélisé par une une v.a. X_i de moyenne 60 kg et d'écart-type 10 kg. Quel est le nombre maximal n de personnes pouvant monter simultanément dans l'ascenseur pour que la charge maximale ne soit dépassée qu'avec une probabilité de 5% ?

Deuxième partie

Statistique descriptive

Chapitre 7

Concepts fondamentaux de la statistique

Résumé – Dans ce chapitre, nous rappelons les concepts fondamentaux de la statistique.

Sommaire

7.1 Définition et champs d'application de la statistique

Le mot « **statistique** » a plusieurs sens. Il peut désigner un ensemble de données (par exemple, les statistiques du chômage au Togo) ou encore l'ensemble des méthodes de collecte, de traitement, d'analyse, d'interprétation et de présentation (diffusion) des données afin de les rendre accessibles. Le mot « statistique » est aussi parfois utilisé pour désigner un paramètre numérique calculé à partir des données.

Les champs d'application de la statistique sont très nombreux. Partout où il y a des données, l'on a besoin de la statistique. La science statistique intervient à toutes les étapes, de la collecte des données jusqu'à leur analyse, leur interprétation et leur diffusion. La statistique intervient aussi bien dans les sciences expérimentales (physique, chimie, biologie, etc...) qu'en informatique, en économie et en gestion, en médecine, en sociologie ou en psychologie. La statistique est souvent à la base de décisions importantes, les décideurs ne regardant pas les données mais plutôt les informations qui en sont extraites. Le statisticien doit donc s'imposer une démarche rigoureuse tout en ne perdant pas de vue l'objectivité.

7.2 VOCABULAIRE DE LA STATISTIQUE

7.2.1 Population, individus et échantillon

Définition 7.1

- L'ensemble Ω sur lequel porte une étude statistique est appelé la **population**.
- Un élément ω de cette population est appelé **individu** ou **unité statistique**.
- Le nombre n d'individus composant la population à étudier est appelé effectif total ou **taille** de la population.

Il faut prêter attention au fait que les termes « population » et « individu » sont employés aussi bien lorsqu'il s'agit d'êtres humains que d'objets quelconques. Il est toujours indispensable de bien définir la population soumise à l'étude statistique.

Exemple 7.2

- Si l'on souhaite étudier l'âge moyen des employés de banque de la ville de Lomé, la population est l'ensemble des employés de banque de Lomé et l'unité statistique est un employé de banque de Lomé.
- Par contre, lorsque l'on souhaite étudier le nombre d'employés des banques de la ville de Lomé, la population est l'ensemble des banques de la ville de Lomé (et non l'ensemble des employés de banque de la ville de Lomé) et l'unité statistique est une banque de la ville de Lomé.

Définition 7.3 On appelle **échantillon** toute partie (tout sous-ensemble) de la population soumise à l'étude statistique.

Les notions de population, individus et échantillons sont illustrées par la figure 7.1.

Définition 7.4

- Le statisticien fait des relevés sur les individus d'une population : ce sont les **observations**.
- Lorsque l'étude statistique consiste à faire des observations sur tous les individus d'une population, on parle de **recensement**.
- Lorsque l'étude statistique consiste à faire des observations sur un échantillon de la population, on parle de **sondage**.

L'utilisation d'un sondage plutôt que d'un recensement s'explique par le fait qu'un recensement dans une population de grande taille est très couteux en temps et en ressources financières. Par exemple, lors de la campagne pour une élection présidentielle, il est souhaitable pour les candidats d'avoir, sur une base journalière, une idée sur les intentions de vote afin d'apprécier l'acceptation de leurs programmes par la population

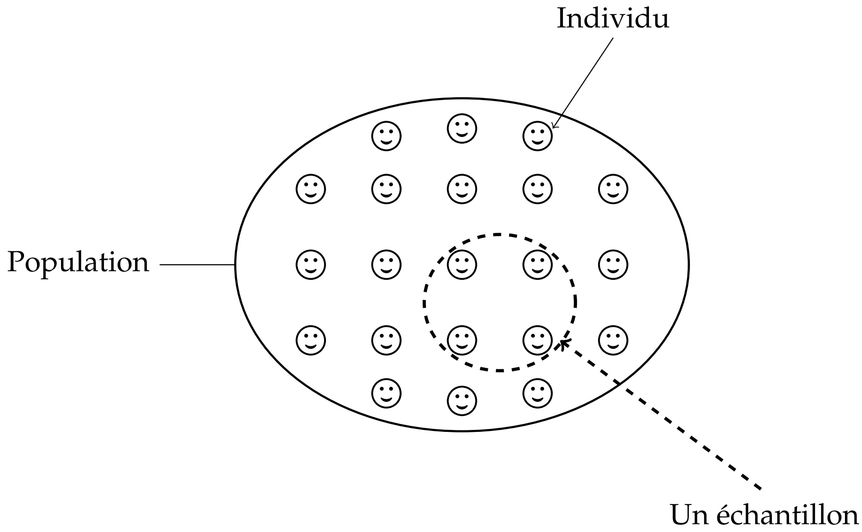


FIGURE 7.1 – Population, individus et échantillon.

et de faire des ajustements si nécessaire. Mais il n'est pas possible d'interroger des millions de personnes en une journée et ce, chaque jour. L'étude des méthodes de sélection d'un échantillon dit « représentatif » et de la généralisation des résultats obtenus sur ledit échantillon à la population entière fait l'objet de la théorie des sondages et de la théorie de l'estimation.

7.2.2 Caractères ou variables statistiques

Définition 7.5 Les caractéristiques ou propriétés étudiées sur les individus sont appelées **caractères** ou **variables**. Mathématiquement, un **caractère** est une application χ de l'ensemble Ω (la population) dans l'ensemble C des valeurs possibles dudit caractère, qui associe à chaque individu ω de Ω la valeur $\chi(\omega)$ que prend ce caractère sur l'individu ω. Les valeurs possibles d'un caractère sont appelées ses **modalités**.

Remarque 7.6 Les modalités d'un caractère doivent être incompatibles (un individu ne peut appartenir à deux modalités ou plus) et exhaustives (chaque individu appartient à une modalité car tous les cas ont été prévus). Ainsi, chaque individu doit appartenir à une et une seule modalité.

Il y a deux types de caractères : quantitatif et qualitatif. La distinction entre les deux types de caractères est très importante puisque les méthodes d'analyse diffèrent suivant la nature du caractère étudié.

7.2.2.1 Caractères quantitatifs

Un caractère est dit **quantitatif** lorsque ses modalités ont des quantités numériques. Par exemple l'âge, la taille, la température, le revenu, le nombre d'enfants d'un couple, la durée de travail sont toutes des variables quantitatives.

Les modalités d'une variable quantitative sont parfois si nombreuses que pour des raisons de commodité, elles sont regroupées en classes. Par exemple, dans l'étude de l'âge des individus d'un pays, la variable « age » peut être définie selon des tranches d'âge exprimant des modalités telles que 0-4 ans, 5-9 ans, 10-17 ans, etc. On distingue deux types de variables quantitatives : les variables **discrètes** et les variables **continues**.

Une variable quantitative X est dite **discrète** si l'ensemble de ses modalités est au plus dénombrable, c'est-à-dire si ses valeurs peuvent être énumérées sous la forme d'une liste de chiffres $x_1, x_2, \ldots$. Parmi les exemples de variables quantitatives discrètes, on peut citer : le nombre d'enfants d'une famille, la répartition des étudiants togolais étudiant à l'étranger dans les pays de l'UEMOA en 2022-2023, le nombre d'accidents mensuels sur les routes togolaises en 2022, etc. . ..

Une variable quantitative est dite **continue** si ces modalités ne sont pas au plus dénombrables c'est-à-dire qu'on ne peut pas les numéroter ou si ses modalités sont regroupées par classes (intervalles de $\mathbb{R}$). Une telle variable peut prendre toutes la valeurs dans un (plusieurs) intervalle(s) de $\mathbb{R}$. Parmi les exemples de variables quantitatives continues, on peut citer : la taille d'une personne, le poids des individus d'une population, la durée de vie d'un composant électronique, etc. . ..

Remarque 7.7 Pour une variable quantitative continue, le nombre de classes doit respecter certains critères. Le nombre de classes à retenir dépend de la précision des mesures et de la taille de la population étudiée. Il ne doit être ni trop grand (sinon le problème de départ qui concerne l'abondance des données n'est pas résolu) ni trop petit (car il y aurait perte d'informations).

7.2.2.2 Caractères qualitatifs

Une variable est dit **qualitative** lorsque ses modalités sont des mots ou des lettres n'exprimant pas de quantité. Une variable qualitative n'ayant que deux modalités est dite **dichotomique**.

On distingue deux types de variables qualitatives : les variables **nominales** et les variables **ordinales**.

Une variable qualitative est dite **nominale** si ses modalités ne sont pas ordonnées. Par exemple on peut citer : le sexe, la couleur des yeux, le loisir préféré des individus d'une population, la marque de voiture, le numéro de téléphone, etc. . .

Une variable qualitative est dite **ordinale** si ses modalités peuvent être ordonnées. Par exemple on peut citer : la mention à un examen ("Passable", "Assez-bien", "Bien", "Très-bien"), la cotation attribuée par un jury (par exemple A = "Très satisfaisant", B = "Satisfaisant", . . .), le degré de satisfaction, etc. . .

Remarque 7.8 Il arrive (par exemple pour des raisons de traitement informatique) de coder les modalités d'une variable qualitative avec des valeurs numériques. Dans le cas de la variable "sexe", par exemple, on pourrait noter 0 pour la catégorie « homme » et 1 pour la catégorie « femme ». Il est important de souligner que les nombres utilisés pour de tels codages sont non-quantitatifs. Le fait d'attribuer des valeurs numériques pour représenter les diverses modalités d'une variable qualitative ne signifie pas que ces nombres possèdent des propriétés arithmétiques. Ces codes ne servent qu'à identifier les modalités de manière pratique. On n'imagine mal le calcul du sexe moyen dans une population.

Application 7.9 Le recensement général de la population et de l'habitat (RGPH) du Togo en 2010 a donné les résultats suivants (voir Tableau 7.1).

TABLE 7.1 – RGPH 2010 (source : `http://data.gouv.tg`).

Région	Maritime*	Plateaux	Centrale	Kara	Savanes
Nombre d'hommes	1 248 354	678 191	308 443	376 111	397 996
Nombre de femmes	1 351 601	696 974	309 428	393 829	430 228
Nombre d'habitants	2 599 955	1 375 165	617 871	769 940	828 224

* Lomé - Commune inclus.

Préciser la population étudiée, les individus, les caractères et leur nature.

Application 7.10 Indiquer de quel type sont les variables présentées ci-dessous : (qualitatives nominales, qualitatives ordinales, quantitatives discrètes ou quantitatives continues).

(a) L'état-civil des habitants togolais.

(b) La taille des étudiants de notre université.

(c) Le nombre de pages d'un support de cours.

(d) Le nombre de ventes d'un appareil électro-ménager.

(e) Le sexe des élèves passant le BAC.

(f) La nationalité des élèves d'une classe.

(g) Le nombre de télévisions par ménage.

(h) Le degré de qualification du personnel d'une entreprise.

(i) La couleur des yeux des étudiants.

(j) Le nombre de jours de pluie pendant le mois d'août.

7.2.3 Série statistique

Définition 7.11 Les observations d'un ou plusieurs caractères sur toute la population forment une **série statistique**.

Les séries statistiques les plus élémentaires sont les séries à une variable (séries statistiques **univariées**) et les séries à deux variables (séries statistiques **bivariées**).

7.3 Statistique descriptive et statistique inférentielle

Les méthodes statistiques peuvent être regroupées en deux grands sous-groupes : la **statistique descriptive** et la **statistique inférentielle** :

- la **statistique descriptive** s'intéresse au traitement des données pour en dégager des renseignements quantitatifs ou qualitatifs et ne s'applique que si les données ont été collectées sur la population en entier.

- la **statistique inférentielle** (encore appelée **statistique mathématique**) a pour fonction d'aider à extrapoler, à partir d'un échantillon tiré de la population à étudier, le comportement de la population dans son ensemble. Cette statistique utilise des modèles théoriques (lois de probabilité) et nécessite la recherche d'échantillons « représentatifs » (qui représentent le mieux possible la diversité de la population entière) constitués au hasard.

Remarque 7.12 La statistique inférentielle dépasse le cadre de ce cours. Dans la suite de cette partie, nous nous concentrerons sur la statistique descriptive.

CHAPITRE 8

SÉRIES STATISTIQUES À UNE VARIABLE

Résumé – Ce chapitre présente les méthodes usuelles de traitement des séries statistiques à une variable. L'objectif est de savoir représenter graphiquement de telles séries et calculer les indicateurs numériques qui y sont liés.

Sommaire

8.1 DESCRIPTION D'UNE SÉRIE STATISTIQUE À UNE VARIABLE

8.1.1 Description d'une série quantitative

Soit X une variable quantitative observée sur une population Ω de taille n. Si X est discrète, on note par $\{x_1, \dots, x_p\}$ l'ensemble de ses modalités. Nous supposons que les valeurs $x_1, \dots, x_p$ sont ordonnées c'est-à-dire :

$$x_1 < \cdots < x_p.$$

Définition 8.1 Si X est continue, les modalités sont regroupées en intervalles appelés **classes** pour rendre les données plus lisibles. L'ensemble C des modalités est partagé en classes $C_1, \dots, C_p$, la classe C_i étant de la forme $C_i = [a_{i-1}, a_i[$ où $a_{i-1} < a_i$. Les nombres a_{i-1} et a_i sont appelés les **bornes** de la classe C_i, $c_i = \dfrac{a_{i-1} + a_i}{2}$ son **centre** et $l_i = a_i - a_{i-1}$ son **amplitude**. Par convention, a_0 (resp. a_p) désigne la plus petite (resp. la plus grande) valeur prise par X.

Définition 8.2 Soit X une variable discrète (resp. continue) décrite par ses modalités (resp. ses classes) $x_1, \cdots, x_p$ (resp. $C_1 = [a_0, a_1[, \ldots, C_p = [a_{p-1}, a_p[$). On appelle :

1) **effectif** de la valeur x_i (resp. de la classe C_i) : le nombre n_i d'individus pour lesquels $X = x_i$ (resp. $X \in C_i$ c'est-à-dire $a_{i-1} \leqslant X < a_i$) ;
2) **fréquence** de la valeur x_i (resp. de la classe C_i) : le rapport $f_i = \dfrac{n_i}{n}$;
3) **effectif cumulé croissant** associé à x_i (resp. à a_i) :
$$n_{ic} = n_1 + \cdots + n_i = \sum_{k=1}^{i} n_k;$$
en d'autres termes, n_{ic} est le nombre d'individus pour lesquels X prend les modalités inférieures ou égales à x_i (resp. a_i) ;
4) **effectif cumulé décroissant** associé à x_i (resp. à a_{i-1}) :
$$n_{id} = n_i + \cdots + n_p = \sum_{k=i}^{p} n_k;$$
en d'autres termes, n_{id} est le nombre d'individus pour lesquels X prend les modalités supérieures ou égales à x_i (resp. a_i).
5) On définit les **fréquences cumulées croissantes** et les **fréquences cumulées décroissantes** en divisant les effectifs cumulés correspondants par n. Ainsi $f_{ic} = \dfrac{n_{ic}}{n}$ et $f_{id} = \dfrac{n_{id}}{n}$, pour tout $i = 1, \ldots, p$.
6) On appelle **distribution statistique** de la variable X, la suite de couples (x_i, n_i) (resp. (C_i, n_i)), $i = 1, \ldots, p$.

Remarque 8.3

1) Les effectifs associés aux différentes modalités (ou classes) vérifient la relation
$$n_1 + \cdots + n_p = \sum_{i=1}^{p} n_i = n.$$
2) Chaque fréquence f_i est un nombre décimal compris entre 0 et 1 et peut donc être exprimée en pourcentage. De plus, la somme des fréquence est égale à 1 ou 100%.
3) On notera aussi que $n_{pc} = n_{1d} = n$ et $f_{pc} = f_{1d} = 1$ ou 100%.
4) Il est souvent plus commode de définir la distribution statistique d'une variable par un tableau de données de la forme :

TABLE 8.1 – Exemple de distribution statistique d'une variable

Modalités	x_1	$\ldots$	x_p
Effectifs	n_1	$\ldots$	n_p

TABLE 8.2 – Nombre d'enfants pour 40 couples.

Nombre d'enfants x_i	0	1	2	3	4
Nombre de couples n_i	12	20	5	2	1

Exemple 8.4 On a relevé le nombre d'enfants de 40 couples. Les résultats obtenus sont donnés dans le tableau 8.2.

Nous pouvons compléter ce tableau avec les fréquences et les effectifs cumulés.

x_i	0	1	2	3	4
n_i	12	20	5	2	1
n_{ic}	12	32	37	39	40
n_{id}	40	28	8	3	1
f_i (%)	30	50	**12.5**	5	2.5
f_{ic} (%)	30	80	**92.5**	97.5	100
f_{id} (%)	100	70	**20**	7.5	2.5

Quelle interprétation pouvons-nous donner aux valeurs en gras ? 12.5% des couples interrogés ont exactement 2 enfants, 92.5% de ces couples ont 2 enfants ou moins, 20% des couples ont 2 enfants ou plus.

Exemple 8.5 On a noté l'âge (arrondi à l'année près) de 48 salariés d'une entreprise. La série statistique brute (données obtenues pendant l'enquête) est donnée ci-dessous :

43	29	57	45	50	29	37	59	46	31	46	24
33	38	49	31	62	60	52	38	43	26	41	52
60	49	52	41	38	26	37	59	57	41	29	33
33	43	46	57	46	33	46	49	57	57	46	43

Le tableau de données ci-dessous donne la distribution statistique de l'âge des 48 salariés en considérant les classes d'âge $[20,30[$, $[30,40[$, $[40,45[$, $[45,50[$, $[50,55[$, $[55,60[$ et $[60,65[$:

TABLE 8.3 – Distribution statistique de l'âge des salariés.

C_i	$[20,30[$	$[30,35[$	$[35,40[$	$[40,45[$	$[45,50[$	$[50,55[$	$[55,60[$	$[60,65[$
n_i	6	6	5	7	10	4	7	3

Les effectifs cumulés sont donnés par :

C_i	$[20,30[$	$[30,35[$	$[35,40[$	$[40,45[$	$[45,50[$	$[50,55[$	$[55,60[$	$[60,65[$
n_i	6	6	5	7	10	4	7	3
n_{ic}	6	12	17	**24**	34	38	45	48
n_{id}	48	42	36	**31**	24	14	10	3

Quelle interprétation pouvons-nous donner aux valeurs en gras ? 24 salariés ont moins de 45 ans, 31 salariés ont 40 ans ou plus.

8.1.2 Description d'une série qualitative

Pour une variable qualitative, les notions d'effectif et de fréquence restent valables en considérant les différentes modalités dudit caractère. Néanmoins, le calcul des effectifs cumulés et des fréquences cumulées n'a de sens que pour les variables qualitatives ordinales.

Exemple 8.6 Reprenons les données du tableau 7.1 (page 79). L'effectif total de la population est $n = 6\,191\,155$. Le tableau des fréquences est le suivant :

TABLE 8.4 – Tableau de fréquences associé aux données du RGPH 2010.

Région	Maritime*	Plateaux	Centrale	Kara	Savanes
Nombre d'habitants	2 599 955	1 375 165	617 871	769 940	828 224
Fréquences	0.420	0.222	0.100	0.124	0.134

* Lomé - Commune inclus.

On peut y lire notamment que la région maritime abrite à elle seule 42% de la population togolaise (en 2010).

8.2 Représentations graphiques

Elles permettent d'avoir rapidement une vue d'ensemble d'un tableau de données et de mettre en évidence certains faits essentiels.

8.2.1 Variables quantitatives discrètes : diagramme en bâtons et polygone des effectifs/fréquences

Soit (x_i, n_i), $i = 1, \ldots, p$ une série statistique quantitative discrète et soient $f_1, \ldots, f_p$ les fréquences respectives de chaque modalité. Le **diagramme en bâtons** des effectifs (resp. des fréquences) de cette distribution statistique est constitué de p segments verticaux d'abscisses x_i et de hauteur n_i (resp. f_i).

Le **polygone des effectifs** (resp. des **fréquences**) est obtenu à partir du diagramme en bâtons en joignant par un segment les sommets des bâtons.

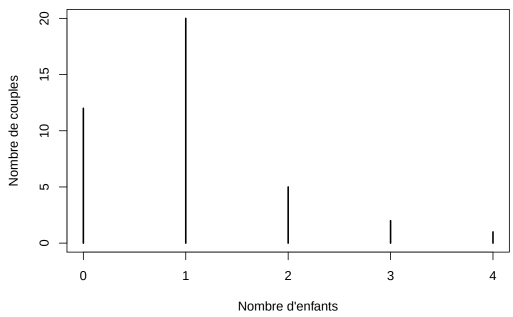
Exemple 8.7 La figure 8.1 donne le diagramme à bâtons et le polygone des effectifs associés au tableau 8.2.

8.2.2 Variables quantitatives continues

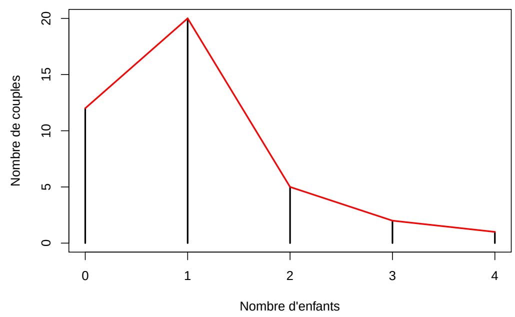
Soit $([a_{i-1}, a_i[, n_i)$, $i = 1, \ldots, p$ une série statistique quantitative continue et soient $f_1, \ldots, f_p$ les fréquences respectives associées à chaque classe.

8.2.2.1 Histogramme

C'est un graphique formé de p rectangles juxtaposés tels que le rectangle associé à la classe i ($i = 1, \ldots, p$) a une largeur égale à l'amplitude de la classe $[a_{i-1}, a_i[$ et une hauteur



(a) Diagramme à bâtons



(b) Polygone des effectifs

FIGURE 8.1 – Diagramme en bâtons et polygone des effectifs associés au tableau 8.2.

égale à

$$h_i = \frac{f_i}{l_i}, \qquad i = 1, \ldots, p. \tag{8.1}$$

Ainsi, l'aire totale de l'histogramme est égale à 1.

L'histogramme donne rapidement une image de l'allure globale de la distribution ; il montre l'étalement des données et apporte ainsi des renseignements sur la dispersion et sur les valeurs extrêmes ; il permet de déceler, éventuellement, des valeurs aberrantes.

Exemple 8.8 Reprenons le tableau 8.3 et complétons-le avec les hauteurs des rectangles composant l'histogramme.

TABLE 8.5 – Données relatives à l'histogramme de l'âge des salariés.

C_i	$[20, 30[$	$[30, 35[$	$[35, 40[$	$[40, 45[$	$[45, 50[$	$[50, 55[$	$[55, 60[$	$[60, 65[$
n_i	6	6	5	7	10	4	7	3
f_i	$6/48$	$6/48$	$5/48$	$7/48$	$10/48$	$4/48$	$7/48$	$3/48$
l_i	10	5	5	5	5	5	5	5
$h_i \simeq$	0.0125	0.025	0.0208	0.0292	0.0417	0.0167	0.0292	0.0125

L'histogramme est représenté sur la figure 8.2.

Remarque 8.9 Une particularité de l'histogramme est que l'échelle verticale n'a pas de signification particulière. On peut vérifier qu'un histogramme représente les fréquences par la surface des blocs et non par leur hauteur.

8.2.2.2 Polygone des effectifs (fréquences) cumulé(e)s

Le polygone des effectifs (fréquences) cumulé(e)s d'une distribution quantitative continue permet de représenter les effectifs (fréquences) cumulés croissants ou décroissants. Nous nous attardons sur le polygone des fréquences qui est plus utilisé. Le principe de construction de celui des effectifs est le même à condition de remplacer les fréquences

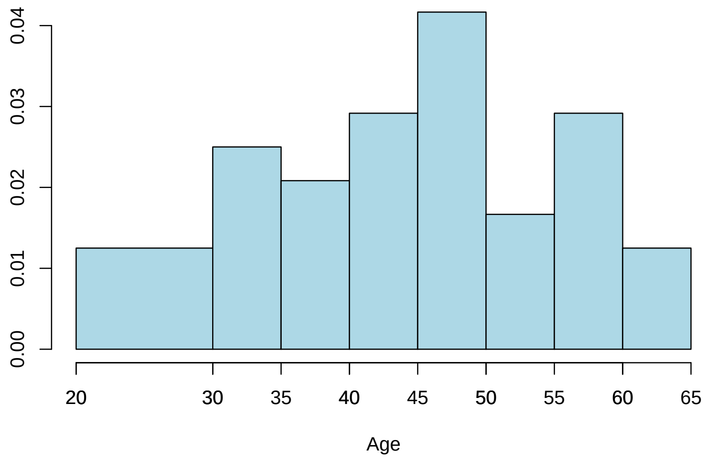


FIGURE 8.2 – Histogramme de l'age des salariés.

par les effectifs.

Le polygone des fréquences cumulées croissantes (resp. décroissantes) est obtenu en plaçant dans un repère orthogonal, les points de coordonnées (a_i, f_{ic}) (resp. (a_{i-1}, f_{id})) pour $i = 1, \ldots, p$. On joint par des segments de droite les points (a_i, f_{ic}) (resp. (a_{i-1}, f_{id})) pour $i = 1, \ldots, p$. Pour les points (a_0, f_{0c}) (resp. (a_p, f_{dp})), on adopte la convention $f_{0c} = 0$ (resp. $f_{0d} = 1$) et $f_{pc} = 1$ (resp. $f_{pd} = 0$) pour les fréquences cumulées croissantes (resp. décroissantes).

Exemple 8.10 Complétons les données du tableau 8.3 par les fréquences cumulées croissantes et décroissantes.

TABLE 8.6 – Fréquences cumulées de l'âge des salariés.

C_i	$[20, 30[$	$[30, 35[$	$[35, 40[$	$[40, 45[$	$[45, 50[$	$[50, 55[$	$[55, 60[$	$[60, 65[$
n_i	6	6	5	7	10	4	7	3
f_i	$6/48$	$6/48$	$5/48$	$7/48$	$10/48$	$4/48$	$7/48$	$3/48$
f_{ic}	$6/48$	$12/48$	$17/48$	$24/48$	$34/48$	$38/48$	$45/48$	1
f_{id}	1	$42/48$	$36/48$	$31/48$	$24/48$	$14/48$	$10/48$	$3/48$

Les polygones des fréquences cumulées croissantes et décroissantes sont superposés sur la figure 8.3.

Le polygone des fréquences cumulées croissantes est la courbe représentative de la

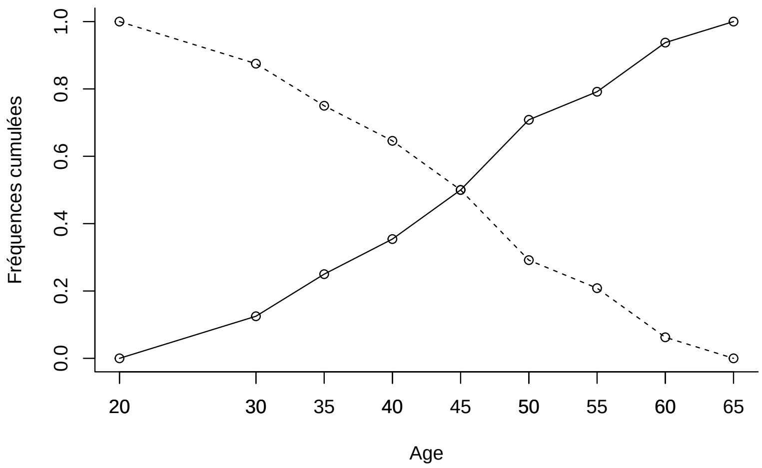


FIGURE 8.3 – Polygones des fréquences cumulées croissantes (ligne continue) et décroissantes (en tirets) de l'âge des salariés.

fonction F_X définie par :

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < x_0; \\ \\ \displaystyle\sum_{j=1}^{i-1} f_j + \frac{x - a_{i-1}}{a_i - a_{i-1}} f_i & \text{si} \quad a_{i-1} \leqslant x < a_i, \quad i = 1 \ldots, p; \\ \\ 1 & \text{si} \quad x \geqslant x_p. \end{cases} \tag{8.2}$$

La valeur $F_X(x)$ est nulle pour les valeurs de x inférieures à la plus petite valeur a_0 et égale à 1 pour les valeurs de x supérieures à la plus grande valeur a_p.

Pour tout $x \in \mathbb{R}$, $F_X(x)$ est une estimation de la proportion des individus pour lesquels $X \leqslant x$. Pour ce faire, on fait l'hypothèse fondamentale que **les valeurs sont équiréparties (uniformément réparties) au sein de chaque classe**.

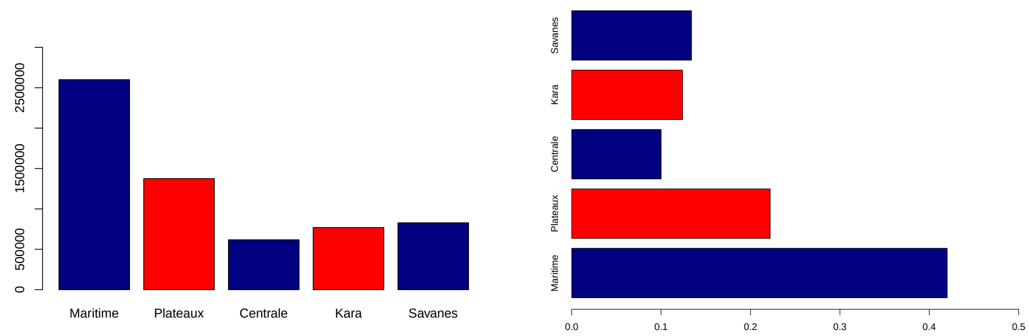
8.2.3 Variables qualitatives

8.2.3.1 Diagramme à barres

Le **diagramme à barres verticales** des effectifs ou des fréquences est une représentation graphique de la distribution d'une variable qualitative sur laquelle l'axe des abscisses correspond aux différentes modalités de la variable et l'axe des ordonnées aux effectifs ou fréquences associées.

Il est aussi possible de faire **diagramme à barres horizontales** des effectifs ou des fréquences. Dans ce cas, c'est plutôt l'axe des ordonnée qui correspond aux différentes modalités de la variable et l'axe des abscisses aux effectifs ou fréquences associées.

Exemple 8.11 La figure 8.4 donne les diagrammes à barres verticales des effectifs et à barres horizontales des fréquences associés au tableau 7.1.



(a) Diagramme à barres verticales (effectifs) (b) Diagramme à barres horizontales (fréquences)

FIGURE 8.4 – Diagrammes à barres verticales (effectifs) et à barres horizontales (fréquences) associés au tableau 7.1.

8.2.3.2 Diagramme à points

Le **diagramme à points** est une représentation graphique associée à une variable qualitative sur laquelle l'axe des ordonnées correspond aux différentes modalités et l'axe des abscisses aux effectifs ou fréquences associés.

Exemple 8.12 Le diagramme à points associé aux données du tableau 7.1 est donné par la figure 8.5.

8.2.3.3 Diagramme circulaire

Le **diagramme circulaire** est une représentation sous forme de cercle telle qu'à chacune des modalités est associée une portion circulaire du diagramme proportionnelle à sa fréquence. Ainsi, les angles (compris entre 0 et 360°) correspondants à chaque portion sont obtenus par la formule :

$$\alpha_i = 360 \times \frac{n_i}{n} = 360 \times f_i, \quad i = 1, \ldots, p. \tag{8.3}$$

Exemple 8.13 Reprenons les données du tableau 7.1 et complétons ce tableau par les angles α_i. On obtient le tableau suivant :

Le diagramme circulaire associé est donné par la figure 8.6.

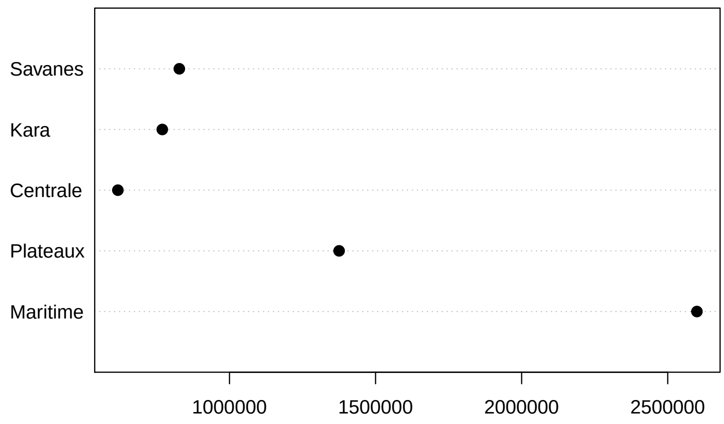


FIGURE 8.5 – Diagramme à points associé aux données du RGPH 2010.

TABLE 8.7 – Angles du diagramme circulaire associé aux données du RGPH 2010.

Région	Maritime*	Plateaux	Centrale	Kara	Savanes
f_i	0.420	0.222	0.100	0.124	0.134
α_i	151°	80°	36°	45°	48°

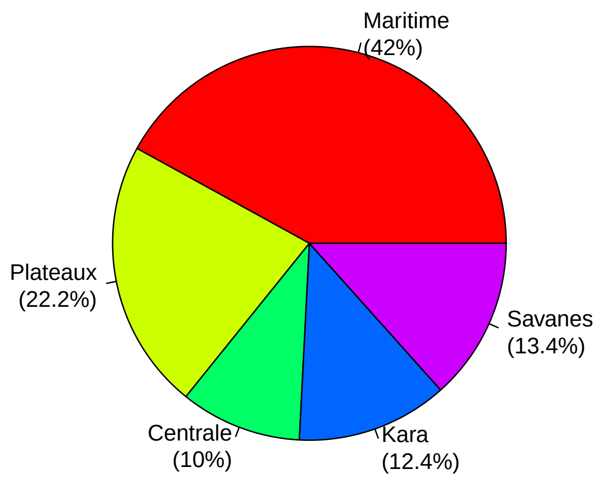


FIGURE 8.6 – Diagramme circulaire associé aux données du RGPH 2010.

8.3 CARACTÉRISTIQUES NUMÉRIQUES D'UNE DISTRIBUTION

Le tableau de données complété avec les fréquences et les représentations graphiques donne les premières informations sur le caractère et la population étudiés. Pour les caractères quantitatifs en particulier, il existe certaines quantités qui permettent de mieux caractériser la distribution statistique obtenue.

Remarque 8.14 Dans toute la suite de cette section, sauf mention contraire, nous considérons un caractère quantitatif X étudié sur une population de taille n. Les observations seront notées $x_1, \ldots, x_n$ et les modalités $x_1, \ldots, x_p$.

8.3.1 Les caractéristiques de position

Encore appelées **caractéristiques de tendance centrale**, elles donnent une idée de l'ordre de grandeur des valeurs constituant la série.

8.3.1.1 Mode et classe modale

Définition 8.15 Soit X le caractère étudié.

1) Si X est une variable quantitative discrète ou qualitative, on appelle **mode** de la série statistique obtenue et on note $\mathrm{Mo}(X)$, l'une des modalités $x_1, \ldots, x_p$ dont la fréquence est maximale.
2) Si X est une quantitative continue, une **classe modale**, $\mathrm{Mo}(X)$, est une classe de densité (rapport fréquence/amplitude) maximale. Un **mode** est alors le centre de l'une de ces classes.

Une distribution statistique est dite **unimodale** si elle a un seul mode et **multimodale** si elle en a plusieurs.

8.3.1.2 Moyenne arithmétique (distribution discrète)

Définition 8.16 Lorsque l'on dispose des observations sur chacun des n individus de la population sous la forme $x_1, \ldots, x_n$, la **moyenne arithmétique** est définie par :

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{8.4}$$

Lorsque la série statistique est donnée sous la forme $(x_1, n_1), \ldots, (x_p, n_p)$ en donnant les modalités et les effectifs correspondants, la moyenne arithmétique est définie par la somme pondérée suivante :

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{p} n_i x_i = \sum_{i=1}^{p} \frac{n_i}{n} x_i = \sum_{i=1}^{p} f_i x_i. \tag{8.5}$$

Proposition 8.17 *Si X et Y sont deux variables quantitatives et si a et b sont des réels alors on a :*

$$\overline{X+Y} = \overline{X} + \overline{Y}, \qquad \overline{aX+b} = a\overline{X} + b \qquad \text{et} \qquad \overline{aX} = a\overline{X}.$$

Remarque 8.18 La moyenne arithmétique est la mesure de tendance centrale des variables quantitatives la plus utilisée. Chaque observation a le même poids qui est $1/n$. La moyenne arithmétique est sensible aux valeurs extrêmes.

8.3.1.3 Moyenne arithmétique (distribution continue)

Pour les variables continues, on suppose que les observations d'une classe sont uniformément réparties à l'intérieur de ladite classe. Cela revient à supposer que chaque centre est une bonne approximation de la moyenne de sa classe. Dans l'approximation de la moyenne, on considère les centres des différentes classes :

$$\overline{X} \simeq \frac{n_1c_1 + \cdots + n_pc_p}{n} = \frac{1}{n}\sum_{i=1}^{p} n_ic_i. \tag{8.6}$$

Remarque 8.19 Généralement quand l'on parle de moyenne, on pense tout de suite à la moyenne arithmétique. Mais il existe d'autres types de moyennes (moyenne géométrique, moyenne harmonique, moyenne quadratique) qui peuvent être plus adaptés dans certains contextes particuliers. Ces autres types de moyennes ne sont pas abordés dans ce cours.

8.3.1.4 Médiane

Définition 8.20 La médiane est la valeur notée m_e qui partage la série des observations **classées par ordre croissant** en deux parties comprenant exactement le même nombre d'observations de part et d'autre de m_e.

Dans la pratique, le calcul de la médiane utilise le tableau des effectifs (ou fréquences) cumulés croissants. La médiane correspond à un effectif cumulé croissant de $n/2$ ou une fréquence cumulée croissante de 50%.

■ **Cas d'une distribution discrète**.

Considérons les observations classées par ordre croissant :

$$x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(n-1)} \leqslant x_{(n)}.$$

On distingue deux cas suivant la parité de l'effectif de la population n :

- Si n est impair, alors on peut écrire $n = 2k+1$ où k est un entier naturel non nul. On peut décomposer la série comme suit :

$$\underbrace{x_{(1)} \leqslant \cdots \leqslant x_{(k)}}_{k \text{ observations}} \leqslant x_{(k+1)} \leqslant \underbrace{x_{(k+2)} \leqslant \cdots \leqslant x_{(2k+1)}}_{k \text{ observations}}.$$

La médiane est donc la valeur

$$m_e = x_{(k+1)} = x_{(\frac{n+1}{2})}.$$

- Si n est pair, alors on peut écrire $n = 2k$ où k est un entier naturel non nul. On peut décomposer la série comme suit :

$$\underbrace{x_{(1)} \leqslant \cdots \leqslant x_{(k)}}_{k \text{ observations}} \leqslant \underbrace{x_{(k+1)} \leqslant \cdots \leqslant x_{(2k)}}_{k \text{ observations}}.$$

La médiane peut être prise égale à toute valeur comprise entre $x_{(k)}$ et $x_{(k+1)}$. Il est d'usage de choisir la moyenne de ces deux valeurs comme médiane :

$$m_e = \frac{x_{(k)} + x_{(k+1)}}{2}.$$

Exemple 8.21

1) Considérons les dix observations suivantes : 1, 1, 1, 2, 2, 2, 2, 3, 3, 4. Le nombre 10 est pair et sa moitié vaut 5. La médiane est la moyenne de la 5e et de la 6e observations soit $m_e = \frac{2+2}{2} = 2$.
2) Considérons les sept observations suivantes : 0, 0, 1, 1, 2, 3, 4. Le nombre 7 est impair et $\frac{7+1}{2} = 4$. La médiane est la 4e observation soit $m_e = 1$.

■ **Cas d'une distribution continue**.

Considérons une distribution continue $([a_0, a_1[, n_1), \ldots, ([a_{p-1}, a_p[, n_p)$. Le calcul de la médiane se fait à l'aide des effectifs cumulés croissants ou des fréquences cumulées croissantes, la médiane étant la valeur correspond à l'effectif cumulé croissant $n/2$ ou à la fréquence cumulée croissante 50%. On distingue deux cas :

- S'il existe une classe $C_i = [a_{i-1}, a_i[$ telle que $n_{ic} = n/2$ ou $f_{ic} = 50\%$, alors $m_e = a_i$;
- Sinon il existe deux classes adjacentes $C_i = [a_{i-1}, a_i[$ et $C_{i+1} = [a_i, a_{i+1}[$ telles que $n_{ic} < n/2 < n_{(i+1)c}$ ou $f_{ic} < 50\% < f_{(i+1)c}$. Ainsi $a_i < m_e < a_{i+1}$ et on calcule une valeur approximative de m_e par interpolation linéaire. Pour faire ce calcul, on suppose que **la distribution est uniforme à l'intérieur de chaque classe**.

L'interpolation linéaire consiste à dire que les trois points $A(a_i, n_{ic})$, $B(a_{i+1}, n_{(i+1)c})$ et $M(m_e, n/2)$ sont approximativement alignés.

En utilisant le théorème de Thalès, on obtient

$$\frac{\frac{n}{2} - n_{ic}}{n_{(i+1)c} - n_{ic}} = \frac{m_e - a_i}{a_{i+1} - a_i}$$

d'où

$$m_e = a_i + (a_{i+1} - a_i)\frac{\frac{n}{2} - n_{ic}}{n_{(i+1)c} - n_{ic}}. \tag{8.7}$$

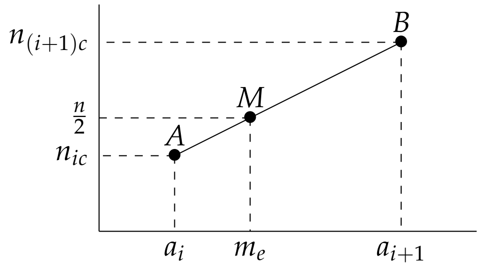


FIGURE 8.7 – Illustration de l'interpolation linéaire pour l'approximation de la médiane

Cette dernière formule peut être réécrite en utilisant les fréquences :

$$m_e = a_i + (a_{i+1} - a_i)\frac{\frac{1}{2} - f_{ic}}{f_{(i+1)c} - f_{ic}}. \tag{8.8}$$

Remarque 8.22 Pour les distributions continues, la médiane est l'abscisse du point d'intersection des polygones des fréquences cumulées croissantes et décroissantes.

8.3.1.5 Les quartiles

Définition 8.23 Considérons une série statistique **classée par ordre croissant**. Les quartiles sont trois valeurs notées Q_1, Q_2 et Q_3 qui partagent la distribution en quatre parties comprenant exactement le même nombre d'observations. Plus précisément, le **premier quartile** (resp. **le second quartile**, resp. le **troisième quartile**) noté Q_1 (resp. Q_2, resp. Q_3) est la valeur correspondant à une fréquence cumulée croissante 25% (resp. 50%, resp. 75%) ou à l'effectif cumulé croissant $\frac{n}{4}$ (resp. $\frac{n}{2}$, resp. $\frac{3n}{4}$).

Remarque 8.24 En particulier, le second quartile est confondu à la médiane.

■ **Cas d'une distribution discrète**.

Ici aussi, supposons les observations classées par ordre croissant :

$$x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(n-1)} \leqslant x_{(n)}.$$

On distingue deux cas :

- Si n est multiple de 4, alors on peut écrire $n = 4k$ où k est un entier naturel non nul. Le premier quartile peut être pris égal à toute valeur comprise entre $x_{(k)}$ et $x_{(k+1)}$. Il est d'usage de choisir la moyenne de ces deux valeurs comme médiane :

$$Q_1 = \frac{x_{(k)} + x_{(k+1)}}{2}.$$

  Idem pour le troisième quartile :

$$Q_3 = \frac{x_{(3k)} + x_{(3k+1)}}{2}.$$

- Si n n'est pas multiple de 4, alors on note k et k' les parties entières respectives de $n/4$ et $3n/4$. Les premier et troisième quartiles sont définis par : $Q_1 = x_{(k+1)}$ et $Q_3 = x_{(k'+1)}$.

Exemple 8.25

1) Considérons les dix observations suivantes : $1, 1, 1, 2, 2, 2, 2, 3, 3, 4$. Le nombre 10 n'est pas multiple de 4. On a $10/4 = 2.5$ et $3 \times 10/4 = 7.5$. On a donc $Q_1 = x_{(3)} = 1$ et $Q_3 = x_{(8)} = 3$.
2) Considérons les huit observations suivantes : $0, 0, 1, 1, 2, 3, 4, 5$. Le nombre 8 est multiple de 4 et $\frac{8}{4} = 2$. Le premier quartile est la moyenne de la 2e et de la 3e observations soit $Q_1 = 0.5$. On vérifie de même que $Q_3 = 3.5$.

■ **Cas d'une distribution continue**.

Le calcul du quartile Q_1 (resp. Q_3) se fait en utilisant la même méthode que celle décrite pour la médiane en prenant soin de remplacer la fréquence cumulée croissante 50% par 25% (resp. 75%) et l'effectif cumulé croissant $n/2$ par $n/4$ (resp. $3n/4$).

8.3.1.6 Les quantiles

Définition 8.26 Soit X une variable quantitative étudiée sur une population et soit $\alpha \in [0, 1]$. On appelle **quantile** (ou **fractile**) d'**ordre** α, la valeur x_α de la variable X telle que la proportion d'observations inférieures ou égales à x_α soit égale à α.

Les quantiles les plus usuels sont la médiane et les quartiles. La médiane est le quantile d'ordre 50%, les quartiles Q_1 et Q_3 sont les quantiles d'ordres respectifs 25% et 75%.

8.3.2 Les caractéristiques de dispersion

Les caractéristiques de dispersion informent sur la répartition des modalités du caractère étudié autour des caractéristiques de tendance centrale.

8.3.2.1 Variance et écart-type

■ **Cas d'une distribution discrète**.

Définition 8.27 Lorsque l'on dispose des observations $x_1, \ldots, x_n$, la **variance** (ou écart quadratique moyen) est définie par :

$$\mathbb{V}\mathrm{ar}(X) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{X})^2. \tag{8.9}$$

Lorsque la série statistique est donnée sous la forme $(x_1, n_1), \ldots, (x_p, n_p)$ en donnant les modalités et les effectifs correspondants, la moyenne quadratique est définie par :

$$\mathbb{V}\mathrm{ar}(X) = \frac{1}{n} \sum_{i=1}^{p} n_i (x_i - \overline{X})^2. \tag{8.10}$$

La racine carrée de la variance notée σ_X est appelée **écart-type** de X :

$$\sigma_X = \sqrt{\mathbb{V}\mathrm{ar}(X)}. \tag{8.11}$$

Dans la pratique, on utilise la formule suivante dite formule de König-Huygens :

Théorème 8.28 (Formule de König-Huygens) *On a :*

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{X})^2 = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right) - \overline{X}^2 = \overline{X^2} - (\overline{X})^2 \tag{8.12}$$

ou encore

$$\frac{1}{n} \sum_{i=1}^{p} n_i (x_i - \overline{X})^2 = \left(\frac{1}{n} \sum_{i=1}^{p} n_i x_i^2 \right) - \overline{X}^2 = \overline{X^2} - (\overline{X})^2. \tag{8.13}$$

où $\overline{X^2}$ désigne la moyenne des carrés. En d'autres termes, la variance est égale à la moyenne (arithmétique) des carrées moins le carré de la moyenne.

■ **Cas d'une distribution continue**.

Lorsqu'on ne dispose que des classes, il n'est pas possible de trouver une valeur exacte de la variance. En supposant que les données sont uniformément distribuées à l'intérieur de chaque classe, on obtient une approximation de la variance par la formule

$$\mathbb{V}\mathrm{ar}(X) \simeq \frac{1}{n} \sum_{i=1}^{p} n_i (c_i - \overline{X})^2$$

où $c_1, \ldots, c_p$ sont les centres des différentes classes.

Remarque 8.29 La variance et l'écart-type sont les caractéristiques de dispersion les plus utilisée. L'écart-type présente l'avantage de s'exprimer dans la même unité que la variable étudiée.

Proposition 8.30 (Propriétés de la variance)

1) La variance est toujours un nombre réel positif (puisque c'est une somme de carrés).

2) La variance est nulle si, et seulement si, X possède une seule valeur.

3) Si X est une variable quantitative et si a et b sont deux réels, on a

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

8.3.2.2 Autres caractéristiques de dispersion

- **Etendue** : C'est la différence entre la plus grande et la plus petite des valeurs prises par le caractère : $e(X) = \max(X) - \min(X)$.

- **Etendue interquartile** : C'est le réel noté EIQ et défini par $\text{EIQ} = Q_3 - Q_1$.

- **Intervalle interquartile** : C'est l'intervalle $I = [Q_1, Q_3]$. Il comporte 50% des observations équitablement réparties autour de la médiane Q_2. Un intervalle I petit implique une forte concentration des valeurs du caractère autour de la médiane.

- **Ecart absolu moyen** : C'est le réel défini par

$$e_m = \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{X}|$$

lorsque l'on dispose des observations $x_1, \ldots, x_n$ et par

$$e_m = \frac{1}{n}\sum_{i=1}^{p}n_i|x_i - \overline{X}|.$$

lorsque la série statistique est donnée sous la forme $(x_1, n_1), \ldots, (x_p, n_p)$ avec les modalités et les effectifs. L'écart moyen donne l'ordre de grandeur des déviations autour de la moyenne.

- **Coefficient de variation** : Il est défini pour les variables positives comme le rapport de l'écart-type avec la moyenne arithmétique :

$$\text{CV} = \frac{\sigma_X}{\overline{X}}.$$

Le coefficient de variation est un nombre sans unité et est invariable par changement d'échelle sur les données. Il permet donc de comparer les dispersions du caractère sur des populations ayant des valeurs moyennes très différentes ou de comparer des variables quantitatives différentes même si les données ne sont pas exprimées dans les mêmes unités. Un CV petit indique une distribution du caractère moins dispersée.

Exemple 8.31 Reprenons les données du tableau 8.2.

x_i	0	1	2	3	4
n_i	12	20	5	2	1

Calculons la moyenne, l'écart-type, la médiane, les quartiles Q_1 et Q_3.

- Moyenne :

$$\overline{X} = \frac{1}{40}(12 \times 0 + 20 \times 1 + 5 \times 2 + 2 \times 3 + 1 \times 4) = 1.$$

- Ecart-type :

$$\mathbb{V}\mathrm{ar}(X) = \frac{1}{40}(12 \times 0 + 20 \times 1 + 5 \times 4 + 2 \times 9 + 1 \times 16) - 1 = 0.85$$

d'où $\sigma_X = \sqrt{0.85} \simeq 0.9219544$.

- Médiane : rangeons les données dans l'ordre croissant. Comme le nombre d'observations est pair ($n = 40$),

$$m_e = \frac{x_{(20)} + x_{(21)}}{2} = 1.$$

- Quartiles Q_1 et Q_3 : $n = 40$ est multiple de 4 et $k = \frac{n}{4} = 10$. Ainsi,

$$Q_1 = \frac{x_{(10)} + x_{(11)}}{2} = 0 \quad \text{et} \quad Q_3 = \frac{x_{(30)} + x_{(31)}}{2} = 1.$$

Exemple 8.32 Reprenons les données du tableau 8.3. Complétons le tableau avec quelques données utiles et calculons l'étendue, la moyenne, l'écart-type, la médiane, les quartiles Q_1 et Q_3.

C_i	$[20,30[$	$[30,35[$	$[35,40[$	$[40,45[$	$[45,50[$	$[50,55[$	$[55,60[$	$[60,65[$
c_i	25	32.5	37.5	42.5	47.5	52.5	57.5	62.5
n_i	6	6	5	7	10	4	7	3
n_{ic}	6	12	17	24	34	38	45	48

- Etendue : $e(X) = 65 - 20 = 45$.
- Moyenne :

$$\overline{X} \simeq \frac{1}{48}\sum_{i=1}^{8} n_i c_i \simeq 43.85417.$$

- Ecart-type : $\mathbb{V}\mathrm{ar}(X) \simeq 122.9058$ d'où $\sigma_X \simeq 11.08629$.
- Médiane : elle correspond à l'effectif cumulé croissant 24. Il n'y a pas besoin d'interpolation linéaire car l'effectif cumulé croissant 24 est associé à 45. Ainsi, $m_e = 45$ ans.
- Quartile Q_1 : Q_1 correspond à un effectif cumulé croissant de $48/4 = 12$ et comme pour la médiane, on lit directement $Q_1 = 35$ ans.
- Le quartile Q_3 correspond à un effectif cumulé croissant de $3 \times \frac{48}{4} = 36$. Cette valeur n'apparaît pas dans les effectifs cumulés croissants mais on voit que $34 < 36 < 38$ d'où on déduit que $50 < Q_3 < 55$. On utilise la formule d'interpolation linéaire :

$$Q_3 = 50 + (55 - 50) \times \frac{36 - 34}{38 - 34} = 52.5.$$

8.4 EXERCICES

Exercice 8.1 La répartition des élèves d'un lycée en fonction de la langue vivante étudiée est donnée par le tableau suivant :

Langue	Anglais	Allemand	Espagnol	Italien	Autres	Total
Effectifs	934	351	205	69	41	1600

1) Préciser la population étudiée, l'unité statistique, le caractère, sa nature et ses modalités.
2) Compléter le tableau avec les fréquences.
3) Représenter cette distribution par un diagramme à secteurs.
4) Préciser le mode du caractère étudié.
5) Quel est le pourcentage d'élèves qui n'étudient pas l'anglais ?

Exercice 8.2 Un institut de statistique a réalisé une enquête sur le nombre de salariés de 40 entreprises de Lomé-la-belle. Les résultats suivants rangés dans l'ordre croissant sont les suivants :

20	22	24	30	32	36	37	39	40	41
43	44	45	45	47	48	50	51	51	52
52	53	53	55	56	58	59	59	61	62
63	64	66	75	76	79	82	86	90	99

1) Définir la population, l'unité statistique, le caractère étudié et sa nature.
2) Etablir la distribution des entreprises selon le nombre de salariés. Pour ce faire, on définit cinq (5) classes de la forme $[a_i, a_{i+1}[$ d'amplitudes respectives 20, 10, 10, 20, 20. La valeur minimale est 20 et la valeur maximale est 100. On complètera le tableau obtenu en y ajoutant les fréquences et les fréquences cumulées.
3) Tracer l'histogramme des fréquences de cette distribution.
4) Calculer le mode, la médiane, la moyenne et l'écart-type du nombre de salariés.
5) Calculer à nouveau la médiane et la moyenne mais cette fois-ci en utilisant les données brutes et non les classes. Que remarquez-vous ?

Exercice 8.3 Le tableau suivant donne la répartition du personnel d'une société suivant le salaire mensuel en kFCFA (milliers de FCFA) :

Salaire	[150,160[	[160,170[	[170,180[	[180,190[	[190,200[	[200,250[	[250,300[
Effectif	5	13	15	18	14	4	5

1) Calculer la moyenne, la variance, l'écart-type, la médiane et les quartiles Q_1 et Q_3.
2) Déterminer le pourcentage de personnel dont le salaire est inférieur à 200 kFCFA.

Exercice 8.4 Dans un laboratoire pharmaceutique, une machine automatique fabrique en grande quantité des suppositoires contenant du paracétamol. On veut contrôler la qualité de la fabrication sur une période donnée. Dans ce but, pendant le fonctionnement de la machine, on prélève de temps à autre un suppositoire dont on mesure la masse X (en mg) de paracétamol qu'il contient. Les résultats des mesures de l'échantillon prélevé sont donnés dans le tableau suivant :

Masse (mg)	$[145, 155[$	$[155, 165[$	$[165, 175[$	$[175, 185[$	$[185, 195[$
Effectifs	7	30	43	16	4

Faire une étude graphique et un résumé statistique de la masse de paracétamol des suppositoires fabriqués dans ce laboratoire.

Exercice 8.5 On observe l'arrivée des clients à un bureau de poste pendant un intervalle de temps donné. En répétant 100 fois l'observation, on obtient les résultats suivants :

Nombre d'arrivées	1	2	3	4	5	6
Effectif	15	25	26	20	7	7

Calculer la moyenne, la variance, l'écart-type, le coefficient de variation, la médiane, les quartiles Q_1 et Q_3 et l'intervalle inter-quartile.

CHAPITRE 9

SÉRIES STATISTIQUES À DEUX VARIABLES

Résumé – Ce chapitre présente les méthodes usuelles de traitement des séries statistiques à deux variables. L'objectif est de savoir faire une étude statistique descriptive sur les données observées sur une population dans un cadre bivarié et plus spécifiquement, de savoir faire des prévisions à l'aide d'un ajustement linéaire ou non linéaire.

Sommaire

9.1 GÉNÉRALITÉS

Dans une enquête statistique, il est rare de ne s'intéresser qu'à une seule variable statistique. Une « intuition » des spécialistes d'un domaine peut conduire à observer deux ou plusieurs variables chez une même population et à rechercher l'existence d'un lien de dépendance entre ces variables. Dans ce chapitre, nous nous focalisons sur le cas de deux variables.

Définition 9.1 Soient Ω une population de taille n, X et Y deux variables statistiques observées sur Ω. On appelle série statistique à deux variables X et Y (ou à deux dimensions ou série statistique double), l'application qui à chaque individu $\omega_i \in \Omega$ $(i = 1, \ldots, n)$ associe le couple (x_i, y_i) des valeurs prises respectivement par X et Y pour ledit individu.

L'ensemble des couples (x_i, y_i), $i = 1, \ldots, n$ est appelée **distribution statistique** du couple (X, Y). Cette distribution statistique du couple (X, Y) est souvent présentée sous deux formes : (1) **données brutes** ou **données non groupées** et (2) **données groupées** ou **tableau de contingence**).

- Le tableau de **données brutes** est un tableau à deux lignes de la forme

X	x_1	$\cdots$	x_i	$\cdots$	x_n
Y	y_1	$\cdots$	y_i	$\cdots$	y_n

ou à trois lignes ayant la forme

Individu	1	$\cdots$	i	$\cdots$	n
X	x_1	$\cdots$	x_i	$\cdots$	x_n
Y	y_1	$\cdots$	y_i	$\cdots$	y_n

- Le **tableau de contingence** est un tableau à double entrée de la forme

X \ Y	y_1	$\cdots$	y_j	$\cdots$	y_q	$n_{i\bullet}$
x_1	n_{11}	$\cdots$	n_{1j}	$\cdots$	n_{1q}	$n_{1\bullet}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	n_{i1}	$\cdots$	n_{ij}	$\cdots$	n_{iq}	$n_{i\bullet}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_p	n_{p1}	$\cdots$	n_{pj}	$\cdots$	n_{pq}	$n_{p\bullet}$
$n_{\bullet j}$	$n_{\bullet 1}$	$\cdots$	$n_{\bullet j}$	$\cdots$	$n_{\bullet q}$	n

où pour tous $i = 1, \ldots, p$ et $j = 1, \ldots, q$, n_{ij} est le nombre d'individus vérifiant simultanément $X = x_i$ et $Y = y_j$,

$$n_{i\bullet} = \sum_{j=1}^{q} n_{ij} = \text{nombre d'individus pour lesquels } X = x_i,$$

$$n_{\bullet j} = \sum_{i=1}^{p} n_{ij} = \text{nombre d'individus pour lesquels } Y = y_j$$

et

$$n = \sum_{i=1}^{p} \sum_{j=1}^{q} n_{ij} = \sum_{i=1}^{p} n_{i\bullet} = \sum_{j=1}^{q} n_{\bullet j}.$$

La dernière colonne (en rouge) et la dernière ligne (en bleu) du tableau (appelées **marges**) contiennent respectivement les effectifs totaux par ligne et par colonne. Elles donnent respectivement la **distribution marginale** de X définie par le tableau

X	x_1	x_2	$\cdots$	x_p	Total
Effectif $n_{i\bullet}$	$n_{1\bullet}$	$n_{2\bullet}$	$\cdots$	$n_{p\bullet}$	n

et celle de Y définie par :

Y	y_1	y_2	$\cdots$	y_q	Total
Effectif $n_{\bullet j}$	$n_{\bullet 1}$	$n_{\bullet 2}$	$\cdots$	$n_{\bullet q}$	n

Exemple 9.2 On mesure sur la même parcelle de terrain le rendement Y en quintaux par hectare (q/ha) en fonction de la quantité X d'engrais répartie sur la parcelle mesurée en kg/ha. Les données sont consignées dans le tableau suivant :

Quantité x_i	100	200	300	400	500	600	700
Rendement y_i	40	50	50	70	65	65	80

Exemple 9.3 Dans une université, on a enquêté auprès de 5375 étudiants n'ayant aucun lien de parenté entre eux, sur leurs habitudes tabagiques (variable Y) et celles de leurs parents (variable X). Les données obtenues sont consignées dans le tableau suivant :

Statut de l'étudiant / Nombre de parents fumeurs	Fumeur	Non fumeur	Total
Deux	400	1380	1780
Un seul	416	1823	2239
Aucun	188	1168	1356
Total	1004	4371	5375

Les distributions marginales de X et Y sont respectivement données par les tableaux

Nombre de parents fumeurs	Deux	Un seul	Aucun	Total
Effectif	1780	2239	1356	5375

et

Statut de l'étudiant	Fumeur	Non fumeur	Total
Effectif	1004	4371	5375

9.2 Caractéristiques numériques

9.2.1 Cas des données non groupées

Les moyennes et variances des variables X et Y se calculent par les formules usuelles comme suit :

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\mathbb{V}\mathrm{ar}(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\overline{X})^2, \quad \mathbb{V}\mathrm{ar}(Y) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{Y})^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - (\overline{Y})^2,$$

Les écarts types sont respectivement définies par

$$\sigma_X = \sqrt{\mathbb{V}\mathrm{ar}(X)} \qquad \text{et} \qquad \sigma_Y = \sqrt{\mathbb{V}\mathrm{ar}(Y)}.$$

Les valeurs caractéristiques de la liaison entre les variables X et Y sont

- la **covariance** de X et Y définie par

$$\mathrm{Cov}(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y}) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{XY}.$$

- et le **coefficient de corrélation linéaire** entre X et Y défini par :

$$\rho_{XY} = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\mathbb{V}\mathrm{ar}(X)}\sqrt{\mathbb{V}\mathrm{ar}(Y)}} = \frac{\operatorname{Cov}(X,Y)}{\sigma_X\sigma_Y} \in [-1,1].$$

Exemple 9.2 (suite) On vérifie que $\overline{X} = 400$, $\overline{Y} = 60$, $\frac{1}{n}\sum_{i=1}^{n} x_i y_i = 26357.143$ donc $\operatorname{Cov}(X,Y) = 2357.143$. De plus, $\mathbb{V}\mathrm{ar}(X) = 40000$, $\mathbb{V}\mathrm{ar}(Y) = 164.2857$ donc $\sigma_X = 200$ et $\sigma_Y = 12.8174$. Finalement, $\rho_{XY} \simeq 0.92$.

9.2.2 Cas des tableaux de contingence (données groupées)

Dans ce cas aussi, les moyennes, variances et écarts types des variables X et Y se calculent par les formules usuelles comme suit :

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{p} n_{i\bullet} x_i, \qquad \overline{Y} = \frac{1}{n}\sum_{j=1}^{q} y_{\bullet j},$$

$$\mathbb{V}\mathrm{ar}(X) = \frac{1}{n}\sum_{i=1}^{p} n_{i\bullet}(x_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{p} n_{i\bullet} x_i^2 - (\overline{X})^2,$$

$$\mathbb{V}\mathrm{ar}(Y) = \frac{1}{n}\sum_{j=1}^{q} n_{\bullet j}(y_j - \overline{Y})^2 = \frac{1}{n}\sum_{j=1}^{q} n_{\bullet j} y_j^2 - (\overline{Y})^2,$$

$$\sigma_X = \sqrt{\mathbb{V}\mathrm{ar}(X)} \quad \text{et} \quad \sigma_Y = \sqrt{\mathbb{V}\mathrm{ar}(Y)}.$$

La covariance est calculée à l'aide de la formule

$$\operatorname{Cov}(X,Y) = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}(x_i - \overline{X})(y_j - \overline{Y}) = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij} x_i y_j - \overline{X}\,\overline{Y}.$$

Remarque 9.4 Lorsque les modalités des variables X et/ou Y sont des intervalles (ou classes), les x_i et/ou y_j sont remplacés par les centres des classes.

L'analyse descriptive d'un tableau de contingence est facilitée par le calcul de deux tableaux des fréquences conditionnelles appelés **profils-lignes** et **profils-colonnes**.

Définition 9.5 On appelle :

- *profils-lignes*, le tableau obtenu en divisant chaque ligne par la somme de ladite ligne c-est-à-dire le tableau formé des fréquences

$$\frac{n_{ij}}{n_{i\bullet}} = \textit{proportion d'individus de la modalité } x_i \textit{ ayant la modalité } y_j;$$

- *profils-colonnes*, le tableau obtenu en divisant chaque colonne par la somme de ladite colonne c-est-à-dire le tableau formé des fréquences

$$\frac{n_{ij}}{n_{\bullet j}} = \textit{proportion d'individus de la modalité } y_j \textit{ ayant la modalité } x_i.$$

Remarque 9.6 Dans le tableau des profils-lignes, la somme de chaque ligne vaut 1 et la somme des colonnes n'a pas de sens et doit donc être évitée. De même, dans le tableau des profils-colonnes, la somme de chaque colonne vaut 1 et la somme des lignes n'a pas de sens et doit aussi être évitée.

Exemple 9.3 (suite)

- Le tableau des profils-lignes arrondis à trois chiffres après la virgule est donné par :

Statut de l'étudiant / Nombre de parents fumeurs	Fumeur	Non fumeur	Total
Deux	0.2247	0.7753	1
Un seul	0.1858	0.8142	1
Aucun	0.1386	0.8614	1

  On peut interpréter la première ligne comme suit : parmi les étudiants dont les deux parents sont fumeurs, il y en a 22.47% qui sont fumeurs et 77.53% qui sont non-fumeurs.

- Le tableau des profils-colonnes arrondis à quatre chiffres après la virgule est donné par :

Statut de l'étudiant / Nombre de parents fumeurs	Fumeur	Non fumeur
Deux	0.3984	0.3157
Un seul	0.4143	0.4171
Aucun	0.1873	0.2672
Total	1	1

  On peut interpréter la première colonne comme suit : parmi les étudiants fumeurs, 39.84% ont les deux parents qui sont fumeurs, 41.43% ont un seul parent fumeur et 18.73% n'ont aucun parent fumeur.

9.3 Représentation graphique

On représente souvent un tableau de données brutes (ou non groupées) par un nuage de points dans un repère orthogonal du plan. Dand le cas d'un tableau de contingence, il est plus utile de représenter les profils-lignes et/ou profils-colonnes par un diagramme à barres juxtaposées.

Exemple 9.2 (suite) Le nuage de points associé aux données de l'exemple 9.2 est donné par la figure 9.1 (page 106).

Exemple 9.3 (suite) Les diagrammes à barres juxtaposées associés aux profils-lignes et profils-colonnes de l'exemple 9.3 sont donnés par la figure 9.2 (page 106).

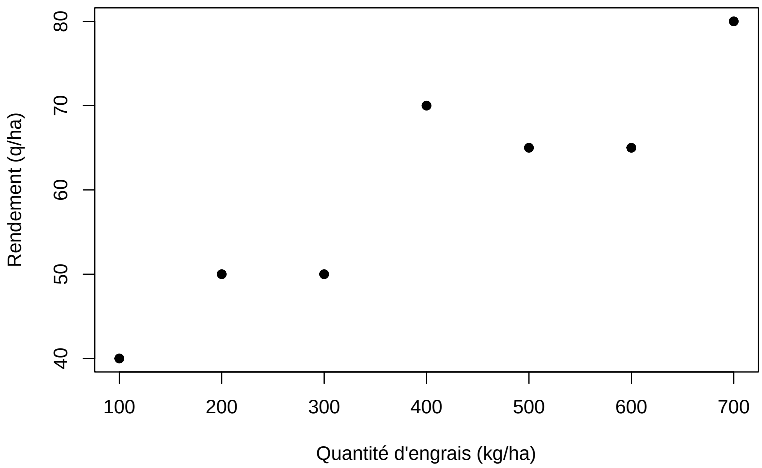
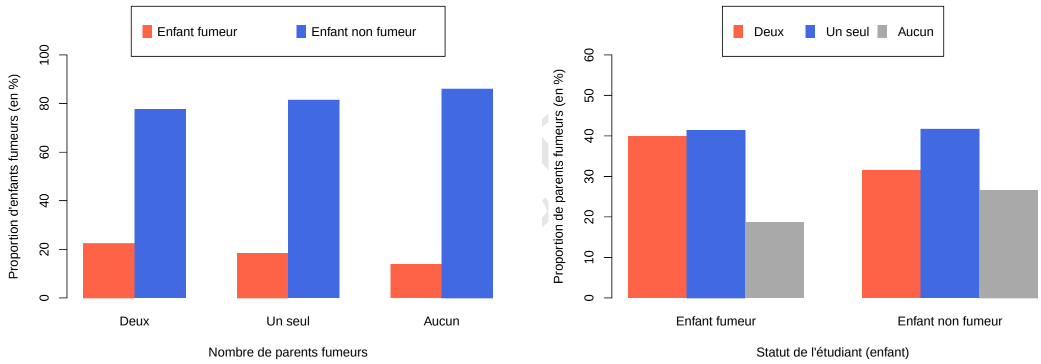


FIGURE 9.1 – Nuage de points des données du rendement



(a) Profils-lignes (b) Profils-colonnes

FIGURE 9.2 – Diagrammes à barres juxtaposées associés aux profils-lignes et profils-colonnes de l'exemple 9.3

9.4 AJUSTEMENT LINÉAIRE

9.4.1 Notion d'ajustement

L'ajustement consiste à chercher une relation de la forme

$$Y = f(X) \tag{9.1}$$

de sorte que la courbe représentative de f passe le « plus près possible » de tous les points du nuage. Dans ce cas, la variable X est appelée **variable explicative** ou **variable endogène** et la variable Y est appelée **variable expliquée** ou **variable exogène**.

Lorsque $f(x) = ax + b$ avec $a, b \in \mathbb{R}$, on parle d'**ajustement linéaire**. Sinon, on parle d'ajustement non linéaire.

L'ajustement linéaire est très utilisé dans la pratique non seulement à cause de sa simplicité mais aussi certains ajustements non linéaires (que nous verrons dans la suite) peuvent être ramenés à un ajustement linéaire au moyen d'une transformation mathématique et/ou d'un changement de variable.

9.4.2 Méthode des moindres carrés ordinaires (MCO)

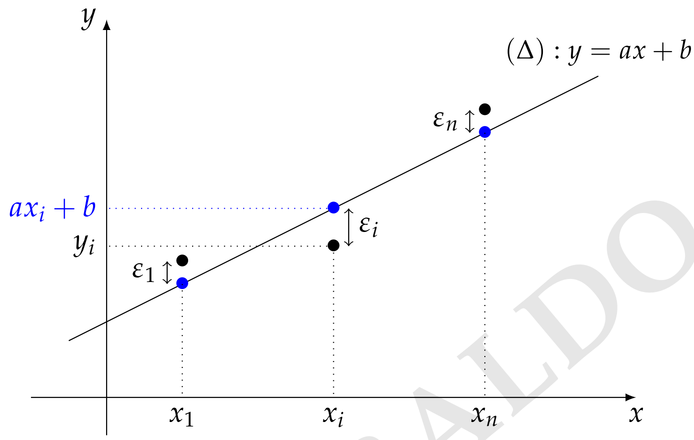


FIGURE 9.3 – Nuage de points et MCO

On cherche la droite (Δ) d'équation $y = ax + b$ (a, $b \in \mathbb{R}$) qui passe le plus près possible de tous les points du nuage.

Définition 9.7 Pour tout $i = 1, \ldots, n$, on appelle :

- **valeur ajustée** de y_i, la quantité $\hat{y}_i = ax_i + b$;
- **résidu** en i, la quantité $\varepsilon_i = y_i - \hat{y}_i = y_i - ax_i - b$.

Soit $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ le vecteur des résidus. Le principe de la méthode des moindres carrés ordinaires (MCO) consiste à rechercher les réels a et b qui rendent minimum la somme des carrés des résidus (SCR)

$$\|\varepsilon\|^2 = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \tag{9.2}$$

qui est une fonction de a et b. On montre le théorème suivant.

Théorème 9.8 *Le couple (a, b) permettant de minimiser la SCR $\|\varepsilon\|^2$ est donné par*

$$a = \frac{\text{Cov}(X, Y)}{\mathbb{V}\text{ar}(X)} \qquad \textit{et} \qquad b = \overline{Y} - a\overline{X}. \tag{9.3}$$

Définition 9.9 La droite d'équation $y = ax + b$, où a et b sont définis par l'équation (9.3), est appelée **droite de régression de Y en X**. Elle est notée $D_{Y/X}$.

Exemple 9.2 (suite) On continue avec l'exemple 9.2.

- En arrondissant à 5 cinq chiffres après la virgule, on obtient l'équation suivante pour la droite de régression de Y en X :

$$(D_{Y/X}) : y = 0.05893\, x + 36.42857.$$

- L'ajustement de la droite de régression au nuage de points est donné par la figure 9.4 (page 108).

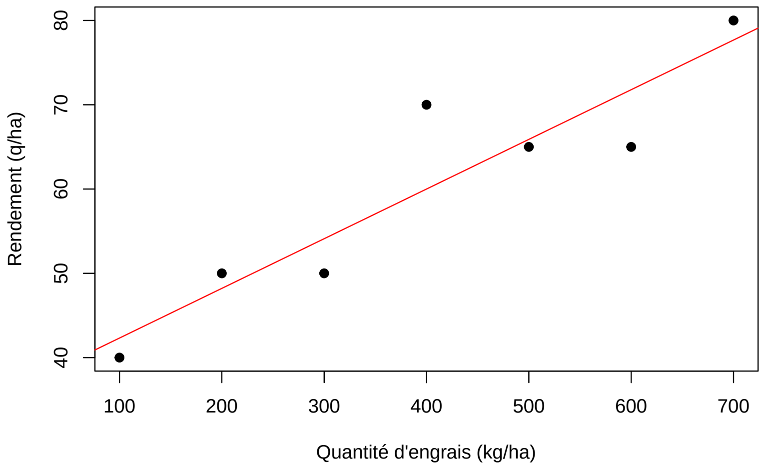


FIGURE 9.4 – Ajustement de la droite de régression sur le nuage de points des données du rendement

- On peut prévoir le rendement pour une quantité d'engrais de 350 kg/ha :

$$\hat{y}(350) = 0{,}058\,93 \times 350 + 36{,}428\,57 \approx 57 \text{ q/ha}.$$

Remarque 9.10 Pour obtenir la droite de régression de X en fonction de Y, il suffit de permuter les rôles des deux variables c'est-à-dire que l'on a

$$(D_{X/Y}) : x = a'y + b', \tag{9.4}$$

où

$$a' = \frac{\mathrm{Cov}(X,Y)}{\mathbb{V}\mathrm{ar}(Y)} \quad \text{et} \quad b' = \overline{X} - a'\overline{Y}.$$

Remarque 9.11 Pour que la droite de régression $(D_{X/Y})$ puisse être déduite de $(D_{Y/X})$, il faudrait aussi que l'on ait

$$x = \frac{y}{a} - \frac{b}{a} = a'y + b'$$

c'est-à-dire $a' = \frac{1}{a}$ (puisque cette relation entraîne $b' = -b/a$). On aurait donc

$$\frac{\mathrm{Cov}(X,Y)}{\mathbb{V}\mathrm{ar}(Y)} = \frac{\mathbb{V}\mathrm{ar}(X)}{\mathrm{Cov}(X,Y)}$$

soit
$$\frac{(\mathrm{Cov}(X,Y))^2}{\mathbb{V}\mathrm{ar}(X)\mathbb{V}\mathrm{ar}(Y)} = 1$$
soit encore $\rho_{XY}^2 = 1$.

9.4.3 Qualité de l'ajustement par MCO

Comment sait-on si un ajustement linéaire est acceptable ou non ? Pour répondre rigoureusement à cette question, nous avons besoin de certaines notions et techniques de statistique inférentielle qui dépassent largement le cadre de ce cours. Toutefois, nous donnons quelques indicateurs pouvant permettre de répondre à cette question.

Définition 9.12 On appelle :

- **variance totale**, la quantité
$$V_T = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{Y})^2 = \mathbb{V}\mathrm{ar}(Y);$$
- **variance expliquée** (par le modèle), la quantité
$$V_E = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - \overline{Y})^2;$$
- **variance résiduelle**, la quantité
$$V_R = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2.$$

Théorème 9.13 *Les trois variances définies plus haut vérifient la relation suivante*
$$V_T = V_E + V_R \tag{9.5}$$
*appelée **équation d'analyse de la variance**.*

La qualité de l'ajustement est jugée par le **coefficient de détermination** défini comme suit.

Définition 9.14 On appelle *coefficient de détermination* de la régression, le coefficient noté R^2 et défini par :
$$R^2 = \frac{V_E}{V_T}.$$

Dans la pratique, pour calculer R^2, on utilise plutôt le théorème suivant.

Théorème 9.15 *Le coefficient de détermination est le carré du coefficient de corrélation linéaire c'est-à-dire :*
$$R^2 = (\rho_{XY})^2.$$

Par la formule (9.5) ou par le fait que $-1 \leqslant \rho_{XY} \leqslant 1$, il est clair que

$$0 \leqslant R^2 \leqslant 1.$$

Plus R^2 est grand (proche de 1), plus la variance expliquée est grande et plus la variance résiduelle est faible. En d'autres termes, plus R^2 est proche de 1, plus on peut dire que la droite de régression s'ajuste « bien » aux données observées. Le coefficient R^2 mesure donc la proportion de la variabilité (ou la proportion de l'information) contenue dans les données que la droite de régression a su expliquer. Il mesure donc aussi l'intensité de la liaison linéaire entre $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$.

Exemple 9.2 (suite et fin) On trouve $R^2 \approx 0.85$ donc on peut dire l'ajustement linéaire explique environ 85% de l'information contenue dans les données.

Remarque 9.16 Le coefficient R^2 permet seulement de juger le degré de liaison linéaire entre X et Y. Il ne faut en aucun cas interpréter la validité du modèle comme représentant le degré de liaison causale ou explicative entre Y et X. Le R^2 peut être très élevé pour d'autres raisons qu'une relation d'explication. Une erreur courante est de croire qu'un coefficient de détermination élevé (et donc un coefficient de corrélation linéaire élevé) induit une relation de causalité (cause à effet) entre deux variables. En réalité, les deux variables peuvent être corrélées à une même troisième variable (le temps par exemple) dont dépendent les deux autres.

9.5 Exemples d'application de l'ajustement linéaire à l'ajustement non linéaire

9.5.1 Ajustement par une fonction puissance

Il s'agit des ajustements de la forme

$$y = bx^a$$

où y et x sont des variables positives et $a \in \mathbb{R}$ et $b \in \mathbb{R}_+$ sont à déterminer. En passant au logarithme, on obtient

$$\ln y = a \ln x + \ln b.$$

On pose $z = \ln y$, $u = \ln x$ et on obtient un ajustement linéaire de la forme $z = au + k$ où a et k sont calculés par la méthode MCO. On en déduit $b = e^k$.

9.5.2 Ajustement par une fonction exponentielle

Il s'agit des ajustements de la forme

$$y = be^{ax}$$

où y est une variable positive et $a \in \mathbb{R}$ et $b \in \mathbb{R}_+$ sont à déterminer. En passant au logarithme, on obtient

$$\ln y = ax + \ln b.$$

On pose $z = \ln y$ et on obtient un ajustement linéaire de la forme $z = ax + k$ où a et k sont calculés par la méthode MCO. On en déduit $b = e^k$.

9.5.3 Ajustement par une fonction logistique

Il s'agit des ajustements de la forme

$$y = \frac{1}{1 + e^{-(ax+b)}}$$

où y est une variable positive et $a \in \mathbb{R}$ et $b \in \mathbb{R}_+$ sont à déterminer. On a :

$$\begin{aligned}
y = \frac{1}{1 + e^{-(ax+b)}} &\iff \frac{1}{y} = 1 + e^{-(ax+b)} \\
&\iff \frac{1}{y} - 1 = e^{-(ax+b)} \\
&\iff \frac{1-y}{y} = e^{-(ax+b)} \\
&\iff \ln\left(\frac{1-y}{y}\right) = -(ax+b) \\
&\iff \ln\left(\frac{y}{1-y}\right) = ax+b.
\end{aligned}$$

En posant

$$z = \ln\left(\frac{y}{1-y}\right),$$

on obtient un ajustement linéaire de la forme $z = ax + b$ où a et k sont calculés par la méthode MCO.

9.6 Exercices

N.B. L'on pourra consulter l'annexe B pour une illustration graphique des différents ajustements.

Exercice 9.1 L'observation des prix et des quantités sur un marché de la tomate a donné les résultats suivants :

Quantités x en kg	10	20	35	50	70	90	110	130
Prix y au kg en kFCFA	5	3.75	2.75	2.25	1.75	1.25	0.8	0.5

où 1 kFCFA = 1000 FCFA. Ainsi, une quantité de 35 kg de tomates est vendue au prix de 2750 FCFA le kg.

1) Représenter le nuage de points.
2) (a) Déterminer la droite d'ajustement linéaire $y = ax + b$ par la méthode MCO.
   (b) Prévoir le prix d'un kg de tomates pour un achat de 140 kg et en déduire que l'ajustement linéaire n'est pas réaliste.
3) (a) Chercher maintenant un ajustement non linéaire de la forme $y = a \ln x + b$. On pourra poser $z = \ln x$ et se ramener à un ajustement linéaire $y = az + b$.
   (b) Calculer le coefficient R^2 de l'ajustement linéaire entre z et y.
   (c) Prévoir le prix au kg pour un achat de 140 kg.

Exercice 9.2 On a relevé la distance de freinage y d'un véhicule (distance parcourue par le véhicule entre le moment où le conducteur commence à freiner et l'arrêt du véhicule) à décélération constante sur une route plane sèche pour différentes vitesses x en km/h. Les données sont consignées dans le tableau suivant :

x (en km/h)	30	50	90	110	130
y (en m)	5.1	14.1	45.5	68	95

Les experts affirment que la distance d'arrêt d'un véhicule est proportionnelle à son énergie cinétique et donc au carré de sa vitesse. On cherche à confirmer cela à l'aide d'un ajustement par une fonction puissance de la forme $y = kx^{\alpha}$, où k et α sont des constantes à déterminer.

1) Montrer en faisant un changement de variable simple, qu'on peut se ramener à un problème de régression linéaire simple puis déterminer les coefficients k et α par la méthode des moindres carrés.
2) Les résultats obtenus confirment-ils l'affirmation des experts ?
3) Confirmer la qualité du modèle en calculant le coefficient de corrélation linéaire entre y et x^2.
4) Estimer la distance de freinage correspondant à une vitesse de 70 km/h.

Exercice 9.3 Dans une expérience de cinétique chimique, l'on a mesuré la concentration x d'une substance chimique en fonction du temps t et obtenu les valeurs numériques suivantes :

t	20	30	40	50	60	70	80
x	0.05	0.10	0.27	0.52	0.80	0.90	0.95

1) Faire le nuage de points.
2) Au vu de la forme du nuage, l'on se propose d'exprimer la grandeur
$$y = \ln\left(\frac{x}{1-x}\right)$$
comme une fonction linéaire du temps $y = A + Kt$ où la constante K est appelée constante de vitesse de la réaction étudiée. Déterminer les coefficients A et K par la méthode MCO.
3) Donner l'expression de x en fonction de t.

Exercice 9.4 La concentration molaire C d'une espèce chimique est mesurée au cours du temps. On obtient les données suivantes :

t (s)	20	40	60	80	100	120
C (mol.l^{-1})	278×10^{-3}	192×10^{-3}	147×10^{-3}	119×10^{-3}	100×10^{-3}	86×10^{-3}

1) Représenter les données par un nuage de points.
2) Considérons le modèle de régression linéaire $C = at + b$. En observant uniquement la variation de la concentration en fonction du temps, dire quel serait le signe de a et que vaudrait la limite de C quand $t \to +\infty$. En déduire que le modèle de régression linéaire n'est pas réaliste.
3) On propose le modèle $\dfrac{1}{C} = at + b$.
   (a) Calculer la limite de C quand $t \to +\infty$. Ce modèle vous semble-t-il raisonnable ?
   (b) En faisant le changement de variable $x = \frac{1}{C}$, déterminer les coefficients a et b.
   (c) Prévoir la valeur de C pour $t = 180$.

Troisième partie

Annexe

ANNEXE A

TABLE DE LA FONCTION DE RÉPARTITION ET DES QUANTILES DE LA LOI $\mathcal{N}(0,1)$

Table des valeurs de la fonction de répartition de la loi $\mathcal{N}(0,1)$.

$$\Phi(x) = \mathbb{P}(X \leqslant x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-t^2/2)\,\mathrm{d}t.$$

x	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5150	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.60	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.70	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.80	0.7881	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106	0.8133
0.90	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.00	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.10	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.20	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.60	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.00	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.20	0.9861	0.9865	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.30	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.40	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.50	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.60	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.70	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.80	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.90	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

Grandes valeurs de x

x	**3.0**	**3.1**	**3.2**	**3.3**	**3.4**	**3.5**	**3.6**	**3.7**	**3.8**
$\Phi(x)$	0.99865	0.99903	0.99931	0.99952	0.99966	0.99977	0.99984	0.99989	0.99993

Quantiles usuels de la loi $\mathcal{N}(0,1)$.

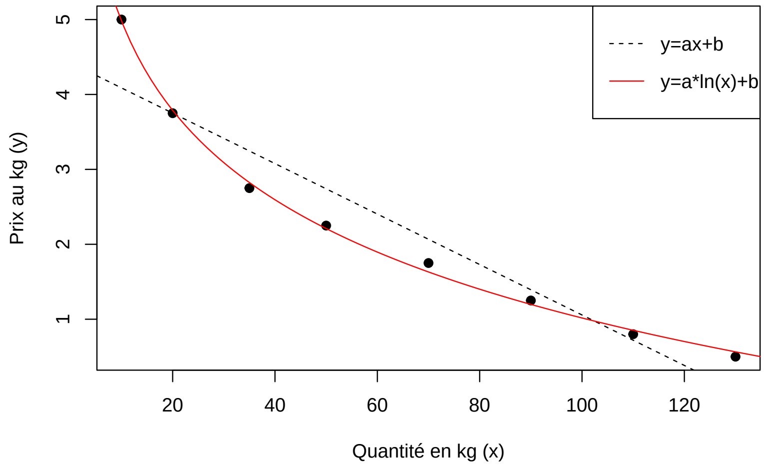
Pour tout $p \in]0,1[$, le quantile d'ordre p de la loi $\mathcal{N}(0,1)$ est $\Phi^{-1}(p) = z$ tel que $\mathbb{P}(X \leqslant z) = p$.

p	0.5	0.75	0.8	0.9	0.95	0.975	0.99	0.995
$\Phi^{-1}(p)$	0	0.67	0.84	1.28	1.64	1.96	2.33	2.58

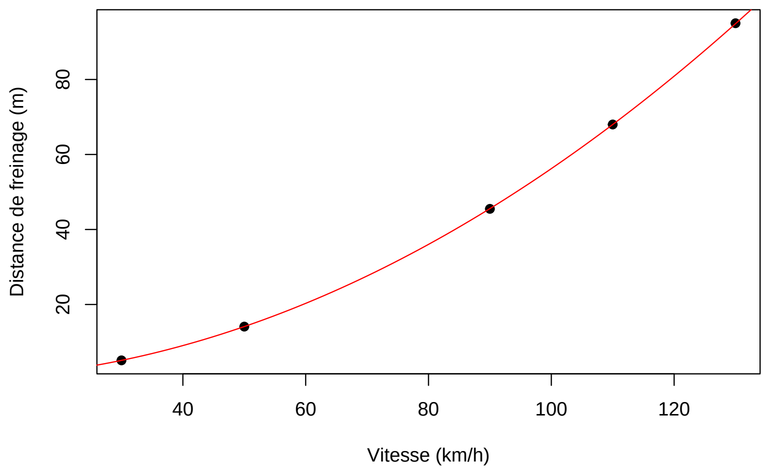
ANNEXE B

ANNEXE POUR LES EXERCICES DU CHAPITRE 9

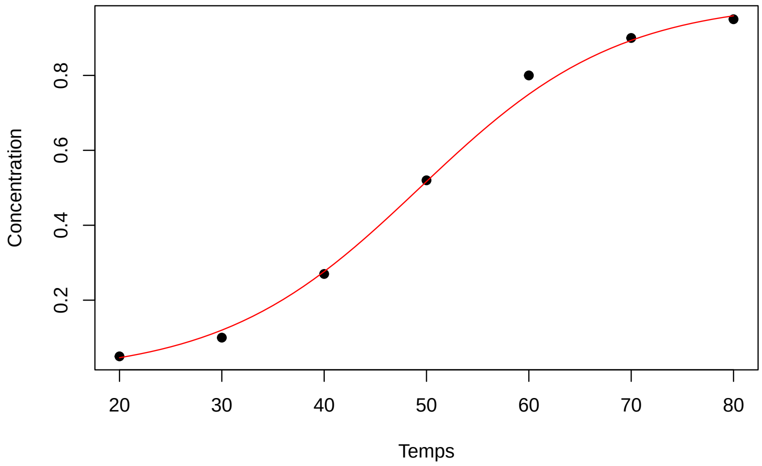
Exercice 9.1



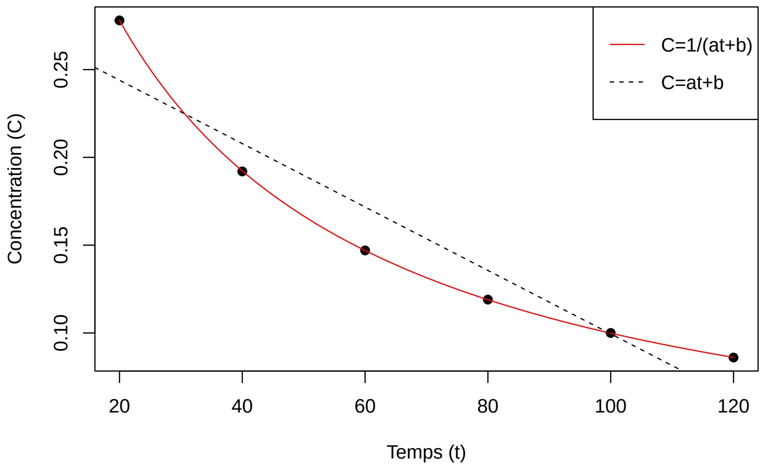
Exercice 9.2



Exercice 9.3



Exercice 9.4



Références bibliographiques

[1] S. Belhaj & A. Ben Aïssa – *Mathématiques pour l'informatique*, Vuibert, Paris, 2013.

[2] E. Cantoni, P. Huber & E. Ronchetti – *Maîtriser l'aléatoire : Exercices résolus de probabilités et statistique*, Springer, Paris, 2006.

[3] Y. Dodge – *Premiers pas en statistique*, Springer, Paris, 2006.

[4] S. Ferrigno, D. Marx, A. Muller-Gueudin, F. Bertrand & M. Maumy-Bertrand – *Mathématiques pour les sciences de l'ingénieur*, Dunod, 2013.

[5] D. Fredon, M. Maumy-Bertrand & F. Bertrand – *Mathématiques : Statistique et probabilités en 30 fiches*, Dunod, 2009.

[6] C. Gautier, A. Warusfel, B. Caminade & S. Nicolas – *Mathématiques «tout-en-un» ECS 2e année : Cours et exercices corrigés*, Dunod, 2008.

[7] K. E. Gneyou – « Probabilités et Statistique », 2017, Notes de cours, Université de Lomé (Togo).

[8] J.-P. Lecoutre – *Statistique et probabilités : Travaux dirigés*, 4ème éd., Dunod, 2008.

[9] — , *Statistique et probabilités (cours et exercices corrigés)*, 6ème éd., Dunod, 2016.

[10] G. Saporta – *Probabilités, analyse de données et Statistique*, 2e éd., Editions Technip, Paris, 2006.

[11] C. Suquet – « Introduction au calcul des probabilités », 2002, Polycopié de cours, Université de Lille (France).

[12] B. Tribout – *Statistique pour économistes et gestionnaires*, Pearson, Paris, 2007.

[13] R. Veysseyre – *Statistique et probabilités pour l'ingénieur*, 2e éd., Dunod, Paris, 2006.

[14] L. Wasserman – *All of Statistics*, Springer, 2003.