



Polycopié de cours destiné aux étudiants de
L1 Mathématiques et Informatique

**Introduction aux probabilités et à
la statistique descriptive**

ACHEMINE Farida



Préface

Ce polycopié est un ouvrage destiné aux étudiants de première année Mathématiques et informatique, mais aussi à tous les étudiants ayant besoin d'outils de probabilités et de statistique descriptive. Le lecteur trouvera une section consacrée au cours, suivie à la fin de chaque chapitre par une section d'exercices corrigés. La plupart de ces exercices ont été tirés de travaux dirigés ou ont été utilisés lors d'évaluations des connaissances.

Nous avons scindé ce document en trois chapitres.

- Nous introduisons dans le chapitre 1 les notions de base et le vocabulaire de la statistique descriptive.
- Le deuxième chapitre est dédié aux représentations numériques des données statistiques.
- Dans le troisième chapitre, nous avons présenté la partie calcul des probabilités et un complément sur l'analyse combinatoire.

À la fin de chaque chapitre, le lecteur découvrira une série d'exercices types accompagnés de leurs corrigés, présentés de manière progressive et détaillée.

Nous espérons que ce polycopié réponde aux attentes des étudiants et qu'il les aidera à réussir.

Mots clés : Statistique descriptive ; Représentation numérique des données ; Calcul des probabilités

Semestre : 02

Unité d'enseignement : Méthodologique.

Matière : Introduction aux probabilités et à la statistique descriptive

Crédits : 3

Coefficient : 2

Objectif de l'enseinemet :

Introduire les notions fondamentales de probabilités et en séries statistiques à une variables.

Connaissances préalables recommandées : Mathématiques de base.

Contenu de la matière :

Chapitre 1 : Notions de base et vocabulaire statistique.

- Concepts de base de la statistique (Population et individu, Variable ou caractère).
- Les tableaux statistiques : Cas de variables qualitatives (Représentation circulaire par des secteurs, Représentation en tuyaux d'orgue, Diagramme en bandes, cas de variables quantitatives (Diagramme en bâtons, Histogramme, Polygone).

Chapitre 2 : Représentation numérique des données.

- Les caractéristiques de tendance centrale ou de position (La Médiane, Les quartiles, Le mode, La moyenne arithmétique, La moyenne arithmétique pondérée, La moyenne géométrique, La moyenne harmonique, La moyenne quadratique).
- Les caractéristiques de dispersion (L'étendu, L'écart-type, L'écart absolu moyen, Le coefficient de variation).

Chapitre 3 : Calcul des probabilités.

- Analyse combinatoire : (Principe fondamental de l'analyse combinatoire, Arrangements, Permutations, Combinaisons).
- Espace probabilisable : (Expérience aléatoire, Evénements élémentaires et composés, Réalisation d'un événement, Evénement incompatible, Système complet d'événements, Espace probabilisable, Concept de probabilité).
- Espace probabilisé : (Définitions, conséquence de la définition, probabilité conditionnelle, événements indépendants, expériences indépendantes).
- Construction d'une probabilité.
- Probabilités conditionnelles, indépendance et probabilités composées (Probabilités conditionnelles, Indépendance, Indépendance mutuelle, Probabilités composés, Formule de Bayes).

Mode d'évaluation : Examen (60%), contrôle continu (40%)

Références .

G. Calot, Cours de statistique descriptive, Dunod, Paris, 1973. .

P. Bailly, Exercices corrigés de statistique descriptive, OPUs Alger, 1993. .

H. Hamdani, Statistique descriptive avec initiation aux méthodes d'analyse de l'information

économique : exercices et corrigés, OPU Alger, 2006. .

K. Redjal, Probabilités, OPU Alger, 2004

Table des matières

1	Éléments de la statistique descriptive	7
1.1	Introduction	7
1.2	Concepts de base de la statistique	7
1.2.1	Variable statistique	8
1.3	Tableaux statistiques et représentations graphiques	9
1.3.1	Cas d'un caractère qualitatif	10
1.3.2	Cas d'une variable quantitative discrète	11
1.3.3	Cas d'une Variable quantitative continue	13
1.3.4	Règle de STURGES	14
1.4	La courbe cumulative	16
1.4.1	Le cas d'une variable quantitative discrète	16
1.4.2	Le cas d'une variable quantitative continue	18
1.5	Série d'exercices N° 1	20
1.6	Solutions des exercices	21
2	Description numérique des données	26
2.1	Introduction	26
2.2	Les paramètres de tendance centrale et de position	26
2.2.1	Le mode	26
2.2.2	La médiane notée M_e	28
2.2.3	Les quartiles	30
2.2.4	La moyenne	32
2.3	Paramètres de dispersion (caractéristiques de dispersion)	38
2.3.1	L'étendu	39
2.3.2	L'écart absolu moyen	39
2.3.3	L'écart interquartile	40
2.3.4	La variance et L'écart-type	40
2.3.5	Coefficient de variation	42

2.4 Exercices (Série 2)	43
2.5 Solution des exercices	45
3 Calcul des probabilités	53
3.1 Introduction	53
3.2 Ensemble fondamental et événement	53
3.3 Espace de probabilités (Le modèle probabiliste de Kolmogorov)	56
3.3.1 Cas où Ω est fini	57
3.4 Probabilité conditionnelle et indépendance	58
3.4.1 Probabilité conditionnelle	58
3.4.2 Formule des probabilités totales	60
3.4.3 Formule de Bayes	61
3.4.4 Indépendance stochastique	62
3.5 Complément du cours : Eléments d'Analyse combinatoire	64
3.5.1 Les outils	64
3.5.2 Principes de base	65
3.5.3 Eléments discernables et éléments indiscernables	66
3.5.4 Les différentes dispositions	67
3.5.5 Les arrangements	67
3.5.6 Permutations	68
3.5.7 Combinaisons	69
3.6 Exercices	71
3.6.1 Exercices sur l'analyse combinatoire (Série 3)	71
3.6.2 Corrigés des exercices	73
3.6.3 Exercices de Probabilités (Série 4)	75
3.6.4 Corrigés des exercices de probabilités	77
Bibliographie	82

Chapitre 1

Eléments de la statistique descriptive

1.1 Introduction

La statistique est un ensemble de méthodes et techniques scientifiques qui sont utilisées à recueillir, de classer et de résumer l'information sur un sujet donné. Ensuite, il s'agit d'analyser et d'expliquer pour prendre une décision éclairée. Elle est utilisée dans l'économie, la biologie, la psychologie, les sciences de l'ingénieur ... etc.

Pour plus de détails, le lecteur est invité à consulter [\[3, 4, 8, 9, 2, 6, 11\]](#)

1.2 Concepts de base de la statistique

Définition 1.1. Population

On appelle population l'ensemble des éléments (objets, personnes, animaux, virus,) sur lesquels porte l'étude statistique. Cet ensemble sera noté Ω .

Exemple 1.1. *On considère l'ensemble des étudiants de la filière mathématiques et informatique (MI). On s'intéresse aux moyennes générales, obtenues par les étudiants, dans le premier semestre. Dans ce cas $\Omega =$ ensemble des étudiants de première année MI.*

Définition 1.2. Individu (ou unité statistique)

Un individu est un élément d'une population et on l'appelle aussi "unité statistique". Il est noté ω , avec $\omega \in \Omega$.

Exemple 1.2. *Si on étudie la production annuelle d'un laboratoire de boîtes de plaquenil (médicament utilisé dans le traitement de la covid-19). Alors, la population est l'ensemble des boîtes produites durant l'année et une boîte constitue un individu.*

Définition 1.3. Echantillon

En générale, il est difficile de faire l'étude de toute la population. Donc, on réduit l'étude à

examiner une partie de cette population. Cette partie est nommée échantillon.

Le nombre d'individus que contient un échantillon est dit taille de l'échantillon.

Remarque 1.1. .

1. **Une enquête statistique** est l'action qui consiste à étudier l'ensemble des individus d'un échantillon ou de la population.
2. **Un Recensement** représente une enquête statistique où l'échantillon étudié est égal à la population.
3. Dans le cas où l'étude est portée sur un échantillon de la population, on dit qu'on fait un **sondage**.
4. **Les données statistiques** englobe les individus sur lesquels porte l'étude statistique, l'ensemble des variables considérées et les observations de ces variables sur les individus.

1.2.1 Variable statistique

Définition 1.4. On peut décrire un individu selon une ou plusieurs caractéristiques qu'on appelle **variables statistiques ou caractères** . Ces variables sont notées par les lettre majuscules X, Y, \dots

X est une application définie sur Ω dans un autre ensemble noté Θ :

$$X : \Omega \longrightarrow \Theta$$

$$\omega \longmapsto X(\omega)$$

Exemple 1.3. (Quelques exemples de Caractères)

La taille, la température, la catégorie socioprofessionnelle, l'âge, la nationalité, la couleur des yeux, le prix, la surface, la distance, le salaire,...

Définition 1.5. Les **modalités** d'une variable statistique sont les différentes valeurs qu'elle peut prendre et on les note par $x_i, y_i \dots$

Exemple 1.4. .

- Si le caractère considéré est "le nombre d'enfants par famille", les modalités peuvent être "0,1,2,3,..."

- Si la variable est "situation familiale", les modalités sont "célibataire, marié, divorcé".

On classe les variables statistiques en deux catégories : la variable statistique qualitative et la variable statistique quantitative.

- Variable statistique qualitative (Caractère qualitatif)

Définition 1.6. Lorsque les modalités d'un caractère ne sont pas mesurables, on qualifie ce caractère de caractère qualitatif. Les modalités d'un caractère qualitatif ne peuvent pas être ordonnées. Par exemple, il est impossible d'ordonner la modalité "marié" et la modalité "divorcé".

Exemple 1.5. Si on considère la variable quantitative : l'état d'une voiture. Les modalités possibles : Ancienne, Dégradée, Neuve.

• Variable statistique quantitative

Définition 1.7. Lorsque les modalités du caractère sont des nombres, on dit que le caractère est quantitatif.

Il y a deux sortes de caractère quantitatif : le caractère discret et le caractère continu.

Définition 1.8. Un caractère quantitatif est dit

- discret si le caractère prend un nombre fini ou dénombrable de valeurs. Par exemple : le nombre d'appartement par quartier dans une ville.
- continu si le caractère prend toutes ses valeurs dans des intervalles de l'ensemble des nombres réels, il est dit continu. Par exemple : "la taille", "le poids",...

Exemple 1.6. Un exemple intéressant pour différencier les variables discrètes des variables continues est de considérer la taille des pieds qui est une variable continue et les pointures des chaussures qui est une variable discrète.

Remarque 1.2. Une variable quantitative discrète qui prend un grand nombre de valeurs distinctes est étudiée comme une variable continue.

1.3 Tableaux statistiques et représentations graphiques

Le but de la statistique descriptive est de résumer l'information contenue dans les observations en utilisant des tableaux, graphiques ou paramètres numériques.

La manière la plus basique de représenter des données statistiques est l'énumération des observations. On appelle cette énumération une série statistique. On note que dans certains cas, des valeurs peuvent être répétées.

Dans les cas où les observations sont nombreuses, il est avantageux de les porter dans un tableau des effectifs ou des fréquences qu'on appelle tableau statistique. La construction du tableau statistique suit le même principe pour les différents caractères, impliquant le comptage des observations. Cependant, la méthode de construction du tableau statistique varie en fonction de la nature de la variable statistique (quantitative continue, quantitative discrète,

qualitative).

Les représentations par des tableaux offre un moyen efficace et simple pour décrire des données statistiques. Elles sont utilisées pour plusieurs raisons, notamment

- simplifier la lecture des informations d'une série statistique.
- comparer deux séries statistiques.

La construction des tableaux dépend de la nature de la variable statistique.

1.3.1 Cas d'un caractère qualitatif

1. On note par n la taille de l'échantillon étudié. Elle est appelée aussi l'effectif total.
2. On note par m_i les différentes modalités du caractère.
3. n_i est l'effectif partiel (ou absolu), c'est le nombre d'individus ayant la modalité m_i . On a $\sum_{i=1}^{i=k} n_i = n$, avec k est le nombre des modalités de la variable statistique.
4. $f_i = \frac{n_i}{n}$ est la fréquence relative ou fréquence partielle. C'est la proportion d'individus qui possède la modalité m_i . On a alors $\sum_{i=1}^{i=n} f_i = 1$.

Le tableau statistique d'un caractère qualitatif est donné par.

Modalité m_i du caractère C	m_1	m_2	...	m_k	Total
Nombre d'individus (effectif) n_i	n_1	n_2	...	n_k	n

On utilise deux types de représentation graphique.

(1) Les diagrammes circulaires consistent à diviser un disque en secteurs. Chaque secteur représente une modalité observée. Et dont la surface est proportionnelle à l'effectif de la modalité.

La modalité m_i occupe un secteur de degré $\theta_i = f_i \times 360^\circ = \frac{n_i}{n} 360^\circ$.

(2) Représentation cartésienne (diagramme en tuyaux d'orgue). Les modalités sont placées le long de l'axe horizontal, tandis que les fréquences ou les effectifs sont données sur l'axe vertical.

Voir l'exemple ci-dessous.

Exemple 1.7. On considère la représentation de 200 jeunes selon leurs sport préférés.

sport préféré m_i	football	Natation	Athlétisme	Tennis	Total
n_i	60	50	45	45	200
f_i	0,3	0.25	0.225	0.225	1
$\theta_i = 360 \times f_i$	108	90	81	81	360

Le diagramme circulaire est donné par la FIGURE 1.1.

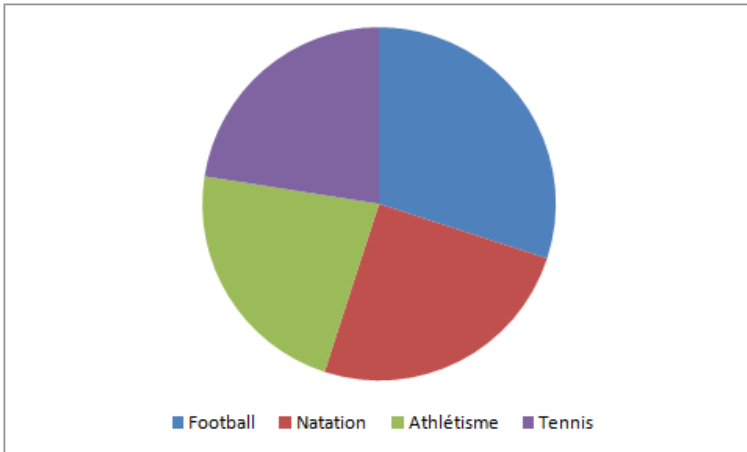


FIGURE 1.1 – Représentation de 200 jeunes selon leurs sports préférés.

Le diagramme en tuyaux d'orgue est donné par la FIGURE 1.2.

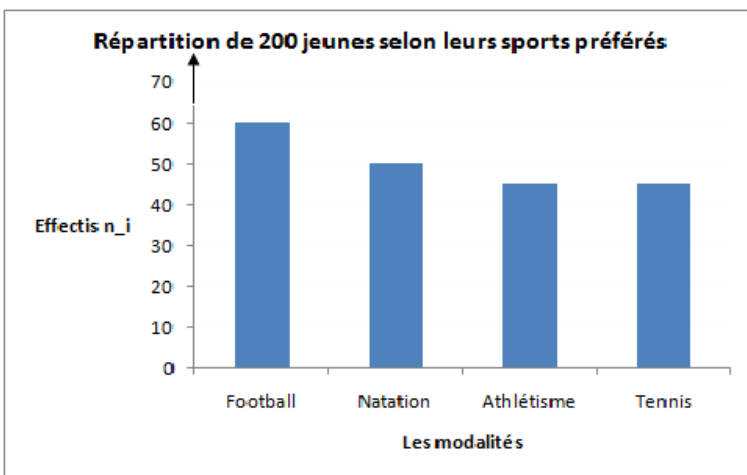


FIGURE 1.2 – Représentation de 200 jeunes selon leurs sports préférés.

1.3.2 Cas d'une variable quantitative discrète

Noutons X la variable statistique étudiée, avec $X(\Omega) = \{x_1, x_2, \dots, x_k\}$.

x_1, x_2, \dots, x_k sont les différentes valeurs de la variable quantitative discrète X dans un échantillon de taille n .

1. On note par n_i le nombre d'individus associés à la modalité x_i . On l'appelle effectif partiel

de x_i , $i = 1, 2, \dots, k$. $\sum_{i=1}^{i=k} n_i = n$, où n est la taille de l'échantillon.

2. $f_i = \frac{n_i}{n}$ est la fréquence relative de la modalité x_i , $i = 1, 2, \dots, k$. On a $\sum_{i=1}^{i=k} f_i = 1$

La forme général d'un tableau statistique pour une variables quantitative discrète est la suivante

x_i	x_1	x_2	...	x_k	Total
Effectifs n_i	n_1	n_2	...	n_k	n

Pour la représentation d'une variable quantitative discrète, on utilise un graphique appelé le diagramme en bâtons. Les valeurs de la variable statistique (V.S) x_i , $i = 1, \dots, k$ sont représentées sur l'axe des abscisses et les fréquences (ou les effectifs) sont représentées sur l'axe des ordonnées. Dans le diagramme en batons, on dessine pour chaque modalité de la variable un trait de hauteur égale à la fréquence (ou effectif) de cette modalité.

On complète ce diagramme par un polygone des fréquences ou des effectifs : ce qui revient à relier par des segments de droites les points (x_i, f_i) .

Exemple 1.8. Répartition des familles selon le nombre d'enfants.

x_i	0	1	2	3	4	5	6	Total
n_i	6	15	12	11	8	3	1	56

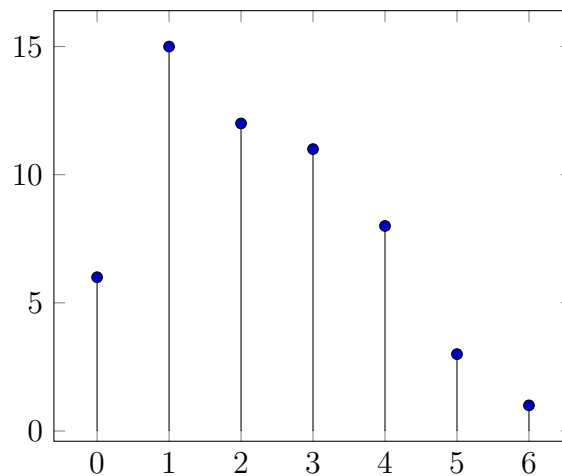


FIGURE 1.3 – Représentation de 56 familles selon le nombre d'enfants.

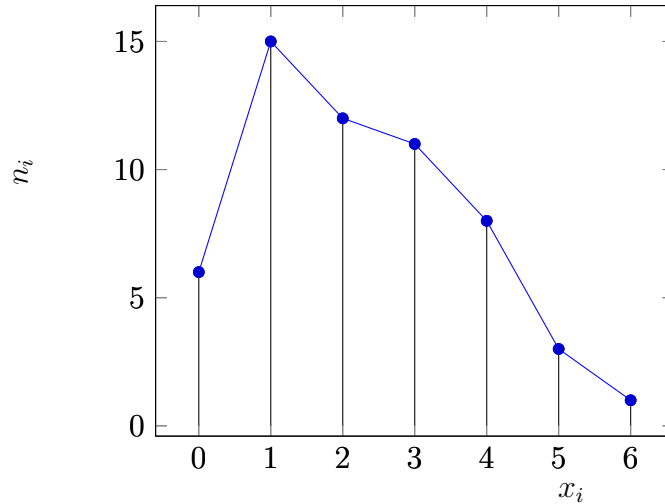


FIGURE 1.4

Le polygone des effectifs est donné dans la FIGURE 1.4.

1.3.3 Cas d'une Variable quantitative continue

Soit X la variable statistique étudiée, avec $(\omega) = [a, b]$.

Dans le cas où la variable statistique (V.S.) est continue, on regroupe les modalités en k classes disjointes qu'on notera $C_i = [e_i, e_{i+1}[$, $i = 1, \dots, k-1$, avec $[a, b] = [e_1, e_2[\cup [e_2, e_3[\cup \dots \cup [e_{k-1}, e_k]$.

1. $\bigcup_{i=1}^{i=k-1} [e_i - e_{i+1}[$ contient toute les modalités possibles de la série statistique.
2. On note par e_i la borne inférieure de la classe $[e_i, e_{i+1}[$.
3. On note par e_{i+1} la borne supérieure de la classe $[e_i, e_{i+1}[$.
4. Par convention la borne supérieure d'une classe n'est pas incluse dans cette classe sauf pour la dernière qui peut être incluse (intervalle fermé).
5. $a_i = e_{i+1} - e_i$ est l'amplitude de la classe $C_i = [e_i, e_{i+1}[$.
6. On note $x_i = \frac{e_{i+1} + e_i}{2}$ le centre de la classe C_i .
7. on note n_i l'effectif de la classe C_i . C'est le nombre d'individus dont les valeurs du caractère est dans la classe $[e_i, e_{i+1}[$.
8. Dans un échantillon de taille n , $f_i = \frac{n_i}{n}$ est la fréquence relative de la classe $[e_i, e_{i+1}[$.

La tableau statistique est comme suit

Classes	$[e_1, e_2[$	$[e_2, e_3[$...	$[e_k, e_{k+1}[$	Total
Effectifs n_i	n_1	n_2	...	n_k	n

La représentation graphique d'une V.S. continue est appelée histogramme.

(1) Dans le cas où les amplitudes des classes sont égales à a , les différentes classes sont représentées sur l'axe des abscisses. Un rectangle de largeur égale à a est dessiné au-dessus de chaque classe C_i , dont la hauteur est proportionnelle à l'effectif n_i (ou à la fréquence f_i).

(2) Dans le cas où les amplitudes des classes sont différentes. On commence par corriger les effectifs comme suit. L'effectif de C_i : $h_i = n_i \times \frac{\text{la plus petite amplitude de la classe } C_i}{\text{l'amplitude}}$. Ensuite, l'histogramme est tracé comme dans (1).

(3) Le polygone statistique des fréquences (ou des effectifs) : est tracé en reliant les milieux des sommets des rectangles de l'histogramme des fréquences (ou des effectifs) en utilisant des segments de droite.

Exemple 1.9. Soient les taux de glycémie observés sur un échantillon de 100 patients représentés dans le tableau statistique :

Classes	$[0.8, 0.85[$	$[0.85, 0.90[$	$[0.90, 0.95$	$[0.95, 1[$	Total
n_i	15	20	35	30	100

La représentation graphique de cette série est donnée dans la FIGURE 1.5.

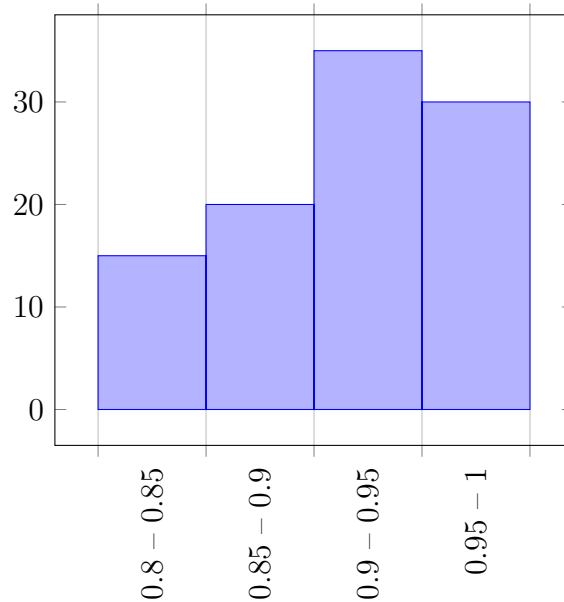


FIGURE 1.5 – Représentation de 100 patients selon leurs taux de glycémie

1.3.4 Règle de STURGES

Soient x_1, x_2, \dots, x_n les valeurs observées d'une variable statistique continue.

Question : En combien de classes partager les valeurs de la variable X ?

Réponse : L'une des réponses à cette question est donnée par la règle de STURGES (Herbert STURGES (1882-1958)). Il a donné une valeur approximative pour le nombre de classe k en fonction de la taille n de l'échantillon $1 + 3.3\log_{10}(n)$.

Le nombre de classes est égal à l'entier le plus proche de $1 + 3.3\log_{10}(n)$. En utilisant cette formule, on obtient la table ci-dessous.

La table de STURGES est donnée comme suit.

Nombre de données (n)	Nombre de classe
12 à 23	5
24 à 46	6
47 à 93	7
94 à 187	8
188 à 376	9
377 à 756	10
757 à 1519	11
1520 à 3053	12
3054 à 6135	13
6136 à 12328	14
12329 et plus	15

Par exemple, pour un échantillon de taille $n = 100$, la règle de STURGES donne le nombre 7.6, on doit donc prendre le nombre de classes égale à 8.

Pour séparer une série de données brutes en classes, on procède de la manière suivante.

→ Déterminer le nombre de classes visées (REGLE DE STURGES) ;

→ Déterminer l'amplitude approximative des classes

$$\frac{x_{max} - x_{min}}{\text{nombre de classes}};$$

où x_{min} = la plus petite modalité de la variable X et x_{max} = la plus grande modalité de X .

Arrondir à une amplitude définitive a facile à lire.

(Par exemple si on trouve $\frac{x_{max}-x_{min}}{\text{nombre de classes}} = 2.101$, alors pour faciliter la lecture on prend $a = 2$).

→ Déterminer la borne inférieure de la première classe notée e_1 . Elle est fixée de manière à contenir la valeur minimum prise par la variable : $e_1 \leq x_{min}$; (inférieur ou égale à la plus petite valeur de la série statistique).

→ Créer les classes :

- construire la première classe : $[e_1, e_1 + a[$;

- pour trouver les classes suivantes, il suffit d'ajouter à chaque fois l'amplitude a ;
- Trouver l'effectif de chaque classe en faisant la somme des effectifs des éléments contenus dans chaque classe.

1.4 La courbe cumulative

Définition 1.9. Soit X une variable statistique quantitative discrète. Soient x_1, x_2, \dots, x_n les différentes valeurs de X . On appelle

- **effectif cumulé croissant de x_i** , la somme des effectifs partiels de x_i et les effectifs de toutes les valeurs de la variable qui sont inférieures à x_i . Il est défini par la formule

$$N_i = n_1 + n_2 + \dots + n_i;$$

- **fréquence cumulé de x_i** , est définie par la somme

$$F_i = f_1 + f_2 + \dots + f_i.$$

Remarque 1.3. .

- Pour une série statistique quantitative continue. En d'autres termes, les valeurs sont regroupées en classes. L'effectif cumulé d'une classe (la fréquence cumulée) est donné en additionnant les effectifs partiels (fréquences cumulée) de cette classe et de toutes les classes qui précèdent cette classe.
- On définit aussi l'effectif cumulé de la première valeur (de la première classe) par l'effectif partiel de cette valeur (classe). Et l'effectif cumulé de la dernière valeur (classe) est donné par la taille de l'échantillon n .
- Dans une série statistique, on définit la fréquence cumulée de la première valeur (la première classe) par la fréquence de cette valeur (classe). On définit aussi la fréquence cumulée de la dernière valeur (la dernière classe) par le nombre égal à 1.

1.4.1 Le cas d'une variable quantitative discrète

Soit X une variable quantitative discrète. On note x_1, x_2, \dots, x_k ses modalités, ordonnées par ordre croissant.

Définition 1.10. On appelle fonction de répartition de la variable X , la fonction F définie de

1.4. LA COURBE CUMULATIVE

\mathbb{R} à valeur dans $[0, 1]$ comme suit :

$$F(x) = \begin{cases} 0 & \text{si } x < x_1 \\ F_i & \text{si } x_i \leq x < x_{i+1} \\ 1 & \text{si } x_k \leq x \end{cases}$$

On rappelle que la fonction de répartition d'une variable quantitative discrète est constante par morceaux.

La représentation graphique de la fonction de répartition est dite courbe cumulative.

Exemple 1.10. Dans un village de 100 familles. On s'intéresse à la variable X : Le nombre de personnes par famille.

On obtient le tableau statistique suivant

x_i	2	3	4	5	6	7	8	Total
n_i	10	15	12	18	15	14	16	100
$f_i = \frac{n_i}{100}$	0.1	0.15	0.12	0.18	0.15	0.14	0.16	1
F_i	0.1	0.25	0.37	0.55	0.7	0.84	1	/

La fonction de répartition est donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 2 \\ 0.1 & \text{si } 2 \leq x < 3 \\ 0.25 & \text{si } 3 \leq x < 4 \\ 0.37 & \text{si } 4 \leq x < 5 \\ 0.55 & \text{si } 5 \leq x < 6 \\ 0.7 & \text{si } 6 \leq x < 7 \\ 0.84 & \text{si } 7 \leq x < 8 \\ 1 & \text{si } 8 \leq x \end{cases}$$

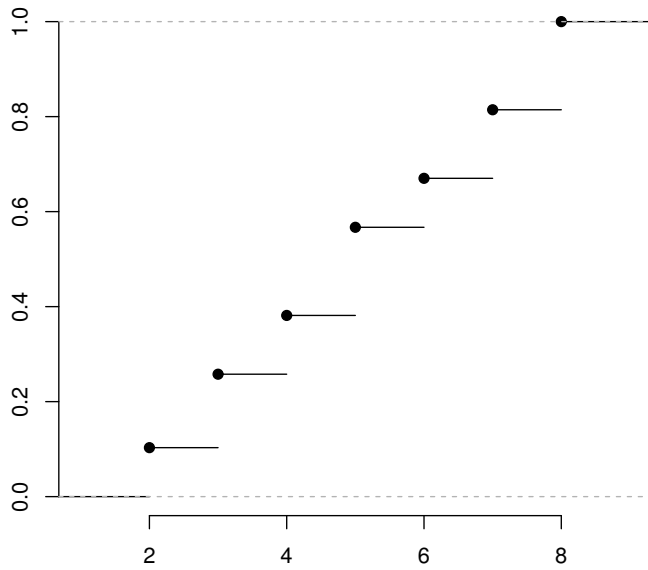


FIGURE 1.6 – La courbe cumulative de la variable X

1.4.2 Le cas d'une variable quantitative continue

Soit X une variable statistique quantitative continue dont la répartition des observations en classes est donnée par : $[e_1, e_2[, [e_2, e_3[, \dots, [e_k, e_{k+1}[$.

Définition 1.11. On appelle fonction de répartition de la variable X la fonction $F(\cdot)$ définie de \mathbb{R} dans $[0, 1]$, par

$$F(x) = \begin{cases} 0 & \text{si } x < e_1 \\ F_{i-1} + \frac{x-e_i}{e_{i+1}-e_i} \times f_i & \text{si } e_i \leq x < e_{i+1} \\ 1 & \text{si } e_k \leq x \end{cases}$$

Exemple 1.11. On a relevé la taille en cm de 20 élèves.

Classes	$[104, 108[$	$[108, 112[$	$[112, 116[$	$[116, 120[$	Total
n_i	4	5	8	3	20
$f_i = \frac{n_i}{20}$	0.2	0.25	0.4	0.15	1
F_i	0.2	0.45	0.85	1	/

La fonction de répartition est donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 104 \\ \frac{x-104}{4} \times 0.2 & \text{si } 104 \leq x < 108 \\ 0.2 + \frac{x-108}{4} \times 0.25 & \text{si } 108 \leq x < 112 \\ 0.45 + \frac{x-112}{4} \times 0.4 & \text{si } 112 \leq x < 116 \\ 0.85 + \frac{x-116}{4} \times 0.15 & \text{si } 116 \leq x < 120 \\ 1 & \text{si } 120 \leq x \end{cases}$$

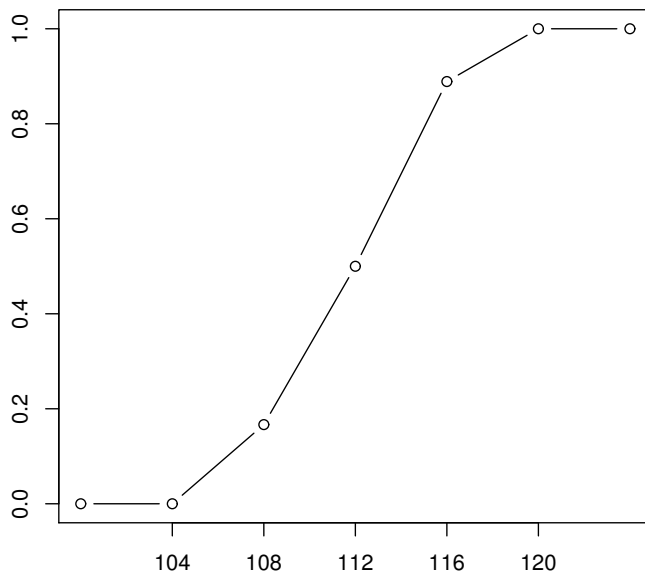


FIGURE 1.7 – La courbe cumulative de la variable X

1.5 Série d'exercices N° 1

Exercice 1.1. *Quelle est la nature de la variable statistique dans les cas suivants :*

- *Couleur des yeux.*
- *Le nombre de maison dans un village.*
- *Distance*
- *Nationalité*
- *Nombre d'enfants par famille*
- *Lieu de résidence.*
- *Marque de voiture.*
- *Le salaire d'employés d'une banque.*
- *La durée de vie des ampoules.*
- *La rigidité des ressorts.*
- *La masse d'une prothèse dentaire.*

Exercice 1.2. *120 clients d'une cafétéria ont répondu à un questionnaire sur leur satisfaction du service. Les résultats de cette enquête sont donnés comme suit*

4 clients ont répondu par "très peu". 5 des clients ont répondu par "un peu". 24 clients ont répondu par "moyennement". 36 clients ont répondu par "bien". 25 on répondu par "très bien". 26 clients ont répondu par "excellent".

- 1. Déterminez la population et la variable étudiées.*
- 2. Précisez la nature de la variable et ses modalités.*
- 3. Dresser le tableau statistique de cette série (effectifs).*
- 4. Donner les deux types de représentation graphique de cette variable.*

Exercice 1.3. *La liste suivante est donne les prénoms d'un groupe d'étudiants, suivis entre parenthèses du nombre de livres que chacun d'entre eux a lus au cours de l'année dernière :*

Ali (2), Kamel (3), Lamia (1), Lamys (3), Ines (2), Karim (2) Sami (0), Sonia (1), Farida (3), Laura (3), rayan (0), Rachel (3), Marie (1), Massy (3), Jugurtha (4), Masten (4), Massinissa (2), Vanessa (4), Sofia (2), Djura (1), Samia (3), Ilyas (2); Amelia (4), Emilie (3), Lina (2).

- 1. Déterminez la population et la variable étudiées.*
- 2. Précisez la nature et les modalités de la variable.*
- 3. Dresser le tableau statistique de cette série (effectifs, fréquences)*
- 4. Représenter la distribution des effectifs par un diagramme en bâtons.*
- 5. Donner le polygone des effectifs.*

Exercice 1.4. *On considère un échantillon d'étudiantes. On a mesuré le poids de chacune. Les poids ont été arrondies à l'entier. Nous obtenons les résultats suivants :*

53 -53 -53 -53 -58 -58 -58 -58 -63 -63 -63 -63 -64 -64 -65 -65 -65 -65- 67- 67- 67 -67- 70- 70- 70- 72- 72 -75- 75 -75- 75 -77- 77- 77- 85 -85 -85 -87.

1. *Quelle est la population étudiée ? Quel est le caractère étudié ? Quelle est sa nature ?*
2. *En utilisant la règle de STURGES, répartir les données en classes de même amplitude.*
3. *Donner le tableau statistique des effectifs et des fréquences.*
4. *Donner l'histogramme des effectifs.*

1.6 Solutions des exercices

Solution de l'exercice 1.1. .

- *"Couleur des yeux" est une variable qualitative.*
- *"Le nombre de maison dans un village" est une variable quantitative discrète.*
- *"Distance" est une variable quantitative continue.*
- *"Nationalité" est une variable qualitative.*
- *"Nombre d'enfants par famille" est une variable quantitative discrète.*
- *"Lieu de résidence" est une variable qualitative.*
- *"Marque de voiture" est une variable qualitative.*
- *"Le salaire d'employés d'une banque" est une variable quantitative continue.*
- *"La durée de vie des ampoules" est une variable quantitative continue.*
- *"La rigidité des ressorts" est une variable quantitative continue.*
- *"La masse d'une prothèse dentaire" est une variable quantitative continue.*

Solution de l'exercice 1.2. .

1. *La population étudiée Ω est l'ensemble des clients de la cafétéria.*

La variable étudiée X est "La satisfaction des clients de la cafétéria".

2. *La variable est qualitative.*

Les modalités de la variable sont : "très peu", "un peu", "moyennement", "bien", "très bien", "excellent".

3. *Le tableau statistique associé à la distribution des effectifs.*

<i>Satisfaction</i>	<i>très peu</i>	<i>un peu</i>	<i>moyennement</i>	<i>bien</i>	<i>très bien</i>	<i>excellent</i>	<i>Total</i>
<i>Effectifs n_i</i>	<i>4</i>	<i>5</i>	<i>24</i>	<i>36</i>	<i>25</i>	<i>26</i>	<i>120</i>

4. *Les deux types de représentation graphique de cette variable :*

(1) *Le diagramme circulaire est donné par la FIGURE 1.8.*

1.6. SOLUTIONS DES EXERCICES

On calcule l'angle θ_i correspondant à l'effectif n_i , $i = 1, 2, \dots, 6$.

$$\theta_i = f_i \times 360 = \frac{n_i}{n} \times 360.$$

Satisfaction	très peu	un peu	moyennement	bien	très bien	excellent	Total
Effectifs n_i	4	5	24	36	25	26	$n=120$
$\theta_i = \frac{n_i}{n} \times 360$	12	15	72	108	75	78	360

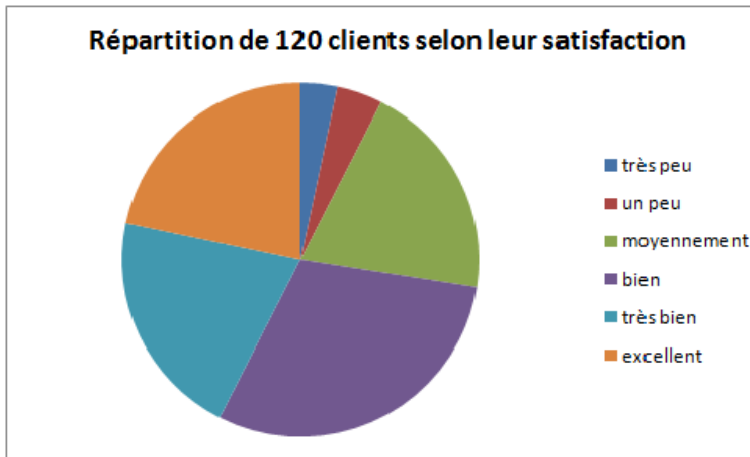


FIGURE 1.8 – Le diagramme circulaire.

(2) Le diagramme en tuyau d'orgue est donné par la FIGURE 1.9.

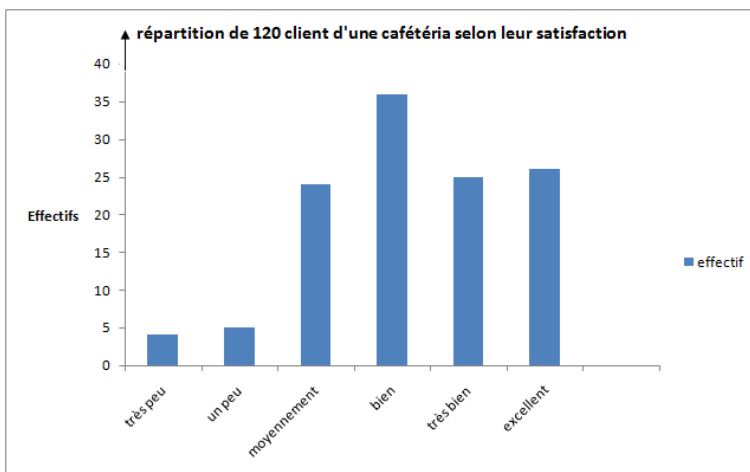


FIGURE 1.9 – Le diagramme en tuyaux d'orgue

Solution de l'exercice 1.3. .

1. La population étudiée Ω est le groupe d'étudiants.

La variable étudiée est $X =$ nombre de livres que chacun d'entre eux a lus au cours de l'année dernière.

2. La variable étudiée est quantitative discrète.

L'ensemble des modalités est $X(\Omega) = \{0, 1, 2, 3, 4\}$.

3. Le tableau statistique associé contient trois lignes :

-la première ligne contient les modalités x_i de X ;

-la deuxième ligne contient les effectifs n_i associés à chacune de ces modalités ; -la troisième ligne contient les fréquences $f_i = \frac{n_i}{n}$.

Le tableau statistique associé à X est le suivant :

x_i	0	1	2	3	4	Total
n_i	2	4	7	8	4	25
f_i	0.08	0.16	0.28	0.32	0.16	1

Le diagramme en bâtons est donné par la FIGURE 1.10.

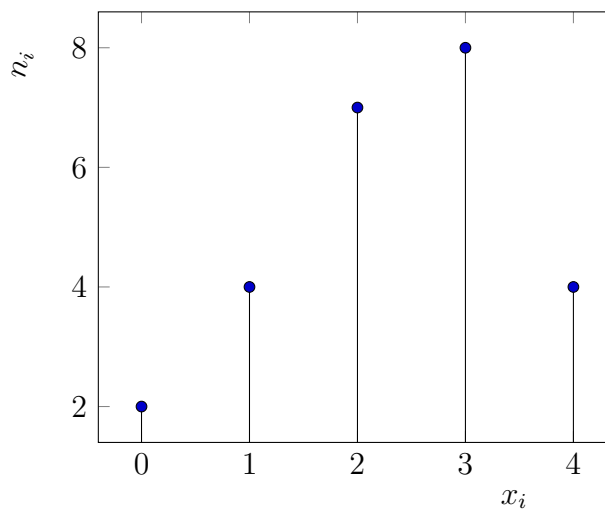


FIGURE 1.10

5) Le polygone des effectifs est donné par FIGURE 1.11 (lignes brisées de couleur bleue) :

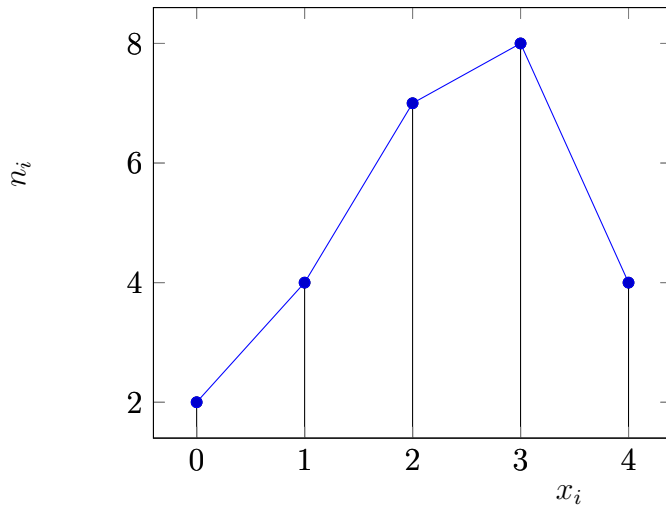


FIGURE 1.11

Solution de l'exercice 1.4. .

1. La population étudiée Ω : l'ensemble des étudiantes .

Le caractère étudié X : le poids.

La nature du caractère : il est quantitatif continu.

2. On calcule $1 + 3.3\log_{10}(40) = 6.28$.

En appliquant la règle de STURGES, on partage les valeurs de la série en 6 classes.

L'amplitude des classes est donnée par $a = \frac{x_{max} - x_{min}}{\text{nombre de classes}} = \frac{87 - 53}{6} = 5.66$ et on arrondi à 6.

La répartition des données en classes : $[53 - 59[$, $[59 - 65[$, $[65 - 71[$, $[71 - 77[$, $[[77 - 83[$ et $[83 - 89[$.

3. On obtient le tableau statistique suivant

Classes	$[53-59[$	$[59-65[$	$[65-71[$	$[71-77[$	$[[77-83[$	$[83-89[$	Total
Effectifs n_i	9	6	12	6	3	4	40
Fréquences $f_i = \frac{n_i}{40}$	0.225	0.15	0.3	0.15	0.075	0,1	1

4. La représentation graphique : L'histogramme des effectifs : FIGURE 1.12.

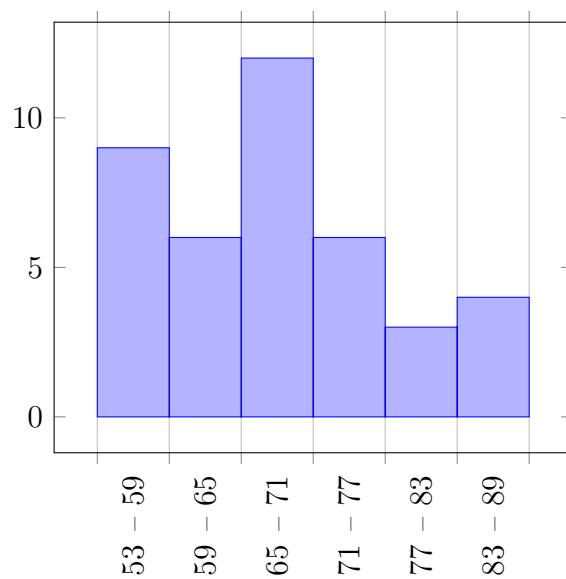


FIGURE 1.12

Chapitre 2

Description numérique des données

2.1 Introduction

Dans le chapitre 1, on s'est focalisé sur la manière de résumer les données d'une série statistique dans des tableaux de fréquences (ou des effectifs) et selon la nature des variables. Aussi, nous avons donné les différentes représentations graphiques. Pour plus de détail, voir les références [4, 8, 9, 2, 6, 16, 14].

Dans ce chapitre, on présentera deux types de valeurs numériques pour caractériser les variables statistiques :

- Les paramètres de tendance centrale ou de position qui sont utilisés à mesurer l'ordre de grandeur des observations.
- Les paramètres de dispersion qui servent à mesurer la dispersion (l'étalement) des observations autour d'un paramètre de position (de tendance centrale).

2.2 Les paramètres de tendance centrale et de position

Les mesures de tendance centrale sont utilisés pour résumer un ensemble de données concernant une variable quantitative. Elles identifient une valeur autour de laquelle des données ont tendance à se regrouper.

Les trois paramètres de tendance centrale les plus utilisés sont : la moyenne, la médiane et le mode. Ils s'expriment dans la même unité que celle de la variable statistique.

2.2.1 Le mode

Définition 2.1. *Le mode*

Le mode est la valeur de la variable à laquelle est associé le plus grand effectif ou la plus grande

fréquence. C'est la valeur la plus observée. Il est dénoté par M_O .

Le mode est un outil d'analyse statistique utilisé pour identifier la valeur la plus fréquente d'une variable statistique. Il est possible de trouver des distributions ayant deux ou plusieurs modes. Elles sont dites distribution bimodales ou plurimodales.

Calcul du Mode M_o : On distingue deux cas.

→ Dans le cas d'une variable quantitative discrète, le mode est déterminé directement à partir du tableau des fréquences (ou des effectifs).

Exemple 2.1. On considère la V.S. donnée par le tableau statistique :

x_i	1	2	3	4	5	6	7	Total
n_i	1	4	3	7	1	5	2	23

Il est clair qu'à partir du tableau, $M_o = 4$.

→ Dans le cas d'une variable quantitative continue, on commence par définir la classe modale.

Définition 2.2. La classe modale est la classe ayant l'effectif le plus élevé. On utilise l'effectif ramené à l'unicité d'amplitude.

Si $[e_i, e_{i+1}[$ est la classe modale, alors

$$M_o = e_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times a_i,$$

avec

e_i est la borne inférieure de la classe modale ;

a_i est l'amplitude de la classe modale ;

$\Delta_1 = n_i - n_{i-1}$ la différence d'effectif entre la classe modale et la classe précédente ;

$\Delta_2 = n_i - n_{i+1}$ la différence d'effectif entre la classe modale et la classe suivante ;

n_i est l'effectif de la classe modale ;

n_{i-1} est l'effectif de la classe qui précède la classe modale ;

n_{i+1} est l'effectif de la classe qui suit la classe modale.

Remarque 2.1. On utilise la même formule pour calculer le mode avec

$\Delta_1 = f_i - f_{i-1}$ la différence des fréquences entre la classe modale et la classe précédente ;

$\Delta_2 = f_i - f_{i+1}$ la différence des fréquences entre la classe modale et la classe suivante.

Exemple 2.2. On considère le tableau statistique suivant.

Classes	[25-30[[30-35[[35-40[[40-45[[45-50[Total
n_i	10	35	55	60	50	200

Toutes les classes ont la même amplitude, la classe modale est [40 – 45]. Le mode est donné par

$$\begin{aligned}
 M_o &= e_i + a_i \times \frac{\Delta_1}{\Delta_1 + \Delta_2} \\
 &= 40 + 5 \times \frac{60 - 55}{(60 - 55) + (60 - 50)} \\
 &= 40 + 5 \times \frac{5}{5 + 10} \\
 &= 41,7.
 \end{aligned}$$

Remarque 2.2. Le choix du paramètre de tendance centrale dépend de l'objectif visé dans l'analyse d'une série statistique. Par exemple, dans le contexte économique, les fabricants d'une marque de téléviseurs cherchent à identifier les modèles les plus vendus pour établir une stratégie de production. Dans de telles situations, le mode se révèle particulièrement intéressant en tant que mesure de description des données.

2.2.2 La médiane notée M_e

Définition 2.3. La médiane

La médiane est la valeur de la V.S. qui sépare les données, triées par ordre croissant, en deux parties d'effectifs égaux. 50 % des valeurs sont supérieures à la médiane et 50 % des valeurs sont inférieures à la médiane. Il s'agit de la valeur de la variable statistique qui correspond à l'ordonnée $\frac{1}{2}$ dans la courbe cumulative c-à-d $F(M_e) = \frac{1}{2}$.

Calcul de la médiane : On distingue deux cas :

→ Dans le cas où les valeurs prises par la variable statistique ne sont pas regroupées en classes (Variable quantitative discrète) :

Soient x_1, x_2, \dots, x_n les n valeurs d'une variable statistique rangées par ordre croissant ($x_1 \leq x_2 \leq \dots \leq x_n$). Deux cas sont possibles :

Si $n = 2p$ (n est pair), on a $M_e = \frac{x_p + x_{p+1}}{2}$.

Si $n = 2p + 1$ (n est impair), on a $M_e = x_{p+1}$.

Exemple 2.3. L'ensemble des nombres 3, 3, 3,4, 4, 5, 6, 8, 8, 10, 11 a pour médiane 5. En effet, on a le nombre de données est $n = 11 = 2 \times 5 + 1$ est impaire donc la médiane est

$$M_e = x_6 = 5.$$

$$\overbrace{3,3,3,4,4}^{50\%}, \underbrace{5}_{M_e=5}, \overbrace{6,8,8,10,11}^{50\%}$$

Exemple 2.4. L'ensemble des nombres 5, 5, 6, 9, 11, 12, 15, 19.

On a $n = 8 = 2 \times 4$ est pair, donc $M_e = \frac{x_4+x_5}{2} = (9 + 11)/2 = 10$.

$$\overbrace{5, 5, 6}^{50\%}, \underbrace{9, 11}_{M_e=(9+11)/2}, \overbrace{12, 15, 19}^{50\%}.$$

→ Dans le cas où les valeurs de la variables sont regroupées en classes (Variable quantitative continue) :

$$M_e = e_i + \frac{0.5 - F_i}{F_{i+1} - F_i} \times a_i = e_i + \frac{\frac{1}{2} - F_i}{f_i} \times a_i,$$

où

$[e_i; e_{i+1}[$ la classe médiane : la classe pour laquelle l'effectif cumulé dépasse pour la 1ère fois

$N_i \geq \frac{n}{2}$ (la fréquence cumulée dépasse pour la 1ère fois ≥ 0.5);

e_i est la borne inférieure de la classe médiane;

a_i est l'amplitude de la classe médiane;

f_i est la fréquence de la classe médiane;

F_i est la fréquence cumulée de la classe qui précède la classe médiane;

F_{i+1} est la fréquence cumulée de la classe médiane.

Autre formulation

$$M_e = e_i + \frac{\frac{n}{2} - S}{n_{me}} \times a_i,$$

avec

n est la taille de l'échantillon étudié;

n_{me} est l'effectif de la classe médiane $[e_i - e_{i+1}[$;

S est l'effectif cumulé de la classe qui précède la classe médiane.

Exemple 2.5.

classes	$[10-20[$	$[20-30[$	$[30-40[$	$[40-50[$	$[50-60[$	$[60-70[$	Total
n_i	10	15	20	35	15	5	100
f_i	0.1	0.15	0.2	0.35	0.15	0.05	1
F_i	0.1	0.25	0.45	0.8	0.98	1	-

La classe médiane est celle qui a l'effectif cumulé $\geq \frac{n}{2} = 50$ ($F_i \geq 0.5$).

On a $F_3 = 0.45 < 0.5$ et $F_4 = 0.8 \geq 0.5$, la classe médiane est $[40 - 50[$.

$$M_e = 40 + \frac{0.5 - 0.45}{0.35} = 41.42$$

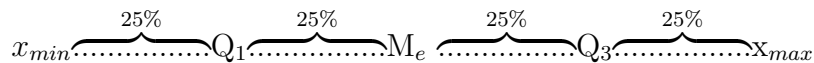
2.2.3 Les quartiles

Dans une série de données triées par ordre croissant, la médiane est la valeur qui divise l'ensemble des données en deux parties égales. En étendant ce concept, on intéresse dans cette section aux valeurs qui divisent l'ensemble des données en quatre parties égales. Ces valeurs sont notées Q_1 , Q_2 et Q_3 . Elles sont dites le premier, le deuxième et le troisième quartile, respectivement. Notons que le deuxième quartile Q_2 coïncide avec la médiane.

Question : Que peut-on déduire de la connaissance des paramètres Q_1 et Q_3 d'une série ?

Réponse :

- Au moins un quart (25%) des valeurs de la série sont inférieures ou égales à Q_1 .
- Au moins trois quarts (75%) des valeurs de la série sont inférieures ou égales à Q_3 .
- Environ (50%) des valeurs de la série se trouvent dans l'intervalle interquartile $[Q_1, Q_3]$.



Calcul des quartiles

Question : Comment déterminer les quartiles Q_1 et Q_3 d'une série ?

→ Dans le cas où les valeurs prises par la variable statistique ne sont pas regroupées en classe (Variable quantitative discrète) :

Soient x_1, x_2, \dots, x_n les n valeurs d'une variable statistique, ordonnées par ordre croissant : $x_1 \leq x_2 \leq \dots \leq x_n$.

On calcule la fraction $\frac{n}{4}$. Quatre cas sont possibles : selon que le résultat est entier ou pas.

Cas 1 : $n = 4q$:

$$Q_1 = \frac{x_q + x_{q+1}}{2} \quad Q_3 = \frac{x_{3q} + x_{3q+1}}{2}.$$

Cas 2 : $n = 4q + 1$:

$$Q_1 = x_{q+1} \quad Q_3 = x_{3q+1}.$$

Cas 3 : $n = 4q + 2$:

$$Q_1 = x_{q+1} \quad Q_3 = x_{3q+2}.$$

Cas 4 : $n = 4q + 3$:

$$Q_1 = x_{q+1} \quad Q_3 = x_{3q+3}.$$

Exemple 2.6. On observe le nombre d'accident par jours, au niveau d'une intersection, pour une période de 100 jours.

x_i (nombre d'accident)	0	1	2	3	4	5	6	7	Total
n_i (nombre de jours)	13	27	27	19	9	3	1	1	100
N_i	13	40	67	86	95	98	99	100	–

On a $100 = 4 \times 25$, $Q_1 = \frac{1+1}{2} = 1$, $Q_2 = M_e = \frac{2+2}{2} = 2$ et $Q_3 = \frac{3+3}{2} = 3$.

→ Dans le cas où les valeurs de la variable sont regroupées en classes (Variable quantitative continue) :

Supposons que le premier quartile (respectivement le troisième quartile) se situe dans la classe $[e_i, e_{i+1}[$, alors

$$Q_1 = e_i + a_i \frac{0.25 - F_i}{f_i},$$

$$\left(\text{resp. } Q_3 = e_i + a_i \frac{0.75 - F_i}{f_i} \right),$$

avec

e_i est la borne inférieure de la classe correspondante ;

a_i est l'amplitude de la classe correspondante $[e_i, e_{i+1}[$;

f_i est la fréquence de la classe correspondante $[e_i, e_{i+1}[$;

F_i est la fréquence cumulée de la classe qui précède la classe correspondante $[e_i, e_{i+1}[$.

Autres formulations

$$Q_1 = e_i + a_i \frac{\frac{n}{4} - S_1}{n_{Q_1}};$$

$$\left(\text{resp. } Q_3 = e_i + a_i \frac{\frac{3n}{4} - S_2}{n_{Q_3}} \right),$$

avec :

n est la taille de l'échantillon ;

n_{Q_i} est l'effectif de la classe qui contient le quartile Q_i ;

S_1 (resp. S_2) est l'effectif cumulé de la classe qui précède la classe qui contient le quartile Q_1 (resp. Q_3).

Remarque 2.3.

$Q_1 \in [e_i, e_{i+1}[$ signifie que $[e_i, e_{i+1}[$ est la première classe où $F_i \geq 0.25$ ($S_1 \geq \frac{n}{4}$).

$Q_3 \in [e_i, e_{i+1}[$ signifie que $[e_i, e_{i+1}[$ est la première classe où $F_i \geq 0.75$ ($S_2 \geq \frac{3n}{4}$).

Exemple 2.7. On considère un échantillon de 100 pesées en grs de pièces fabriquées dans une entreprise.

classes	[30-40[[40-50[[50-60[[60-70[[70-80[[80-90[Total
n_i	10	10	15	33	25	7	100
N_i	10	20	35	68	93	100	-

Le première effectif cumulé $\geq \frac{100}{4} = 25$ est $N_3 = 0.35$, donc Q_1 est dans la classe [50 – 60[.

$$Q_1 = 50 + 10 \times \frac{\frac{100}{4} - 20}{15} = 53.33grs$$

Le première effectif cumulé $\geq 3 \times \frac{100}{4} = 75$ est $N_5 = 93$, donc Q_3 est dans la classe [70 – 80[.

$$Q_3 = 70 + 10 \times \frac{3 \times \frac{100}{4} - 68}{25} = 72.8grs.$$

2.2.4 La moyenne

La moyenne représente le principal indicateur de tendance d'une série de données. La moyenne peut être qualifiée de différentes manières : arithmétique, géométrique, quadratique, harmonique.

Pour chacune de ces notions, on donnera dans ce qui suit, les expressions mathématiques qui permettent de les calculer.

- La moyenne simple : Le calcul est réalisé à partir d'un tableau où chaque élément ne correspond qu'à une seule information. Ou en se basant sur une série de données qui n'est pas récapitulée dans un tableau.

- La moyenne pondérée : est déterminée en utilisant un tableau statistique. Dans ce cas, l'effectif n_i correspond au poids de l'observation x_i .

Soit X la variable statistique étudiée. Soit x_1, x_2, \dots, x_n un échantillon de taille n .

Selon la nature de la variable, on peut les placer dans un tableau statistique. On a donc deux cas (voir Chapitre 1) :

On rappelle que dans le cas d'une variable quantitative discrète :

x_i	x_1	x_2	...	x_k	Total
Effectifs n_i	n_1	n_2	...	n_k	n
fréquences f_i	f_1	f_2	...	f_k	1

Dans le cas d'une variable quantitative continue :

Classes	$[e_1; e_2[$	$[e_2; e_3[$...	$[e_k; e_{k+1}[$	Total
Centres des classes	x_1	x_2	...	x_k	-
Effectifs n_i	n_1	n_2	...	n_k	n

Pour le calcul des moyennes dans le cas d'une variable quantitative continue les x_i sont les centres des classes.

La moyenne arithmétique

La moyenne arithmétique est le paramètre de tendance centrale le plus utilisée pour caractériser une série d'observations. On donne la définition de la moyenne comme suit.

- La moyenne arithmétique simple des n observations est définie par la somme des n observations x_i divisée par le nombre n .

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i.$$

- La moyenne arithmétique pondérée : pour des données classées dans un tableau statistique des effectifs (ou des fréquences), la moyenne est définie par

- En utilisant les effectifs :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k n_i x_i.$$

- En utilisant les fréquences relatives :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k n_i x_i = \sum_{i=1}^k \frac{n_i}{n} x_i = \sum_{i=1}^k f_i x_i.$$

Remarque 2.4. .

- $\sum_{i=1}^k n_i (x_i - \bar{X}) = 0$
- $\sum_{i=1}^k n_i (x_i - \alpha)^2 = 0$ est minimal pour $\alpha = \bar{X}$.

Exemple 2.8. Les notes sur 20 obtenues en Statistique par une étudiante sont les suivantes :

Contrôle	Contrôle1	contrôle2	partiel
Notes	12	14	11

La moyenne semestrielle de cette étudiante si les notes ont le même coefficient : $\bar{x} = \frac{12+14+11}{3} = 12.33$.

Si les deux contrôles continus sont affectés du coefficient 2 et le partiel affecté du coefficient 3, $\bar{x} = \frac{2 \times 12 + 2 \times 14 + 3 \times 11}{7} = \frac{85}{7}$.

Exemple 2.9. Une voiture roule durant une heure à la vitesse constante de 110 km/h et durant une heure à la vitesse constante de 130 km/h.

Question : Calculer la vitesse constante à laquelle la voiture aurait dû rouler pendant la durée totale (2h) du trajet pour faire la même distance ?

Réponse : Soit d la distance parcourue, v la vitesse et t le temps. On sait que

$$d = v \times t.$$

La première partie a duré 1 heure en roulant à la vitesse de 110 km/h, ce qui fait que la voiture a fait 110 km.

La deuxième partie a duré 1 heure en roulant à la vitesse de 130 km/h, donc la voiture a parcouru 130 km.

Par conséquent, la vitesse moyenne de la voiture v vérifie : $110 + 130 = v \times (1 + 1)$, ce qui donne $v = \frac{110+130}{2}$ km/h.

Ceci donne la moyenne arithmétique de 110 et de 130.

Définition 2.4. (Symétrie)

Les trois valeurs centrales d'une série statistique sont : la moyenne, le mode et la médiane. Ces paramètres résument l'ensemble des valeurs de la série considérée. Si dans la série statistique le mode est unique, on a :

- Une série de données X est symétrique si les observations sont réparties de façon équiprobable autour de la moyenne \bar{X} . Dans ce cas on a $\bar{X} = M_o = M_e$. Sinon, la série est dissymétrique (assymétrique).
- Une série X est asymétrique à droite si son diagramme s'étale à droite : $M_o < M_e \leq \bar{X}$. Elle possède un étalement à droite et une oblique à gauche.
- Une série est asymétrique à gauche lorsque son diagramme s'étale à gauche : $\bar{X} \leq M_e < M_o$. Elle possède un étalement à gauche et une oblique à droite.

La moyenne géométrique notée \bar{G}

On sert de la moyenne géométrique \bar{G} dans le calcul des taux moyens.

Elle est calculée uniquement pour des nombres positifs.

- La moyenne géométrique simple : La moyenne géométrique de n valeurs positives x_i est la racine n ième du produit de ces valeurs. Notée \bar{G} , c'est à dire

$$\bar{G} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \left[\prod_{i=1}^n x_i \right]^{\frac{1}{n}}.$$

Il est plus simple de calculer $\log \bar{G}$.

$$\log \bar{G} = \frac{1}{n} \log \left(\prod_{i=1}^n x_i \right) = \frac{1}{n} \sum_{i=1}^n \log x_i.$$

$$\bar{G} = e^{\frac{1}{n} \sum_{i=1}^n \log x_i}$$

- La moyenne géométrique pondérée : Pour des données qui sont regroupées dans un tableau statistique des effectifs (avec $x_i > 0$), on a

$$\bar{G} = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_k^{n_k}} = \left[\prod_{i=1}^k x_i^{n_i} \right]^{\frac{1}{n}}$$

$$\log \bar{G} = \frac{1}{n} \log \left(\prod_{i=1}^k x_i^{n_i} \right) = \frac{1}{n} \sum_{i=1}^k n_i \log x_i$$

$$\bar{G} = e^{\frac{1}{n} \sum_{i=1}^k n_i \log x_i}$$

Exemple 2.10. .

Le prix d'un produit a augmenté en janvier d'une année donnée, de 0,8 % et en février de la même année de 0,7 %.

En notant P_1 le prix en janvier, on a

- on obtient au premier du mois de février, $P_2 = P_1 + \frac{0,8}{100} \times P_1 = P_1 \times (1 + 0,008) = P_1 \times 1,008$;

2.2. LES PARAMÈTRES DE TENDANCE CENTRALE ET DE POSITION

-et au 1er mois de mars, on obtient $P_3 = P_2 + \frac{0.7}{100}P_2 = P_2 \times 1.007 = P_1 \times 1.008 \times 1.007$.

- Si on note le pourcentage d'augmentation sur les deux premiers mois par T . Donc,

$$P_2 = P_1 \times \left(1 + \frac{T}{100}\right) \quad \text{et} \quad P_3 = P_2 \times \left(1 + \frac{T}{100}\right),$$

donc

$$P_3 = P_1 \times \left(1 + \frac{T}{100}\right)^2.$$

Le taux T est réalisé l'équation

$$\left(1 + \frac{T}{100}\right)^2 = 1.008 \times 1.007.$$

$1 + \frac{t}{100} = \sqrt{1.008 \times 1.007} = 1.0075$ est la moyenne géométrique de 1.008 et de 1.007.
 $t = 0.75\%$.

La moyenne harmonique

On utilise cette moyenne pour les grandeurs quotients. Par exemple les vitesses.

- Soient n valeurs positives x_i , la moyenne harmonique simple est définie par x_i :

$$\overline{H} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}.$$

- La Moyenne harmonique pondérée : pour des données résumées dans un tableau statistique des effectifs, la moyenne harmonique est définie par

$$\overline{H} = \frac{n}{\sum_{i=1}^k \frac{n_i}{x_i}}.$$

Exemple 2.11. .

Ines se déplace en voiture. Sa vitesse à l'aller est de 10 km/h et celle au retour est de 20 km/h. Soit D la distance en km du trajet à l'aller.

Pour calculer sa vitesse moyenne v sur tout le parcours, on procède comme suit.

- La distance totale parcourue est $2D$.

- Soit t_1 le temps (en h) pour effectuer le trajet de l'aller, soit t_2 le temps (en h) pour effectuer le trajet du retour. On a $D = 10 \times t_1$ et $D = 20 \times t_2$, donc

- le temps pour l'aller : $t_1 = \frac{D}{10}$;
- le temps pour le retour : $t_2 = \frac{D}{20}$;
- le temps total $t_1 + t_2 = \frac{D}{10} + \frac{D}{20}$.
- Si on note par $v_1 = 10\text{km/h}$ et $v_2 = 20\text{ km/h}$. La vitesse moyenne v sur le trajet aller retour vérifie $v = \frac{2 \times D}{t_1 + t_2}$. Donc $t_1 + t_2 = \frac{2 \times D}{v}$. Par conséquent,

$$\frac{2 \times D}{v} = \frac{D}{v_1} + \frac{D}{v_2}.$$

- On obtient

$$v = \frac{2}{\frac{1}{v_1} + \frac{1}{v_2}} = 13,33\text{km/h}$$

. C'est la moyenne harmonique de 10 et de 20.

La moyenne quadratique notée \bar{Q}

- Moyenne quadratique simple : des n observations x_i , est définie par

$$\bar{Q} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}.$$

- La Moyenne quadratique pondérée : pour des données résumées dans un tableau statistique des effectifs, la moyenne quadratique est définie par

$$\bar{Q} = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i x_i^2}.$$

Exemple 2.12. Soient deux disques de rayons respectifs 8 cm et 9 cm.

Questions :

- 1- Calculer la somme des aires de ces disques.
 - 2- Quel est le rayon par lequel il faut remplacer les deux rayons pour obtenir la même somme ?
- L'aire du disque de rayon 8 cm est $A_1 = \pi 8^2\text{ cm}^2$ et l'aire du disque de rayon 9cm est $A_2 = \pi \times 9^2\text{ cm}^2$. Par conséquent, la somme est $\pi \times (8^2 + 9^2)\text{ cm}^2$.
 - On cherche le rayon R qui permet de réaliser $\pi \times (8^2 + 9^2) = \pi \times R^2$, donc $8^2 + 9^2 = 2R^2$, et $R^2 = \frac{8^2+9^2}{2}$.
 - $R > 0$ donc $R = \sqrt{\frac{8^2+9^2}{2}}$.
 - R est simplement la moyenne quadratique de 8 et de 9.

Remarque 2.5. *Sur un même échantillon, on a les inégalités*

$$\overline{H} < \overline{G} < \overline{X} < \overline{Q} \quad \text{et} \quad \overline{G}^2 = \overline{H} \times \overline{X}.$$

2.3 Paramètres de dispersion (caractéristiques de dispersion)

Pour décrire deux séries de données relatives aux même phénomène, les mesures de tendance centrale peuvent s'avérer insuffisantes. Par exemple, si on considère les salaires horaires (en DA) dans une entreprise

495; 497; 500; 503; 505

Et dans une autre entreprise

450; 475; 500; 525; 550

Dans les deux séries de données, la médiane et la moyenne sont identiques et constantes : $\overline{X} = 500$ DA et $M_e = 500$ DA. Cependant, ces deux séries se distinguent et ce qui les distingue réside dans la dispersion ou la variabilité des données. Dans cette situation, la deuxième série présente une dispersion bien plus importante que la première. Par exemple, on constate que leurs écarts à la moyenne $x_i - \overline{X}$ sont différents.

5; 3; 0; 3; 5

50; 25; 0; 25; 50

Par conséquent, il est essentiel de décrire une série statistique en utilisant non seulement des paramètres de position, mais également des paramètres de dispersion.

Les paramètres de dispersion sont employés afin d'évaluer la variabilité de la série. En d'autres termes, il s'agit de décrire la distance entre toutes les données et le paramètre de tendance centrale.

Il existe plusieurs mesures de la dispersion et les plus utilisés sont : L'étendue, l'écart absolu moyen, l'écart interquartile, l'écart type et le coefficient de variation.

2.3.1 L'étendu

Définition 2.5. *L'étendu est l'écart entre la plus grande et la plus petite valeur d'une variable statistique quantitative.*

L'étendu d'un ensemble de valeur x_1, x_2, \dots, x_k d'une variable statistique X est définie par

$$E = \max(x_i) - \min(x_i).$$

Exemple 2.13. *L'étendue de l'ensemble 2, 3, 3, 5, 5, 8, 10, 12, 13, 15, 16, 20, 23 est $E = 23 - 2 = 21$.*

2.3.2 L'écart absolu moyen

Définition 2.6. *L'écart absolu moyen noté (e) de la variable statistique X est définie comme la moyenne arithmétique des valeurs absolues des écarts à la moyenne arithmétique.*

Il est donné par

$e = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{X}|$, si on considère que les n valeurs données x_1, x_2, \dots, x_n sont toutes évaluées.

$e = \frac{1}{n} \sum_{i=1}^k n_i |x_i - \bar{X}|$, la série possède k valeurs distinctes x_1, x_2, \dots, x_k avec des effectifs d'apparitions n_1, n_2, \dots, n_k , respectivement.

$e = \frac{1}{n} \sum_{i=1}^k n_i |x_i - \bar{X}|$, les données de la série sont regroupées en k classes d'effectifs d'apparitions n_1, n_2, \dots, n_k , x_i sont le centre de la classe $[e_i; e_{i+1}[$.

Remarque 2.6. *Si on possède des fréquences relatives f_i au lieu des effectifs n_i , on calcule l'étendu en remplaçant les fractions $\frac{n_i}{n}$ par f_i dans les expressions de l'étendu.*

Exemple 2.14. *Ce paramètre montre la différence de dispersion entre les deux séries*

49; 497; 500; 503; 505

et

450; 475; 500; 525; 550.

On a pour la première série

$$e = \frac{|495 - 500| + |497 - 500| + |500 - 500| + |503 - 500| + |505 - 500|}{5} = \frac{16}{5} = 3,2$$

et pour la deuxième série

$$e = \frac{|450 - 500| + |475 - 500| + |500 - 500| + |525 - 500| + |550 - 500|}{5} = \frac{150}{5} = 30.$$

2.3.3 L'écart interquartile

Définition 2.7. On définit par

l'intervalle interquartile, l'intervalle $[Q_1, Q_3]$. Il contient 50% des observations, en laissant 25% à gauche et 25% à droite. Il contient la moitié des observations.

l'écart interquartile, la différence entre le troisième quartile et le premier quartile

$$EQ = Q_3 - Q_1.$$

Exemple 2.15. On reprend l'Exemple 2.7, on a trouvé $Q_1 = 53.33$ grs et $Q_3 = 72.8$ grs. Donc $EQ = 72.8 - 53.33 = 19.47$ grs.

2.3.4 La variance et L'écart-type

Les mesures de dispersion les plus utilisées sont l'écart-type et la variance.

Définition 2.8. *La variance* est la moyenne des carrés des écarts par rapport à la moyenne arithmétique.

• Dans le cas où les valeurs x_1, x_2, \dots, x_n de la variable statistique X ne se répètent pas, la variance est donnée par

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

Pour la simplicité de calcul, on utilise le théorème de Koenig

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

• Si à chaque valeur x_1, x_2, \dots, x_k de la variable statistique X correspond des effectifs d'apparition n_1, n_2, \dots, n_k , respectivement, la variance est définie par

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{X})^2.$$

- En utilisant les fréquences relatives, la formule de la variance est

$$\sigma^2 = \sum_{i=1}^k f_i(x_i - \bar{X})^2.$$

- Dans le cas où les données sont regroupées en k classes de fréquences d'apparitions n_1, n_2, \dots, n_k , avec x_i est le centre de la classe $[e_i; e_{i+1}[$, la variance est donnée par

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^k n_i(x_i - \bar{X})^2$$

Remarque 2.7. En développant la somme des différences au carré dans la formule

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^k n_i(x_i - \bar{X})^2,$$

on peut simplifier l'expression comme suit $\sigma^2 = \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - \bar{X}^2$.

Définition 2.9. *L'écart-type* est donné comme étant la racine carrée de la variance

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i(x_i - \bar{X})^2}$$

Dans le cas où les valeurs de la variable statistique se répètent, l'écart-type devient

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i(x_i - \bar{X})^2} = \sqrt{\sum_{i=1}^k f_i(x_i - \bar{X})^2}$$

Remarque 2.8.

- La variance est toujours un nombre réel positif $\sigma^2 \geq 0$. Elle mesure la dispersion (la variabilité des données) de la variable statistique autour de la moyenne. Son unité est le carré de l'unité de la variable statistique.
- Il peut être compliqué d'utiliser la variance comme indicateur de dispersion car le carré entraîne une modification des unités. Ainsi, il n'y a pas de signification directe, à la différence de l'écart-type qui se manifeste dans les mêmes unités que la moyenne.
- L'écart-type d'une série est d'autant plus bas (c'est-à-dire proche de 0) qu'elle sera regroupée autour de la moyenne, avec des valeurs très proches.

Exemple 2.16. *On reprend les deux séries de salaires horaires*

495; 497; 500; 503; 505

et

450; 475; 500; 525; 550

On calcule la variance de chacune, on trouve

pour la première série $\sigma^2 = 13,6$, $\sigma = 3.69$

et pour la deuxième série $\sigma_2^2 = 1250$, $\sigma = 35.36$.

Le rapport entre les écarts-types des deux séries est proche du rapport des écarts absolu moyen.

Ces deux paramètres donnent les mêmes informations.

2.3.5 Coefficient de variation

Le coefficient de variation est un coefficient qui ne possède aucune unité et est exprimé en pourcentage. Son utilisation consiste à comparer des séries de valeurs dont les échelles de mesure ne sont pas identiques. Par exemple, il permet de faire une comparaison des différences salariales entre deux pays qui utilisent deux monnaies différentes. La formule donne le coefficient de variation.

$$C_v = \frac{\sigma}{\bar{X}}.$$

Remarque 2.9. *Un coefficient de variation élevé indique une dispersion importante autour de la moyenne arithmétique.*

2.4 Exercices (Série 2)

Exercice 2.1. La représentation des femmes âgées de 40 à 50 ans d'après le nombre de leurs enfants vivants est donnée par le tableau ci-dessous.

nbre d'enfants x_i	0	1	2	3	4	Total
Effectif n_i	20	20	30	10	20	100

- 1 - Quelle est la population observée ?
- 2 - Donner le caractère étudié et préciser sa nature.
- 3 - Dresser le tableau des effectifs, des fréquences relatives, des effectifs cumulés croissants et des fréquences cumulées croissantes.
- 4 - Quel est le pourcentage des femmes ayant moins de 2 enfants ?
- 5 - Quel est le pourcentage des femmes dont le nombre d'enfants vivants est au plus 3 ?
- 6 - Quel est le pourcentage des femmes dont le nombre d'enfants vivants est au moins 2 ?
- 7 - Donner la représentation graphique adéquate.
- 8 - Calculer les paramètres de position : le mode, la médiane, le premier et le troisième quartile et la moyenne arithmétique.
- 9 - Calculer : l'étendu, l'écart-type et le coefficient de variation.

Exercice 2.2. Le montant des ventes d'un magasin et le nombre d'articles vendus pendant une année sont donnés dans le tableau statistique

Classes	[15-20[[20-25[[25-30[[30-35[[35-40[[40-45[Total
Effectifs n_i	70	70	50	70	80	60	400

- 1- Quel est le type de la variable étudiée ?
- 2- Dresser le tableau des effectif cumulés, fréquences relatives, et fréquence cumulées.
- 3- Donner la représentation graphique adéquate.
- 4- Calculer le mode, la médiane, les quartiles et la moyenne arithmétique. Comparer les paramètres \bar{X} , M_e et M_o et donner une interprétation.
- 5- Calculer l'écart interquartile, l'écart type ainsi que le coefficient de variation.
- 6- Donner un intervalle qui contient 50% de valeurs centrales. Que représente les bornes de cet intervalle ? interprétez.

Exercice 2.3. Un agent de police a relevé les vitesses(km/h) suivantes :

120 – 128 – 93 – 102 – 111 – 148 – 118 – 92 – 158 – 136 – 127 – 117
 104 – 119 – 121 – 102 – 125 – 124 – 88 – 138 – 103 – 115 – 132 – 99

2.4. EXERCICES (SÉRIE 2)

115 – 132 – 128 – 141 – 122 – 115 – 107 – 126 – 132 – 144 – 110 – 137

112 – 102 – 126 – 145 – 135 – 107 – 123 – 96 – 108 – 122 – 120 – 124

151 – 152 – 128 – 121 – 118 – 131 – 108 – 141 – 143 – 146 – 108 – 113.

- 1) *Quelle est la population étudiée ? Quelle la variable statistique étudiée ? Quelle est sa nature ? Quelle est la taille de l'échantillon étudié ?*
- 2) *Donner l'étendu de la série statistique.*
- 3) *En utilisant la règle de Sturges, répartir les données en des classes de même amplitude.*
- 4) *Donner le tableau statistique des effectifs, effectifs cumulés, fréquences et fréquences cumulées.*
- 5) *Calculer le mode et la médiane. Donner une interprétation des résultats.*
- 6) *Tracer l'histogramme de la série.*
- 7) *Calculer graphiquement le mode.*
- 8) *Tracer le polygone statistique.*
- 9) *Calculer les quartiles et l'écart interquartile.*
- 10) *Calculer la moyenne et la variance.*
- 11) *Comparer les paramètres \bar{X} , M_o et M_e . Conclure.*

Exercice 2.4. *On a mesuré la taille (en cm) de 100 étudiants de la section C.*

Taille (cm)	[130-140[[140-150[[150-160[[160-170[[170-180[[180-190[[190-200[
Nbre d'étudiants	4	8	14	16	30	20	8

- 1) *Quelle est la population étudiée ? Quelle est la variable étudiée ? Quelle est sa nature ?*
- 2) *Donner le tableau statistique des effectifs, fréquences, effectifs cumulés et fréquences cumulées.*
- 3) *Calculer les mesures de tendance centrale (la moyenne arithmétique, le mode et la médiane), puis comparer ces trois paramètres. La distribution est-elle symétrique ? si non dans quel sens présenterait-elle un étalement ?*
- 4) *Représenter ces 3 paramètres sur l'histogramme.*

2.5 Solution des exercices

Solution de l'exercice 2.1. .

1- La population observée : les femmes âgées de 40 à 50 ans.

2- Le caractère : nombre d'enfants vivants. Sa nature : quantitatif discret.

3- Le tableau statistique

x_i	0	1	2	3	4	Total
n_i	20	20	30	10	20	100
N_i	20	40	70	80	100	/
f_i	0.2	0.2	0.3	0.1	0.2	1
F_i	0.2	0.4	0.7	0.8	1	/

4- Le pourcentage des femmes ayant moins de 2 enfants est :

$$(0.2 + 0.2) \times 100\% = 40\%.$$

5 - Le pourcentage des femmes dont le nombre d'enfants vivants est au plus 3 est

$$(0.2 + 0.2 + 0.3 + 0.1) \times 100\% = 80\%.$$

6 - Le pourcentage des femmes dont le nombre d'enfants vivants est au moins 2 est :

$$(0.3 + 0.1 + 0.2) \times 100\% = 60\%.$$

7- La représentation graphique :

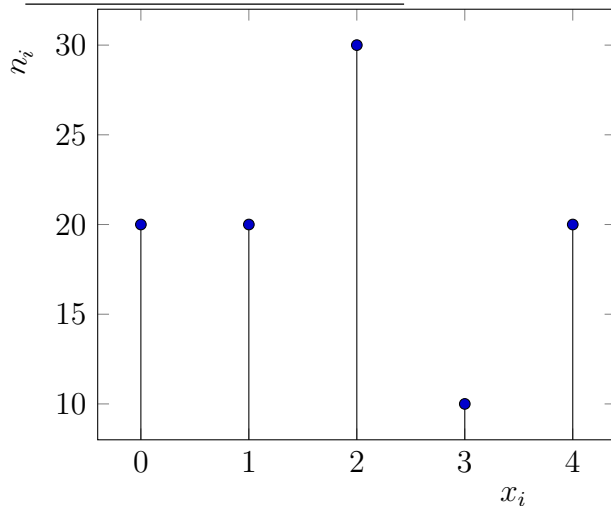


FIGURE 2.1 – Le diagramme en bâtons

2.5. SOLUTION DES EXERCICES

8 - Les paramètres de position

- Le mode M_o : est la modalité la plus répétée est $M_o = 2$ ($n_3 = 30$ est l'effectif le plus élevé).
- La médiane : on a $n = 100$ est pair donc $n = 2 \times 50$, posons $k = 50$.

$$M_e = \frac{x_k + x_{k+1}}{2} = \frac{2 + 2}{2} = 2.$$

(x_{50} et x_{51} sont déterminés en utilisant les effectifs cumulés N_i).

50% des femmes ont un nombre d'enfants ≤ 2 .

- Le premier quartile et le troisième quartile : Soit x_q et x_{q+1} les valeurs des observations classées respectivement en q ème et $(q+1)$ ème position (lorsque les observations sont classées par ordre croissant).

On a $n = 100 = 4 \times 25$, posons $q = 25$.

$$Q_1 = \frac{x_q + x_{q+1}}{2} = \frac{1 + 1}{2} = \frac{1}{2} \quad \text{et} \quad Q_3 = \frac{x_{3q} + x_{3q+1}}{2} = \frac{3 + 3}{2} = 3.$$

(x_{25} , x_{26} , x_{75} et x_{76} sont déterminés en utilisant les effectifs cumulés N_i).

50% des femmes ont le nombre d'enfants entre 1 et 3.

- La moyenne arithmétique \bar{x} :

x_i	0	1	2	3	4	Total
n_i	20	20	30	10	20	100
$n_i x_i$	0	20	60	30	80	190
$n_i x_i^2$	0	20	120	90	320	550

$$\bar{x} = \frac{1}{6} \sum_{i=1}^{i=5} n_i x_i = \frac{190}{100} = 1,9.$$

On a $\bar{X} \leq M_o = M_e$. La série statistique n'est pas symétrique. Elle présente un étalement vers la gauche.

8 - Les paramètres de dispersion :

- L'étendu : $E = \max(x_i) - \min(x_i) = 4 - 0 = 4$.
- L'écart-type : $\sigma = \sqrt{V(X)}$, où $V(X)$ est la variance et on a

$$V(X) = \frac{1}{100} \sum_{i=1}^{i=5} n_i x_i^2 - \bar{x}^2 = \frac{550}{100} - (1,9)^2 = 1,99.$$

2.5. SOLUTION DES EXERCICES

Donc $\sigma = \sqrt{1.89} = 1.37$.

- Le coefficient de variation : donné par $C_v = \frac{\sigma}{\bar{x}} = \frac{1.37}{1.9} = 0.72$

Solution de l'exercice 2.2. .

1- La variable est quantitative continue.

2- Le tableau statistique

Classe	[15-20[[20-25[[25-30[[30-35[[35-40[[40-45[Total
n_i	70	70	50	70	80	60	400
N_i	70	140	190	260	340	400	/
f_i	0.175	0.175	0.125	0.175	0.2	0.15	1
F_i	0.175	0.35	0.475	0.65	0.85	1	/

3- La représentation graphique

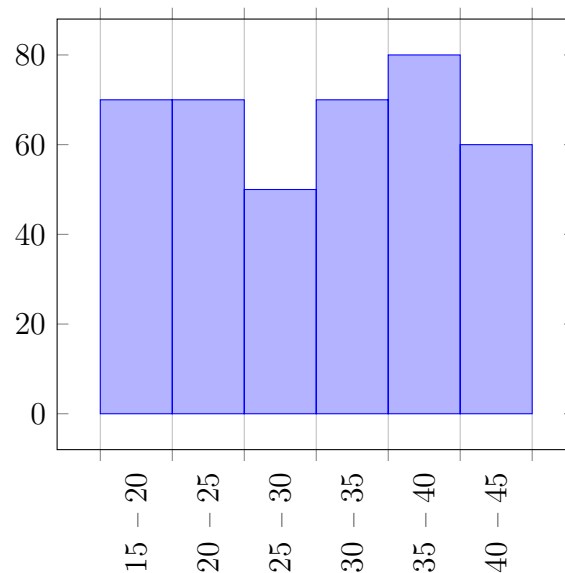


FIGURE 2.2 – L'histogramme des effectifs

4-

-Le mode :

La classe modale est [35 – 40[(la classe possédant le plus grand effectif).

La borne inférieure de la classe modale $e_5 = 35$.

2.5. SOLUTION DES EXERCICES

L'amplitude de la classe modale est $a_5 = 5$.

$$M_o = e_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times a_i = 35 + \frac{80 - 70}{(80 - 70) + (80 - 60)} \times 5,$$

d'où $M_O = 36.67$.

La médiane : La classe médiane est $[30 - 35[$, (la classe pour laquelle la fréquence cumulée dépasse pour la 1ère fois 0.5 - l'effectif cumulé dépasse pour la 1ère fois $\frac{400}{2}$).

On a : $e_4 = 30$ est la borne inférieure de la classe médiane.

$a_4 = 5$ est l'amplitude de la classe médiane.

$S = N_3 = 190$ est l'effectif cumulé de la classe $[25 - 30[$.

$n_{me} = n_4 = 70$ est l'effectif de la classe médiane.

$$M_e = e_i + \frac{\frac{n}{2} - S}{n_{me}} \times a_i = 30 + \frac{200 - 190}{70} \times 5 = 30,71.$$

- Le premier quartile : On a $Q_1 \in [20 - 25[$ (la classe pour laquelle la fréquence cumulée dépasse pour la 1ère fois 0.25), notée $[e_2 - e_3[$.

$n_2 = 70$ est l'effectif de la classe $[20 - 25[$.

$N_1 = 70$ est l'effectif cumulé croissant de la classe $[15 - 20[$.

$a_2 = 5$ est l'amplitude de la classe $[20 - 25[$.

$$Q_1 = e_2 + \frac{\frac{n}{4} - N_1}{n_2} \times a_2 = 20 + \frac{100 - 70}{70} \times 5 = 22.14.$$

Le troisième quartile : On a $Q_3 \in [35 - 40[$ (la classe pour laquelle la fréquence cumulée dépasse pour la 1ère fois 0.75), notée $[e_5 - e_6[$.

$n_5 = 80$ est l'effectif de la classe $[35 - 40[$.

$N_4 = 260$ est l'effectif cumulé de la classe $[30 - 35[$.

$a_5 = 5$ est l'amplitude de la classe $[35 - 40[$.

$$Q_3 = e_5 + \frac{\frac{3 \times n}{4} - N_4}{n_5} \times a_5 = 35 + \frac{300 - 260}{80} \times 5 = 37.5$$

- La moyenne arithmétique :

Classe	[15-20[[20-25[[25-30[[30-35[[35-40[[40-45[Total
Centre(x_i)	17.5	22.5	27.5	32.5	37.5	42.5	/
n_i	70	70	50	70	80	60	400
$n_i x_i$	1225	1575	1375	2275	3000	2550	12000
$n_i x_i^2$	21437.5	35437.5	37812.5	73937.5	112500	108375	389500

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i=6} n_i x_i = \frac{12000}{400} = 30.$$

$\bar{X} < M_e < M_o$, la distribution n'est pas symétrique. Elle présente un étalement vers la gauche.

5-

- L'écart interquartile : $E_Q = Q_3 - Q_1 = 37.5 - 22.14 = 15.30$.

- L'écart type : L'écart-type $\sigma = \sqrt{V(X)}$, où $V(X)$ est la variance de X donnée par

$$V(X) = \frac{1}{n} \sum_{i=1}^{i=6} n_i x_i^2 - \bar{x}^2 = \frac{389500}{400} - 30^2 = 73.75.$$

Donc $\sigma = \sqrt{73.75} = 8.59$.

- Le coefficient de variation : $C_v = \frac{\sigma}{\bar{x}} = \frac{8.59}{30} = 0.286$.

6- L'intervalle qui contient 50% de valeurs centrales est l'intervalle interquartile, délimité par le premier quartile et le troisième quartile : $Q_3 = 37.5$ et $Q_1 = 22.14$.

C'est-à-dire 50% des articles vendus ont un montant compris entre 22.14 et 37.5.

Solution de l'exercice 2.3. .

1) La population étudiée : L'ensemble des voitures.

La variable statistique étudiée : X : La vitesse.

La nature de la variable est quantitative continue.

La taille de l'échantillon est $n = 60$.

2) L'étendu de la série est donné par $e = x_{max} - x_{min} = 158 - 88 = 70$.

3) Le nombre de classes : on calcule $1 + 3.3 \log_{10}(60) = 6.87$, donc le nombre de classe est $k = 7$.

L'amplitude des classes : $a = \frac{x_{max} - x_{min}}{k} = \frac{70}{7} = 10$.

Les classes : $[88 - 98[$, $[98 - 108[$, $[108 - 118[$, $[118 - 128[$, $[128 - 138[$, $[138 - 148[$, $[148 - 158]$.

2.5. SOLUTION DES EXERCICES

4) Le tableau statistique :

Classes	[88-98[[98-108[[108-118[[118-128[[128-138[[138-148[[148-158[Total
n_i	4	8	11	16	10	7	4	60
f_i	0.06	0.13	0.18	0.27	0.17	0.12	0.07	1
N_i	4	12	23	39	49	56	60	/
F_i	0.06	0.19	0.37	0.64	0.81	0.93	1	/

5) • Le mode M_o : $M_o \in [118 - 128[\Rightarrow M_o = 118 + \frac{(16-11)}{(16-11)+(16-10)} \times 10 = 122.54$.

Le plus grand nombre de voiture passent avec une vitesse de 122.54km/h.

• La médiane M_e : $M_e \in [118 - 128[\Rightarrow M_e = 118 + \frac{30-23}{16} \times 10 = 122.37$.

50% des voitures passent avec une vitesse ≤ 122.37 km/h.

6) L'histogramme des effectifs

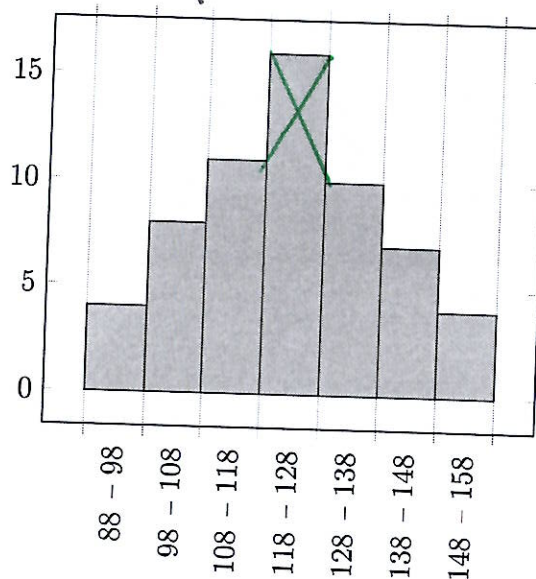


FIGURE 2.3 – L'histogramme des effectifs

7) Le mode graphiquement est donné par l'abscisse du point d'intersection des diagonales (voir la figure).

8) Le polygone statistique des effectifs : (lignes brisées de couleur rouge).

On obtient le polygone des effectifs en joignant par des segments de droites les milieux des bases supérieures des rectangles de l'histogramme.

2.5. SOLUTION DES EXERCICES

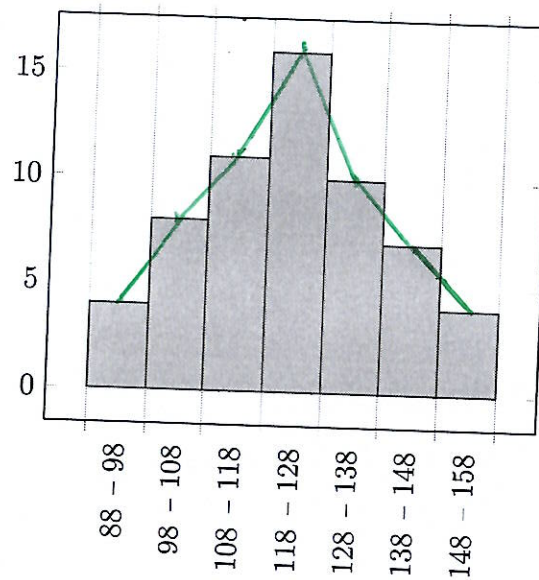


FIGURE 2.4 - Le polygone des effectifs

9) Les quartiles et l'écart interquartile

$$Q_1 \in [108 - 118[\implies Q_1 = 108 + 10 \times \frac{\frac{60}{4} - 12}{11} = 110.72;$$

$$Q_3 \in [128 - 138[\implies Q_3 = 128 + 10 \times \frac{\frac{3 \times 60}{4} - 39}{10} = 134.$$

L'écart interquartile $EQ = Q_3 - Q_1 = 134 - 110.72 = 23.28$.

50% des voitures passent avec une vitesse entre 110.72 km/h et 134 km/h.

10) La moyenne et la variance : on note par x_i , $i = 1, 2, \dots, 7$ les centres des classes

Classes	[88-98[[98-108[[108-118[[118-128[[128-138[[138-148[[148-158[Total
x_i	93	103	113	123	133	143	153	/
n_i	4	8	11	16	10	7	4	60
$n_i x_i$	372	824	1243	1968	1330	1001	612	7350
$n_i x_i^2$	34596	84872	140459	242064	176890	143143	93636	915660

La moyenne arithmétique $\bar{X} = \frac{1}{60} \sum_{i=1}^{i=7} n_i x_i = \frac{7350}{60} = 122.5$.

La variance $V(X) = \frac{1}{60} \sum_{i=1}^{i=7} n_i x_i^2 - \bar{X}^2 = \frac{915660}{60} - (122.5)^2 = 254,75$. La série statistique

2.5. SOLUTION DES EXERCICES

est dispersée.

11) On a $M_o < M_e < \bar{X}$. La série n'est pas symétrique.

Chapitre 3

Calcul des probabilités



3.1 Introduction

Les probabilités servent à modéliser une expérience aléatoire.

On commence par décrire les différentes issues possibles (éventualités) de cette expérience aléatoire. Ensuite, l'objectif est d'associer à chacune de ces éventualités un nombre compris entre 0 et 1 qui mesure la chance qu'elles ont de se réaliser. Ce chapitre est basé sur les références [\[1, 5, 8, 12, 13, 10, 7\]](#)

3.2 Ensemble fondamental et événement

Le mot "Aléatoire" vient du latin "aléa" qui signifie "les dés".

Un phénomène aléatoire est un processus lié au hasard.

Expérience aléatoire : On considère qu'une expérience est aléatoire lorsque elle vérifie les

trois conditions :

- on connaît tous les résultats éventuels de l'expérience ;
- le résultat de l'expérience n'est pas prévisible. En d'autres termes, il est impossible de déterminer avec certitude à l'avance.
- L'expérience peut être reproduite dans les mêmes conditions.

→ L'ensemble des résultats possibles d'une expérience aléatoire est dit **ensemble fondamental (ou univers de l'expérience aléatoire)**. On le note Ω .

La notion de probabilité est utilisée pour évaluer les chances qu'un événement se produise. La représentation précise de l'ensemble fondamental Ω est indispensable pour cette quantification.

Définition 3.1. .

- **Un événement élémentaire** est un singleton $\{\omega\} \subset \Omega$.
- **Un événement aléatoire** (ou un événement) A est un sous ensemble de Ω : $A \subset \Omega$ ($A \in \mathcal{P}(\Omega)$). Où $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω .

Exemple 3.1. On lance un dé cubique équilibré, à six faces numérotées de 1 à 6, et on note le résultat obtenu.

C'est une expérience aléatoire à six issues :

- Il y a six résultats possibles : $\Omega = \{1, 2, 3, 4, 5, 6\}$;
- On ne peut pas prévoir le résultat de cette expérience ;
- et on peut refaire l'expérience plusieurs fois.

Un exemple d'un événement élémentaire est : "on obtient le nombre 6".

Exemple 3.2. On lance une pièce de monnaie et on regarde la face supérieure. Il ya deux résultats possibles : Pile (P) et Face (F).

C'est une expérience aléatoire, on a $\Omega = \{P, F\}$.

Exemple 3.3. On observe la durée de vie d'une lampe.

C'est une expérience aléatoire. On a $\Omega = [0, +\infty[$.

Définition 3.2. .

- $A = \emptyset$ est dit **événement impossible** : c'est un événement qui ne peut pas se réaliser.
- $A = \Omega$ est dit **événement certain**.
- Soit A un événement. Soit ω le résultat de l'expérience. On dit que A se réalise $\iff \omega \in A$.

Exemple 3.4. On reprend l'Exemple 3.1, on a

- Un événement **impossible** : " on tire le nombre 15".
- Un événement **certain** : "on tire un nombre entier positif".

3.2. ENSEMBLE FONDAMENTAL ET ÉVÉNEMENT

Un événement est un élément de $\mathcal{P}(\Omega)$ (l'ensemble des parties de Ω), qui suit la théorie des ensembles. Un lexique spécifique aux événements est présent, distinct du lexique de la théorie des ensembles. Ci-dessus un tableau de correspondance entre les deux vocabulaires.

Notation	Vocabulaire théorie des ensembles	Vocabulaire probabiliste
Ω	ensemble fondamenta	événement certain
\emptyset	ensemble vide	événement impossible
ω	élément de Ω	est dit événement élémentaire
A	sous-ensemble de Ω	événement
$\omega \in A$	ω appartient à A	ω réalise l'événement A
$A \subset B$	A inclus dans B	L'événement A implique l'évènement B
$A \cup B$	réunion de A et B	Un au moins des deux événements est réalisé (A ou B)
$A \cap B$	intersection de A et B	Les deux événement A et B sont tréalisés (A et B)
\bar{A} ou A^c	Complémentaire de A dans Ω	événement contraire de A
$A \cup B = \emptyset$	A et B disjoints	A et B incompatibles

Remarque 3.1. *Les opérations logiques relatives aux événements comme "ou", "et", la "négation" se manifestent par des opérations ensemblistes comme suit : **La réunion, l'intersection et la complémentaire** sont respectivement utilisées.*

- $A \cup B$ se réalise ssi A se réalise ou B se réalise : A ou B .
- $A \cap B$ se réalise ssi A se réalise et B se réalise : A et B .
- \bar{A} se réalise $\iff A$ ne se réalise pas.

Les propriétés de ces opérations telles que la commutativité, l'associativité et la distributivité, telles qu'elles sont vues en théorie des ensembles, demeurent valides pour les événements aléatoires.

Le couple $(\Omega, \mathcal{P}(\Omega))$ est dit **un espace probabilisable**. Mais, même si Ω est fini, le nombre d'éléments de $\mathcal{P}(\Omega)$ est $2^{\text{card}(\Omega)}$ peut être un nombre très élevé. Il vaut mieux dans ce cas de ne considérer qu'une famille restreinte \mathcal{A} de parties de Ω ($\mathcal{A} \subset \mathcal{P}(\Omega)$).

Afin que les résultats des opérations telles que intersection, réunion, complémentaire, restent des événements, cette famille d'événements doit vérifier les conditions suivantes.

Propriétés 3.1. .

1. $\Omega \in \mathcal{A}$;
2. $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$;
3. si $(A_n)_{n \geq 1}$ est une famille dénombrable d'éléments de \mathcal{A} alors, $\bigcup_{n \geq 1} A_n \in \mathcal{A}$.

3.3. ESPACE DE PROBABILITÉS (LE MODÈLE PROBABILISTE DE KOLMOGOROV)

→ \mathcal{A} est dite alors tribu ou σ -algèbre sur Ω .

→ Le couple (Ω, \mathcal{A}) est appelé **espace probabilisable**.

Exemple 3.5. $\forall \Omega$ non vide, on a $\{\emptyset, \Omega\}$ et $\mathcal{P}(\Omega)$ sont des tribus sur Ω .

3.3 Espace de probabilités (Le modèle probabiliste de Kolmogorov)

Définition 3.3. Soient Ω un ensemble fondamental non vide et \mathcal{A} une tribu sur Ω . On appelle une mesure de probabilité p définie sur (Ω, \mathcal{A}) une application :

$$\begin{aligned} p &: \mathcal{A} \rightarrow [0; 1] \\ A &\mapsto p(A) \end{aligned}$$

telle que

1. $0 \leq p(A) \leq 1, \forall A \in \mathcal{A}$
2. $p(\Omega) = 1$.
3. Pour toute suite $(A_n)_{n \geq 1}$ d'événements incompatibles ($A_n \cap A_m = \emptyset, n \neq m$), on a

$$p\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} p(A_n) \text{ (propriété de } \sigma\text{-additivité).}$$

→ Le triplet (Ω, \mathcal{A}, p) est dit **espace probabilisé** ou **espace de probabilités**.

Remarque 3.2. Dans le cas d'un ensemble Ω fini, l'axiome 3 de cette définition est substitué par

pour toute famille $\{A_1, \dots, A_n\}$ d'événements incompatibles ($A_n \cap A_m = \emptyset, n \neq m$), on a

$$p\left(\bigcup_{i=1}^{i=n} A_i\right) = \sum_{i=1}^{i=n} p(A_i) \text{ (axiome d'additivité).}$$

Remarque 3.3. La probabilité d'un événement E désigne la chance que l'événement E se réalise.

Proposition 3.1. On considère deux événements A et B .

1. $p(\bar{A}) = 1 - p(A)$, (car $\forall A \in \mathcal{A}, A \cup \bar{A} = \Omega$, on a $p(\Omega) = p(A) + p(\bar{A}) = 1$)
2. $p(\emptyset) = 1 - p(\Omega) = 0$

3. $\forall A, B \in \mathcal{A}, A \subset B$, on a $\boxed{p(A) \leq p(B)}$ (Inégalité de Boole).
4. $\forall A, B \in \mathcal{A}$, on a $\boxed{p(A \cup B) = p(A) + p(B) - p(A \cap B)}$
5. $\forall A, B \in \mathcal{A}$, on a $\boxed{p(A \setminus B) = p(A) - p(A \cap B)}$.

Remarque 3.4. . Soit E un événement.

- 1) E est dit un événement certain si et seulement si $p(E) = 1$.
- 2) E est dit un événement impossible si et seulement si $p(E) = 0$.

3.3.1 Cas où Ω est fini

Soit Ω un ensemble fini. Si $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, on note $p_i = p(\{\omega_i\})$, $i = 1, \dots, n$, la probabilité de réalisation de ω_i . On a

$$0 \leq p_i \leq 1 \quad \text{et} \quad \sum_{i=1}^{i=n} p_i = 1.$$

$(\Omega, \mathcal{P}(\Omega), p)$ est un espace probabilisé.

On considère un événement A ($A \subset \Omega$), on a $p(A) = \sum_{\omega_i \in A} p_i$. Ceci indique que la probabilité d'un événement A est déterminée en sommant les probabilités de tous les événements élémentaires qu'il contient. Par exemple $A = \{\omega_1, \omega_3\}$, on a $p(A) = p_1 + p_3$.

L'hypothèse d'équiprobabilité : (La probabilité uniforme)

Dans certaines expériences aléatoire, telle que le jet d'un dé équilibré, les événements élémentaires possède la même probabilité de réalisation. Dans ces situations, on affirme que les événements élémentaires sont équiprobables.

Dans cette situation, notons p la probabilité de chaque événement élémentaire. Ce qui donne $p_1 = p_2 = \dots = p_n = \frac{1}{n} = p$, et par conséquent, la probabilité d'un événement A est définie par

$$p(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{|A|}{|\Omega|} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

. Cette probabilité est dite probabilité uniforme sur $(\Omega, \mathcal{P}(\Omega))$.

Un cas favorable à un événement E est un événement élémentaire qui réalise E .

Il est important de noter que cette formule est valide uniquement lorsque les événements élémentaires sont équiprobable. Dans cette situation, on calcule les probabilités en calculant le

cardinal des ensembles en question.

Notation. On note le cardinal d'un ensemble E par $\text{card}(E) = |E|$.

Exemple 3.6. On lance un dé cubique équilibré à six faces numérotées de 1 à 6, et on note le résultat obtenu.

Il y a six résultats possibles : $\Omega = \{1, 2, 3, 4, 5, 6\}$. Le dé est non truqué (bien équilibré), donc on a l'équiprobabilité.

La mesure de probabilité p est définie par

ω_i	1	2	3	4	5	6
$p_i = p(\{\omega_i\})$	1/6	1/6	1/6	1/6	1/6	1/6

Soit $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω , on obtient l'espace probabilisé $(\Omega, \mathcal{P}(\omega), p)$.

- Soit $A = \{1, 6\}$, $p(A) = \frac{|A|}{|\Omega|} = \frac{2}{6} = \frac{1}{3}$.

- Soit A l'événement : "obtenir une face paire". On a $p(A) = \frac{|A|}{|\Omega|} = \frac{|\{2,4,6\}|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$.

Remarque 3.5. Exprimer une expérience aléatoire sous forme d'un modèle probabiliste impliqué

- La définition de l'ensemble fondamental Ω ;
- la définition de la tribu \mathcal{A} ;
- la définition d'une mesure de probabilité sur Ω .

Le choix du modèle lié à une expérience aléatoire doit être fait d'une manière aussi simple que possible, tout en prenant en considération la nature du problème à résoudre.

3.4 Probabilité conditionnelle et indépendance

3.4.1 Probabilité conditionnelle

On considère un espace probabilisé (Ω, \mathcal{A}, p) et un événement B tel que $p(B) \neq 0$.

Il arrive parfois que la probabilité d'un événement A soit déterminée en tenant compte du fait que l'événement B a déjà eu lieu.

L'ensemble des résultats possibles est donné par B . Ainsi, la réalisation de l'événement B entraîne une modification de la probabilité de réalisation d'un événement élémentaire.

Exemple 3.7. On lance 2 fois de suite une pièce de monnaie bien équilibrée.

On considère les événements :

3.4. PROBABILITÉ CONDITIONNELLE ET INDÉPENDANCE

E_1 : "Le premier lancer donne pile" ;

E_2 : "Les deux lancers donnent pile".

Calculer la probabilité d'avoir 2 piles en tenant compte du résultat du 1er lancer qui a donné pile ?

Définition 3.4. Soit (Ω, \mathcal{A}, p) un espace probabilisé. Soit B un événement tel que $p(B) > 0$. La probabilité conditionnelle d'un événement A sachant que B s'est produit, est définie par

$$p(A/B) = \frac{p(A \cap B)}{p(B)}.$$

Remarque 3.6. .

- On note $p_B(A)$ et on lit la probabilité de A sachant B .

- Si $B = \emptyset$, $p_B(\cdot)$ n'est pas définie.

- Il ya uniquement les événements qui ont une partie commune avec B qui peuvent se produire.

- La probabilité conditionnelle p_B constitue une nouvelle mesure de probabilité sur l'espace probabilisable (Ω, \mathcal{A}) . Elle regroupe toutes les caractéristiques d'une probabilité. Ainsi, on dispose des caractéristiques suivantes :

(1) Pour deux événements E_1 et E_2 , nous avons $p(E_1 \cup E_2/B) = p(E_1/B) + p(E_2/B)$, si $E_1 \cap E_2 = \emptyset$.

(2) $p(\bar{E}/B) = 1 - p(E/B)$, pour un événement E .

Exemple 3.8. On lance un dé cubique à six faces numérotées de 1 à 6.

$\Omega = \{1, 2, 3, 4, 5, 6\}$.

$\mathcal{A} = \mathcal{P}(\Omega)$ ensemble des parties de Ω .

On définit la mesure de probabilité p par

ω_i	1	2	3	4	5	6
$p_i = p(\{\omega_i\})$	1/4	1/4	1/8	1/8	1/8	1/8

On considère l'événement B : "un résultat paire". $B = \{2, 4, 6\}$.

On a $p(B) = \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$.

Si on considère l'événement A : "le résultat < 5 ".

On a $A = \{1, 2, 3, 4\}$, alors $p(A/B) = \frac{p(A \cap B)}{p(B)} = \frac{p(\{2, 4\})}{\frac{1}{2}} = \frac{\frac{1}{4} + \frac{1}{8}}{\frac{1}{2}} = \frac{3}{4}$.

Propriétés 3.2. (Formule des Probabilités composées)

Soient A et B deux événements de Ω tels que $0 < p(A), p(B)$, donc

$$p(A \cap B) = p(B/A) \times p(A) = p(A/B) \times p(B).$$

On peut généraliser cette formule à l'intersection de n événements de la manière suivante : $B_1, B_2, \dots, B_n, n \geq 2$, tels que $0 < p(\cap_{i=1}^n B_i)$, on a

$$p(B_1 \cap B_2 \cap \dots \cap B_n) = p(B_1) \times p(B_2/B_1) \times p(B_3/B_1 \cap B_2) \times \dots \times p(B_n/B_1 \cap \dots \cap B_{n-1}).$$

✓

Exemple 3.9. Supposons qu'il y ait deux urnes U_1 et U_2 , contenant chacune 3 boules noires et 5 boules blanches. Une boule est extraite de la première urne U_1 , on enregistre sa couleur et on la remet dans l'urne U_2 . Par la suite, une boule est extraite de la deuxième urne U_2 . Calculer la probabilité d'obtenir deux fois une boule noire ?

Réponse : Notons

\mathcal{N}_1 : "La boule tirée de U_1 est noire".

\mathcal{N}_2 : " La boule tirée de U_2 est noire".

Il est clair qu'on cherche à calculer $p(\mathcal{N}_1 \cap \mathcal{N}_2)$.

On a $p(\mathcal{N}_1) = \frac{3}{8}$ et $p(\mathcal{N}_2/\mathcal{N}_1) = \frac{4}{9}$. Puisque le premier tirage a donné une boule noire, il y a maintenant 4 boules noires dans l'urne U_2 . Par conséquent,

$$p(\mathcal{N}_1 \cap \mathcal{N}_2) = p(\mathcal{N}_1)p(\mathcal{N}_2/\mathcal{N}_1) = \frac{3}{8} \cdot \frac{4}{9} = \frac{1}{6}.$$

3.4.2 Formule des probabilités totales

Il arrive parfois que la probabilité $p(A)$ ne puisse pas être évaluée directement, alors qu'on peut calculer $p(A/B_i)$ et $p(B_i)$ pour $i = 1, 2, \dots, n$. Où $\{B_1, B_2, \dots, B_n\}$ est une partition de Ω .

Dans la proposition ci-dessous, on donne la réponse à la question : Comment calculer $p(A)$ et $p(B_i/A)$?

Proposition 3.2. (*Formule des probabilités Totales : un cas particulier*)

On considère un espace de probabilité (Ω, \mathcal{A}, p) . Soit B un événement tel que $0 < p(B) < 1$. Pour tout événement A , on a

$$p(A) = p(A|B) \times p(B) + p(A|\bar{B}) \times p(\bar{B}).$$

Proposition 3.3. (*Formule des probabilités totales : la généralisation*)

Soit (Ω, \mathcal{A}, p) un espace de probabilité. Soit $\{B_1, B_2, \dots, B_n\}$ une partition de Ω , telle que

$p(B_i) > 0$, pour tout $i = 1, \dots, n$. Alors, pour tout événement A , on a

$$p(A) = \sum_{i=1}^{i=n} p(A/B_i) \times p(B_i).$$

Démonstration. Comme $A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^{i=n} B_i\right) = \bigcup_{i=1}^{i=n} (A \cap B_i)$, on déduit que

$$p(A) = p\left(\bigcup_{i=1}^{i=n} (A \cap B_i)\right) = \sum_{i=1}^{i=n} p(A/B_i) \times p(B_i).$$

□

On rappelle la définition d'une partition de Ω (voir cours d'Algèbre L1).

Définition 3.5. Une famille d'événements $(B_i)_{i \in I}$ est dite une partition (un système complet d'événements) de l'ensemble Ω si les deux conditions suivantes sont satisfaites :

(i) $\cup_{i \in I} B_i = \Omega$;

(ii) pour tous $i \neq j$, $B_i \cap B_j = \emptyset$ (les événements B_i sont deux à deux incompatibles).

Exemple 3.10. (Un exemple dit Le problème du tricheur)

Dans une population, on admet qu'il y a la proportion $p \in]0, 1[$ de tricheurs. On choisit une individu au hasard dans cette population. Il extrait une carte d'un jeu de 52 cartes. Lorsqu'un individu est un tricheur, il est certain qu'il va tirer un as. Calculer la probabilité que cet individu tire un as ?

Réponse :

On considère les événements :

T : "L'individu choisi est un tricheur" ; A : "L'individu tire un as".

On a $\{T, \bar{T}\}$ est un système complet d'événement, $T \cup \bar{T} = \Omega$ et $T \cap \bar{T} = \emptyset$. En utilisant la formule des probabilités totales, on calcule

$$p(A) = p(T)p(A/T) + p(\bar{T})p(A/\bar{T}) = p \times 1 + (1 - p) \times \frac{4}{52} = \frac{1 + 12p}{13}.$$

3.4.3 Formule de Bayes

Proposition 3.4. (Formule de Bayes)

Soit (Ω, \mathcal{A}, p) un espace de probabilité. Soit $\{B_1, B_2, \dots, B_n\} \subset \mathcal{A}$ est une partition de Ω , vérifiant $p(B_i) > 0$, pour tout $i = 1, \dots, n$. Alors pour tout événement $A \in \mathcal{A}$ tel que $p(A) > 0$,

on a

$$p(B_j/A) = \frac{p(A/B_j) \times p(B_j)}{\sum_{i=1}^n p(A/B_i) \times p(B_i)}, \quad \forall j = 1, \dots, n.$$

Pour deux événements A et B de probabilités non nulles, on a

$$p(B/A) = \frac{p(A/B) \times p(B)}{p(A/B) \times p(B) + p(A/\bar{B}) \times p(\bar{B})}.$$

(Motivation : on s'intéresse à calculer la probabilité de B sachant A en utilisant la probabilité de A sachant B).

Démonstration. D'après la définition de la probabilité conditionnelle, on écrit

$$p(B_j/A) = \frac{p(B_j \cap A)}{p(A)} = \frac{p(A/B_j) \times p(B_j)}{p(A)}.$$

On utilise la formule des probabilités totales pour $p(A)$ pour obtenir le résultat recherché. \square

3.4.4 Indépendance stochastique

Soit (Ω, \mathcal{A}, p) un espace de probabilité. Soient A et B deux événements tels que $p(B) > 0$. Affirmer que la réalisation de B n'affecte pas la réalisation de A s'exprime par $p(A/B) = p(A)$.

Définition 3.6. Soient A et B deux événements d'un espace de probabilisé (Ω, \mathcal{A}, p) , de probabilités non nulles ($0 < p(A), p(B) \leq 1$). On dit que A et B sont (stochastiquement) indépendants si et seulement si

$$p(A/B) = p(A) \quad \text{ou} \quad p(B/A) = p(B)$$

Exemple 3.11. On considère un jeu de 32 cartes. On tire sans remise deux cartes.

A : "La carte tirée au premier tirage est un roi"

et B : "La carte tirée au deuxième tirage est un roi"

Si le tirage est sans remise, ces deux événements ne sont pas indépendants.

Dans un tirage avec remise, les événements A et B sont indépendants.

Proposition 3.5. (Caractérisation de deux événements indépendants)

Deux événements $A, B \in \mathcal{A}$ d'un espace probabilisé (Ω, \mathcal{A}, p) . Ils sont dits indépendants si et seulement si

$$p(A \cap B) = p(A) \times p(B).$$

Exemple 3.12. On considère un jeu de 32 cartes. On tire une carte au hasard. Soient les événements :

F : "obtenir une figure".

C : "obtenir un coeur".

L'événement $F \cap C$: "obtenir une dame de coeur ou un valet de coeur ou un roi de coeur".

Et on a $p(F \cap C) = 3/32$; $p(F) = \frac{12}{32} = \frac{3}{8}$ et $p(C) = \frac{8}{32} = \frac{1}{4}$.

On a l'égalité des probabilités : $p(F \cap C) = p(F) \times p(C)$ ce qui montre que les événements F et C sont indépendants.

Théorème 1. Soient (Ω, \mathcal{A}, p) un espace de probabilité et A et B deux événements. Si A et B sont indépendants, alors 1) A et \bar{B} sont indépendants, 2) \bar{A} et B sont indépendants et 3) les événements \bar{A} et \bar{B} sont indépendants.

Démonstration. Soient A et B deux événements indépendants, on a

$A = (A \cup B) \cup (A \cap \bar{B})$, avec $(A \cup B) \cap (A \cap \bar{B}) = \emptyset$, donc

$$p(A) = P((A \cap B)) + P((A \cap \bar{B}))$$

Or A et B sont indépendants, $p(A) = p(A) \times p(B) + p((A \cap \bar{B}))$

$$p(A \cap \bar{B}) = p(A) - p(A) \times p(B) = p(A)(1 - p(B)) = p(A)p(\bar{B}).$$

Ce qui prouve que A et \bar{B} sont indépendants.

En appliquant ce résultat aux événements \bar{B} et A , on trouve que \bar{A} et \bar{B} sont indépendants.

On fait le même raisonnement pour \bar{A} et B . □

Définition 3.7. Soit (Ω, \mathcal{A}, p) un espace de probabilité. Si on considère n événements A_i , $n > 2$, on a les définitions

- on dit que A_1, A_2, \dots, A_n sont indépendants deux à deux si

$$p(A_i \cap A_j) = p(A_i) \times p(A_j), \quad 1 \leq i \neq j \leq n.$$

- On dit que $A_1, A_2, \dots, A_n \in \mathcal{A}$ sont mutuellement indépendants si pour toute partie I de l'ensemble $\{1, 2, \dots, n\}$ on a

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

Exemple 3.13. Dans le cas de 3 événements A, B et C , ils sont :

- mutuellement indépendants si :

$$p(A \cap B) = p(A)p(B), \quad p(A \cap C) = p(A)p(C), \quad p(B \cap C) = p(B)p(C) \quad \text{et} \quad p(A \cap B \cap C) = p(A)p(B)p(C);$$

- indépendants deux à deux si :

$$p(A \cap B) = p(A)p(B), p(A \cap C) = p(A)p(C), p(B \cap C) = p(B)p(C).$$

3.5 Complément du cours : Éléments d'Analyse combinatoire

L'analyse combinatoire est la discipline des mathématiques qui étudie les différentes manières de disposer des objets. Il s'agit d'un ensemble de méthodes et techniques pour évaluer le nombre de dispositions qu'on peut former à l'aide des éléments d'un ensemble fini.

Ces méthodes et techniques sont utilisées afin de calculer le nombre de toutes les réalisations possibles d'une expérience aléatoire. Ce qui fait que cette branche est indispensable dans le calcul des probabilités.

3.5.1 Les outils

Les tableaux, les listes et les arbres de classement.

Les tableaux :

Question : On lance successivement deux dés à 3 faces, numérotées de 1 à 3. Combien y a-t-il d'issues (résultats possibles) ?

	1	2	3
1	(1,1)	(1,2)	(1,3)
2	(2,1)	(2,2)	(2,3)
3	(3,1)	(3,2)	(3,3)

Réponse : Il y a $3 \times 3 = 9$ résultats possibles.

Dans un tableau il n'est pas possible de mettre plus de deux paramètres.

Les listes :

Question : Combien de mots peut on former avec les lettres H, E,L, et D (sans répétition) ?

HEDL	HELD	LHDE	DLEH
HEDL	EHL D	LHED	DELH
ELHD	HLED	LDEH	DHEL
HLDE	LDHE	ELDH	DHLE
HDEL	EDHL	DLHE	LEDH
HDLE	EDLH	LEHD	DEHL

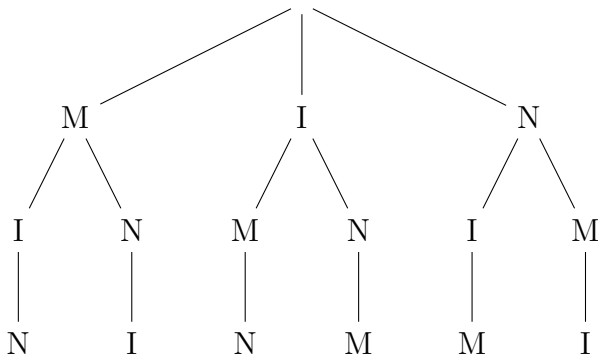
Réponse : On peut former 24 mots différents avec les lettres H , E, L et D.

On note que la liste est très longue et on peut oublier des éléments ou de les représenter plusieurs fois.

L'arbre de classement :

Définition 3.8. *Un arbre de classement est un schéma qui sert à dénombrer tous les résultats possibles d'une expérience aléatoire.*

Question : Combien de mots peut on former avec les lettres M,I et N (sans répétition) ?



Réponse : Le nombre de mots qu on peut écrire avec les lettres M, I et N est : 6.

3.5.2 Principes de base

Le Principe de multiplication : On considère une expérience qui se décompose en m sous expériences. Si la i ème expérience a n_i résultats possibles pour $i = 1, 2, \dots, m$. Le nombre total de résultats possibles de cette expérience est :

$$\prod_{i=1}^{i=m} n_i = n_1 \times n_2 \times \dots \times n_m.$$

Exemple 3.14. *Question :* On veut calculer le nombre de manières possibles de choisir un code de 8 chiffres. *La question :* Combien de possibilités de choisir un code de 8 chiffres ?

Réponse :

Il y a $n_1 = 10$ manières distinctes de choisir le premier chiffre.

Il y a $n_2 = 10$ manières distinctes de choisir le deuxième chiffre.

Il y a $n_3 = 10$ manières distinctes de choisir le troisième chiffre.

Ainsi de suite, il y a $n_8 = 10$ manières distinctes de choisir le 8^{ème} chiffre.

Ce qui donne le nombre total

$$\prod_{i=1}^{i=8} n_i = 10^8.$$

de manières différentes de constituer un code.

Le Principe d'addition

Ce principe se résume comme suit. Dans une expérience, la première opération, notée o_1 , peut être réalisée de n_1 manières différentes. La deuxième opération, notée o_2 , peut être réalisée de n_2 manières différentes, ..., et l'opération o_m peut être réalisée de n_m manières différentes.

On obtient le nombre de façons différentes d'effectuer une de ces opérations est

$$n_1 + n_2 + \dots + n_m.$$

Exemple 3.15. .

Question On s'intéresse au calcul du nombre de menus différents que nous avons avec le choix entre quatre restaurants, le premier possède 13 menus, le second possède 23 menus, le troisième possède 18 menus, et le quatrième possède 32 menus.

Réponse : On peut manger soit au premier restaurant, soit au deuxième, soit au troisième, soit au quatrième.

Le nombre de choix possibles de menus est donc $13 + 23 + 18 + 32 = 86$.

3.5.3 Éléments discernables et éléments indiscernables

Soient e_1, e_2, \dots, e_n les n éléments de l'ensemble Ω , on a $\text{card}(\Omega) = n$.

On dit que les objets e_i et e_j sont indiscernables s'ils sont équivalents.

On dit que les éléments e_i et e_j sont discernables, si ils sont distincts (différents).

Exemple 3.16. Les éléments de Ω peuvent être

1) discernables 2 à 2, $\{1, 2, 4, b, c, d, e, f, g, h, k\}$;

2) tous indiscernables $\{1, 1, 1, 1, 1, 1, 1, 1, 1\}$;

3) discernables ou indiscernables $\{1, 1, 3, a, a, b, b, c, b, g, hd, c, a\}$.

3.5.4 Les différentes dispositions

Soit Ω un ensemble fondamental contenant n éléments. On extrait au hasard k éléments de Ω . On obtient un échantillon de taille k qu'on appellera "une disposition". On s'intéresse au calcul du nombre de manières différentes d'obtenir cette disposition.

Pour obtenir ce nombre, il est nécessaire de spécifier les éléments suivants :

Chaque élément d'une disposition est caractérisé par :

- le nombre d'occurrences d'un élément dans la disposition. Il est possible d'effectuer une disposition avec ou sans répétition en prenant en considération le fait qu'un élément peut être utilisé à plusieurs fois ou une seule fois.
- la place d'un élément dans la disposition. Si la place d'un élément dans une disposition est importante, on qualifie cette disposition d'ordonnée. Sinon, elle est non ordonnée.

Définition 3.9. (*Définitions*)

- 1) Une disposition sans répétition est une disposition où chaque élément peut se présenter soit une fois, soit jamais (1 fois ou 0 fois).
- 2) Dans une Disposition avec répétition, chaque élément peut apparaître plus d'une fois.
- 3) Une disposition ordonnée est une disposition où l'ordre d'obtention d'un élément est significatif. Par exemple dans les situations telles que la composition des chiffres formant un numéro de téléphone ou un code.
- 4) Une disposition non ordonnée est une disposition où l'ordre d'obtention d'un élément n'est pas important. On ne tient pas compte de l'ordre dans la formation de la disposition.

3.5.5 Les arrangements

Soit Ω un ensemble fondamental contenant n éléments. On constitue un échantillon de taille k pris parmi les éléments de Ω .

Définition 3.10. On appelle *arrangement de k éléments parmi n* ($0 \leq k \leq n$) toute disposition ordonnée de k éléments pris parmi les n éléments.

Attention On tient compte de l'ordre des éléments.

Définition 3.11. (*Arrangement avec répétition*)

Si dans un arrangement un élément peut être observé plusieurs fois, l'arrangement est appelé **arrangement avec répétition** de k éléments pris parmi n éléments.

Le nombre d'arrangement avec répétition de k éléments pris parmi n est

$$A_n^k = n^k.$$

Remarque 3.7. Dans le cas d'un arrangement avec répétition, les k éléments de la disposition ne sont pas nécessairement tous discernables. Ce qui correspond à un **tirage avec remise** ou l'**ordre** est important.

Exemple 3.17. .

• Question 1 : Combien de mots de 7 lettres peut on constituer avec un alphabet à 26 lettres ?

Réponse : $26 \times 26 \times 26 \times 26 \times 26 \times 26 \times 26 = 26^7$.

• Question 2 : On jette deux fois de suite un dé à 12 faces. Quel est le nombre résultats possibles en tenant compte de l'ordre ?

Réponse : $12 \times 12 = 12^2 = 144$.

Définition 3.12. (Un arrangement sans répétition).

Un arrangement est dit sans répétition si chaque élément n'est observé qu'une seule fois dans l'arrangement.

Le nombre d'arrangements sans répétition de k élément pris parmi n est

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

Remarque 3.8. Dans un arrangement sans répétition, les k objets de la disposition sont tous discernables. Dans ce cas, on dit que le **tirage est sans remise** et avec **ordre**.

On a $0 \leq k \leq n$ et $n \in \mathbb{N}^*$

Exemple 3.18. .

Question : De combien de façons peut on disposer 4 personnes dans un salon de 15 places ?

Réponse : $\frac{15!}{(15-4)!} =$.

3.5.6 Permutations

Soit Ω un ensemble fondamental contenant n éléments. On extrait un échantillon de taille n parmi les éléments de Ω .

Définition 3.13. (Permutation sans répétition)

On appelle permutation sans répétition de n éléments discernables, un arrangement de n éléments choisis parmi n éléments.

Le nombre de permutations de n éléments discernables est

$$P_n = n!.$$

Exemple 3.19. .

Question 1 : De combien de manière peut on disposer 10 étudiants dans une salle de 10 places ?

Réponse : 10!.

Question 2 : De combien de manières différentes peut on placer 4 personnes l'une à côté de l'autre ?

Réponse : Il y a 4 choix pour la 1ère place, 3 choix pour la 2ème place, puis 2 choix, et 1 choix.

Par conséquent, il y a $4 \times 3 \times 2 \times 1 = 24$ manières différentes de placer ces 4 personnes.

Définition 3.14. (Permutation de n objets, dont k seulement sont distincts (discernables))

C'est une disposition ordonnée de n éléments choisis dans k classes d'éléments distinctes. On note le nombre d'éléments dans la classe i par n_i , $1 \leq i \leq k$, avec $n_1 + n_2 + \dots + n_k = n$.

Le nombre de permutation avec répétition est

$$P_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! \times n_2! \times \dots \times n_k!}.$$

Exemple 3.20. .

Question : Quel est le nombre de mots différents qu'on peut former avec les lettres du mot "STATISTIQUES" ?

Réponse : Il y a $n = 12$ lettres et 7 groupes distincts $k = 7$: $n_1 = 3$ (trois S), $n_2 = 3$ (trois T), $n_3 = 2$ (deux I), $n_4 = 1$ (un E), $n_5 = 1$ (un A), $n_6 = 1$ (un U), $n_7 = 1$ (un Q). On obtient le nombre de mots distincts qui est égal au nombre de permutations avec répétition :

$$\frac{12!}{3!3!2!1!1!1!1!}.$$

3.5.7 Combinaisons

Soit Ω un ensemble fondamental contenant n éléments. On forme un échantillon de taille k pris parmi les n éléments de Ω .

Définition 3.15. (Combinaison sans répétition (sans remise-simultanément)).

Une combinaison sans répétition de k éléments parmi n ($1 \leq k \leq n$) est une disposition non ordonnée et sans répétition, de k éléments choisis parmi n éléments discernables. Le tirage des k éléments est fait simultanément, donc sans remise. La notion d'ordre n'a pas d'importance et une combinaison sans répétition est un ensemble de k éléments discernables.

Le nombre de combinaisons sans répétition de k éléments choisi parmi n est donné par

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

Remarque 3.9. .

- Une combinaison sans répétition correspond à un tirage sans remise et sans tenir compte de l'ordre.

- Le nombre de combinaisons sans répétition de k éléments pris parmi n est égale au nombre de combinaisons sans répétition de $n - k$ éléments pris parmi n .

Exemple 3.21. .

• Question 1 : De combien de façons peut-on choisir un groupe de 35 étudiants parmi un groupe de 120 étudiants ?

Réponse : $C_{120}^{35} = \frac{120!}{35!(120-35)!}$.

• Question 2 : A un examen, un candidat doit répondre à 3 questions sur 5. Calculer le nombre de choix possibles pour le candidat ?

Réponse : $C_5^3 = \frac{5!}{2!(5-3)!}$.

• Question 3 : A un examen, on considère qu'un étudiant doit répondre à 3 questions sur 5. Sous l'hypothèse que la première question est imposée, calculer le nombre de choix possibles pour le candidat ?

Réponse : C_4^2 . Car le candidat est obligé de répondre à la première question. Parmi les 4 question, il ne lui reste que 2 questions à choisir.

Définition 3.16. (Combinaison avec répétition (avec remise)).

Une combinaison avec répétition de k éléments parmi n est une disposition, non ordonnée, de k éléments choisis parmi n éléments distincts et qui peuvent se répéter.

Le nombre de combinaisons avec répétition de k éléments parmi n est

$$K_n^k = C_{(n+k-1)}^k = \frac{(k+n-1)!}{k!(n-1)!}.$$

Remarque 3.10. Une combinaison avec répétitions correspond au cas d'un tirage sans ordre et avec remise.

3.6 Exercices

3.6.1 Exercices sur l'analyse combinatoire (Série 3)

Exercice 3.1. .

Un numéro de téléphone commence par le chiffre 0 suivi de 8 chiffres.

- 1- *Combien de numéros de téléphone peut on former ?*
- 2- *Combien de ces numéros de téléphone ne comportant pas le chiffre 2 ?*
- 3- *Combien de numéros de téléphone qui se terminent par 23 ?*

Exercice 3.2. .

- 1- *De combien de façons peut on aligner 5 livres de statistique, 2 livres d'algèbre et 4 livres d'analyse sur une étagère ?*
- 2- *De combien de façons différentes peut-on placer l'une à côté de l'autre, 10 boules marrons, 3 vertes, 5 blanches et 2 bleues ?*

Exercice 3.3. .

- 1) *Combien de nombres de 4 chiffres différents peut on former à l'aide des chiffres $\{2, 3, 4, 5, 6, 7\}$?*
- 2) *Combien de ces nombres sont multiples de 5 ?*
- 3) *Combien de ces nombres sont plus petit que 5000 ?*

Exercice 3.4. .

Combien de mots de 11 lettres peut-on former avec les 26 lettres de l'alphabet si

- a) *on utilise chaque lettre une seule fois ?*
- b) *on peut réutiliser les lettres ?*

Exercice 3.5. .

De combien de façons différentes peut-on répartir 5 hommes et leurs femmes sur une rangée de 10 chaises dans les cas suivants :

- a) *Sans aucune condition*
- b) *Si chaque homme doit être assis à coté de sa femme ?*

Exercice 3.6. .

Cinq garçons et quatre filles s'assoient sur un banc.

1. *Quel est le nombre de dispositions possibles ?*
2. *Donner le nombre de dispositions possibles si les garçons sont d'un côté et les filles de l'autre.*

3.6. EXERCICES

3. Donner le nombre de dispositions possibles si chaque fille est intercalée entre deux garçons.
4. Traiter les mêmes questions s'ils s'assoient autour d'une table ronde.

Exercice 3.7. On forme un groupe de 7 personnes choisies parmi 31 étudiantes et 27 étudiants.

- 1) Quel est le nombre de façons de constituer ce groupe ?
- 2) Quel est le nombre de façons de constituer ce groupe avec uniquement des étudiantes ?
- 3) Quel est le nombre de façons de constituer ce groupe avec uniquement des personnes du même sexe.
- 4) Quel est le nombre de façons de constituer ce groupe avec au moins un étudiant et au moins une étudiante.

Exercice 3.8. .

Montrer la formule du triangle de Pascal : $n \geq k$, $C_n^k = C_{n-1}^k + C_{n-1}^{k-1}$.

3.6.2 Corrigés des exercices

Solution de l'exercice 3.1. 1) C'est le nombre d'arrangements avec répétition : $\mathcal{A}_{10}^8 = 10^8$.

2) $\mathcal{A}_9^8 = 9^8$.

3) $\mathcal{A}_{10}^6 = 10^6$.

Solution de l'exercice 3.2. .

1-Le nombre de façons différentes d'aligner ces 11 livres est le nombre de permutations sans répétition : $P_{11} = 11!$.

2- Le nombre de façons différentes de placer ces 20 boules est le nombre de permutation avec répétition : $P_{20}(10, 3, 5, 2) = \frac{20!}{10! \times 3! \times 5! \times 2!}$.

Solution de l'exercice 3.3. .

1) C'est le nombre d'Arrangements sans répétition : $A_6^4 = \frac{6!}{2!}$

2) $A_5^3 = \frac{5!}{2!}$

3) $A_3^1 \times A_5^3 = \frac{3!}{2!} \times \frac{5!}{2!}$. On peut choisir le premier chiffre à gauche de A_3^1 façons différentes et les trois autres chiffres sont pris parmi les 5 restant.

Solution de l'exercice 3.4. .

a) C'est le nombre possible d'arrangements sans répétition que l'on peut faire avec un échantillon de 11 éléments parmi 26 : $A_{26}^{11} = 26!/15!$.

b) A chaque position (1 à 11), on peut mettre 26 lettres différentes donc le nombre sera : le nombre d'arrangements avec répétition : $\mathcal{A}_2 26^{11} = 26^{11}$.

Solution de l'exercice 3.5. .

a) Aucune condition : $10!$

b) $5!2!2!2!2!$. Un couple est considéré comme une seule personne.

Solution de l'exercice 3.6. .

1) $(5 + 4) = 9!$

2) $(5!4!) \times 2$

3) $5! \times 4!$ (GFGFGFGFG)

Si la table est ronde on a toutes les place sont identiques :

1) Une personne est fixée et on place les autres : C'est le nombre de permutations sans répétition $(9 - 1)!$

2) $5!4!$

3) Une fille intercalée entre deux garçons ça donne (GFGFGFGFG). C'est à dire deux garçons sont cote à cote. On les considère comme une seule personne. On obtient en permutant les

garçons entre eux (4!), en permutant les filles entre elle (4!), et en permutant les deux garçons entre eux : $4!4!2!$. Comme c'est une table ronde, la réponse est : $\frac{4!4!2!}{8}$

Solution de l'exercice 3.7. .

1) C_{31+27}^7 , nombre de combinaison de 7 choisis parmi 58.

2) C_{27}^7

3) $C_{27}^7 + C_{31}^7$

4) $C_{58}^7 - C_{27}^7 - C_{31}^7$

Solution de l'exercice 3.8.

$$\begin{aligned} C_{n-1}^k + C_{n-1}^{k-1} &= \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-1-k+1)!} = \\ &= \frac{(n-1)!(n-k)}{k!(n-k)!} + \frac{(n-1)!k}{k!(n-1-k+1)!} \\ &= \frac{(n-1)!}{k!} \left(\frac{(n-k)+k}{(n-k)!} \right) = \frac{n!}{k!(n-k)!} = C_n^k \end{aligned}$$

3.6.3 Exercices de Probabilités (Série 4)

Exercice 3.9. Soient A , B et C trois événements d'un espace de probabilité (Ω, \mathcal{A}, p) . Exprimer en fonction de A , B et de C les événements suivants :

1. Aucun événement ne se réalise.
2. Au moins un événement se réalise.
3. Au moins deux événements se réalisent
4. Au plus deux événements se réalisent.
5. Au plus un événement se réalise.
6. Exactement deux événement se réalisent.
7. Les trois événements se réalisent.

Exercice 3.10. Soient A et B deux événements tels que $P(A) = 1/5$ et $P(A \cup B) = 1/2$.

1. Sous l'hypothèse que A et B soient incompatibles. Calculer $P(B)$.
2. Sous l'hypothèse que A et B soient indépendants. Calculer $P(B)$.
3. Calculer $P(B)$ avec l'hypothèse que l'événement A ne peut être réalisé que si l'événement B est réalisé.

Exercice 3.11. ;

I) On jette une pièce de monnaie, bien équilibrée, trois fois de suite.

1) Donner la liste de tous les résultats possibles (i.e. l'ensemble fondamental Ω) en notant P pour pile et F pour face.

2) Préciser l'espace de probabilité.

3) Donner la probabilité des événements suivants.

A : "Le tirage ne comporte que des faces".

B : "Le tirage comporte au moins une fois pile.

II) Reprendre les mêmes questions, si on lance trois pièces de monnaies identiques en même temps.

Exercice 3.12. Un sac contient 5 jetons verts et 4 jetons rouges.

1) On tire successivement et au hasard 3 jetons du sac, sans remettre le jeton tiré. Calculer les probabilités :

a) De ne tirer que 3 jetons verts ;

b) De ne tirer aucun jeton vert ;

c) De tirer au plus 2 jetons verts ;

d) De tirer exactement 1 jeton vert.

2) On tire simultanément et avec remise 3 jetons du sac. Reprendre alors les questions a), b), c) et d).

Exercice 3.13. *Un tireur possède deux tentatives pour atteindre une cible. On suppose qu'il atteint sa cible la première fois dans 75% des cas. Lorsque il échoue dans le premier tir, il réussit le second tir dans 90% des cas.*

Quelle est la probabilité pour qu'il n'atteint pas la cible deux fois de suite ?

Exercice 3.14. *Parmi les étudiants d'un département de mathématiques 40% suivent l'option A_1 , 30% suivent l'option A_2 et 30% suivent l'option A_3 . Chaque étudiant suive une seule option. La proportion d'étudiants qui n'ont pas la moyenne dans l'option A_1 est de 10%, dans l'option A_2 de 5% et dans l'option A_3 de 5%. On choisit un étudiant au hasard.*

1. Calculer la probabilité qu'il n'ait pas la moyenne.

2. Sachant qu'il n'a pas la moyenne, calculer la probabilité a posteriori qu'il ait suivi l'option A_1 , A_2 ou A_3 .

Exercice 3.15. *Dans une population en cas de migraine trois patients sur cinq optent pour le médicament M_1 et deux malades sur cinq choisissent un médicament M qui comportent des effets secondaires :*

Avec le médicament M_1 , on observe 80% des malades qui sont soulagés.

Avec le médicament M , on observe 95% des malades qui sont soulagés.

1. Calculer le taux global de personnes soulagées ?

2. Calculer la probabilité pour un patient d'avoir pris le médicament M_1 sachant qu'il est soulagé ?

Exercice 3.16. *On s'intéresse à l'étude d'une mutation génétique sur des nourrissons. On effectue un dépistage. Il existe 3% des bébés qui possèdent cette mutation. On commence le dépistage par un test qui donne 98% de résultats qui sont positifs pour les bébés ayant la mutation et 3% de résultats qui sont positifs pour les bébés n'ayant pas cette mutation.*

1. Déterminer la probabilité qu'un bébé pris au hasard soit atteint de cette mutation sachant que le test a donné un résultat positif ?

2. Déterminer la probabilité qu'un bébé pris au hasard soit indemne de cette mutation sachant que le test a donné un résultat négatif ?

3.6.4 Corrigés des exercices de probabilités

Solution de l'exercice 3.9. .

1. $\bar{A} \cap \bar{B} \cap \bar{C}$.
2. $A \cup B \cup C$.
3. $(A \cap B \cap C) \cup (\bar{A} \cap B \cap C) \cup (A \cap \bar{B} \cap C) \cup (A \cap B \cap \bar{C})$.
4. $\overline{A \cup B \cup C}$.
5. $(\bar{A} \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C) \cup (A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C})$.
6. $(\bar{A} \cap B \cap C) \cup (A \cap \bar{B} \cap C) \cup (A \cap B \cap \bar{C})$.
7. $A \cap B \cap C$.

Solution de l'exercice 3.10. .

1. A et B incompatibles donc $A \cap B = \emptyset$ d'où $P(A \cup B) = P(A) + P(B) \implies P(B) = 1/2 - 1/5$.
2. A et B indépendants : $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = P(A) + P(B)(1 - P(A))$, d'où $P(B) = \frac{P(A \cup B) - P(A)}{1 - P(A)} = \frac{1/2 - 1/5}{1 - 1/5} = 3/8$.
3. A ne peut être réalisé que si B est réalisé : tous les événements de A sont dans B , $P(A \cap B) = P(A)$, donc $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A) = P(B)$, donc $P(B) = P(A \cup B) = 1/2$.

Solution de l'exercice 3.11. .

1) On lance une pièce de monnaie bien équilibrée, trois fois de suite.

1) L'ensemble fondamental est constitué de triplet ordonnés dont les éléments symbolisent pile (P) et face (F).

$$\Omega = \{(P, P, P), (P, P, F), (P, F, F), (F, F, F), (F, P, P), (F, F, P), (F, P, F), (P, F, P)\}.$$

Il ya $2 \times 2 \times 2 = 8$ resultats possibles.

2) L'espace de probabilité $(\Omega, \mathcal{P}(\Omega), p)$, où $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω et p est la mesure de probabilité donnée par $p(\omega) = \frac{1}{8}$, $\forall \omega \in \Omega$ (car les événements élémentaires sont équiprobables).

- $A = \{(F, F, F)\}$, donc $p(A) = p(\{(F, F, F)\}) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{1}{8}$.

3.6. EXERCICES

• $B = \{(F, F, P), (F, P, F), (P, F, F), (F, P, P), (P, F, P), (P, P, F), (P, P, P)\}$, donc

$$p(B) = p(\{(F, F, P), (F, P, F), (P, F, F), (F, P, P), (P, F, P), (P, P, F), (P, P, P)\}) = \frac{\text{card}(B)}{\text{card}(\Omega)} = \frac{7}{8}.$$

On peut aussi remarquer que $B = \bar{A}$, donc $p(B) = 1 - p(A) = 1 - \frac{1}{8} = \frac{7}{8}$.

II) On lance trois pièces de monnaies identiques et bien équilibrées en même temps.

1) L'ensemble fondamental est constitué de triplet non ordonnés dont les éléments symbolisent pile (P) et face (F).

L'ensemble des résultats possibles : (ensemble fondamental) :

$$\Omega = \{\{P, P, P\}, \{P, P, F\}, \{P, F, F\}, \{F, F, F\}\}.$$

$\text{card}(\Omega) = 4$, il ya 4 résultats possibles.

2) L'espace de probabilité $(\Omega, \mathcal{P}(\Omega), p)$, où $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω et p est la mesure de probabilité donnée par $p(\omega) = \frac{1}{4}$, $\forall \omega \in \Omega$ (car les événements élémentaires sont équiprobables).

3) Les événements élémentaires sont équiprobables, donc

• $p(A) = p(\{\{F, F, F\}\}) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{1}{4}$, $A = \{(F, F, F)\}$.

• $B = \{\{F, F, P\}, \{F, P, P\}, \{P, P, P\}\}$, donc

$$p(B) = p(\{\{P, P, P\}, \{P, P, F\}, \{P, F, F\}\}) = \frac{\text{card}(B)}{\text{card}(\Omega)} = \frac{3}{4}.$$

Solution de l'exercice 3.12. .

1) Tirages successifs sans remise de 3 jetons parmi 9 : un tirage successif et sans remise de 3 jetons parmi 9 est un arrangement sans répétition de 3 éléments parmi 9. Il ya $A_9^3 = 9 \times 8 \times 7 = 504$.

a) Notons A l'événement : "tirer 3 jetons verts". On a $p(A) = \frac{A_5^3}{A_9^3} = \frac{5}{42}$.

b) Notons B l'événement "ne tirer aucun jetons vert" . On a $p(B) = \frac{A_4^3}{A_9^3} = 1/21$.

c) Notons C l'événement "tirer au plus 2 jetons vert". On a deux méthodes.

3.6. EXERCICES

Méthode 1 : On remarque que $C = \bar{A}$. Donc, $p(C) = 1 - p(A) = 1 - \frac{5}{42} = \frac{37}{42}$.

Méthode 2 : On note les événements E_1 : "ne tirer aucun jeton vert"; E_2 : "Tirer un jeton vert et deux jetons rouges"; E_3 : "tirer exactement deux jetons vert".

$$p(C) = \frac{\overbrace{A_4^3}^{\text{Ne tirer aucun vert}} + \overbrace{3 \times A_5^1 \times A_4^2}^{\text{Choix de la place du jeton vert-Choix d'un vert et de deux rouges.}} + \overbrace{3 \times A_5^2 \times A_4^1}^{\text{Tirer exactement deux verts}}}{A_9^3}$$

$$= \frac{37}{42}.$$

Le tirage est successif, donc on induit un ordre. Dans le tirage exactement d'un jeton vert, il faut multiplier par 3, c'est à dire choisir une place pour le jeton vert.

Dans l'événement E_3 , il faut choisir une place pour le jeton rouge, d'où la multiplication par 3 (le nombre de places) dans le calcul de la probabilité.

d) Soit D l'événement "tirer exactement 1 jeton vert". $p(D) = \frac{3 \times A_5^1 \times A_4^2}{A_9^3} = \frac{5}{14}$.

2) Tirages successif et avec remise de 3 jetons parmi 9. : on compte le nombre d'arrangements avec répétitions. $\text{card}(\Omega) = A_9^3 = 9^3$.

a) $p(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{A_5^3}{9^3} = \frac{5^3}{9^3}$.

b) $p(B) = \frac{\text{card}(B)}{\text{card}(\Omega)} = \frac{A_4^3}{9^3} = \frac{4^3}{9^3}$.

c) $p(C) = 1 - p(A)$.

d) $p(D) = \frac{\text{card}(D)}{\text{card}(\omega)} = \frac{3 \times A_5^1 \times A_4^1}{9^3}$.

Solution de l'exercice 3.13. .

Il réussit le premier tir dans 75% des cas et il échoue dans 25% des cas.

Il réussit le second tir dans 90% des cas et il échoue dans 10% des cas.

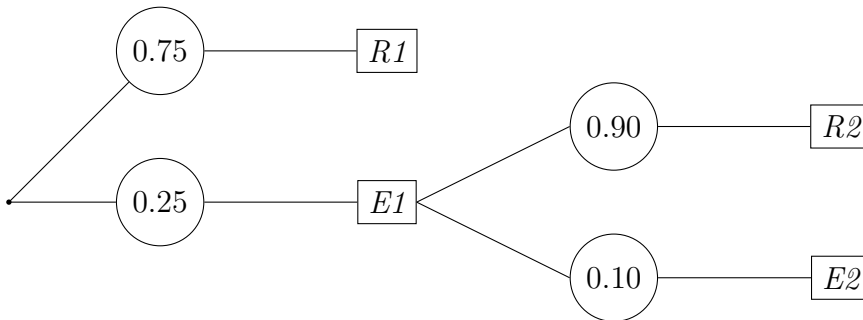
On note les événements ($i = 1, 2$) :

E_i : le tireur échoue au i ème tir.

R_i : le tireur réussit le i ème tir

La probabilité que le tireur échoue deux fois est donnée par $p(E_2 \cap E_1) = p(E_1) \times p(E_2/E_1) = \frac{25}{100} \times \frac{10}{100}$.

L'arbre pondéré



Solution de l'exercice 3.14. Pour $i \in \{1, 2, 3\}$, on note les événements

B_i : "Un étudiant suit l'option A_i ";

C : " Un étudiant n'a pas obtenu la moyenne".

On a $P(A_1) = 0.4$; $P(A_2) = 0.3$; $P(A_3) = 0.3$; $P(C|A_1) = 0,1$; $P(C|A_2) = 0.5$; $P(C|A_3) = 0.5$.

1. $P(C) = P(C|A_1) \times P(A_1) + P(C|A_2) \times P(A_2) + P(C|A_3) \times P(A_3) = 0.1 \times 0.3 + 0.5 \times 0.3 = 0.18$.

2. $P(A_i|C) = \frac{P(C|A_i) \times P(A_i)}{P(C)}$, $i = 1, 2, 3$. On obtient,
 $P(A_1|C) = \frac{P(C|A_1) \times P(A_1)}{P(C)} = \frac{0.1 \times 0.3}{0.18} = 0.16$.

Solution de l'exercice 3.15. On note les événements suivants :

S : "Une personne est soulagée"

M_1 : "une personne ayant pris le médicament M_1 ".

M : "Une personne ayant pris le médicament M "

1. Le taux de personnes soulagées :

On a $p(M) = 2/5$; $p(M_1) = 3/5$.

On a $S = S \cap \Omega = S \cap (M_1 \cup M)$ ($M_1 \cap M = \emptyset$ et $M_1 \cup M = \Omega$). De la formule des probabilités totales :

$$p(S) = p(S \cap (M_1 \cup M)) = p((S \cap M_1) \cup (S \cap M)) = p(S \cap M_1) + p(S \cap M) = p(S/M_1)p(M_1) + p(S/M)p(M)$$

$$p(S) = \frac{3}{5} \times 0.8 + \frac{2}{5} \times 0.95 = 0.86.$$

2. La probabilité pour un patient d'avoir pris M_1 sachant qu'il est soulagé :

$$p(M_1/S) = \frac{p(M_1 \cap S)}{p(S)} = \frac{p(M_1) \times p(S/M_1)}{P(S)} = \frac{\frac{3}{5} \times 0.80}{0.86} = 0.41$$

Bibliographie

- [1] J. P. Lecoutre, Statistique et probabilités, Dunod, Paris, 5eme Edition, 2012.
- [2] G. Calot, Cours de statistique descriptive, Dunod, Paris, 1973.
- [3] P. Bailly, Exercices corrigés de statistique descriptive, OPUs Alger, 1993.
- [4] H. Hamdani, Statistique descriptive avec initiation aux méthodes d'analyse de l'information économique : exercices et corrigés, OPU Alger, 2006.
- [5] K. Redjda, Probabilités, OPU Alger, 2004.
- [6] C. Leboeuf, Jean Louis Roque et Jean Guegan, Cours de probabilités et de statistiques 2ème édition, ellipses (1996).
- [7] M. Nemiche. Exercices corrigés de statistiques et probabilités, 2015.
<https://www.uv.es/nemiche/cursos/polycopies/NemicheStu3.pdf>
- [8] A. Perrut, cours de probabilités et statistiques.
https://math.univ-lyon1.fr/irem/IMG/pdf/PolyTunis_A_Perrut.pdf
- [9] A. Chekroun. Statistiques descriptives et exercices.
<http://math.univ-lyon1.fr/chekroun/Files/chekrounstatistiques.pdf>
- [10] S. Picchione. Combinatoire et probabilités 4eme année. 2012-2013.
<file:///C:/Users/Personal/Downloads/dokumen.tipscombinatoires-et-probabilites.pdf>
- [11] Zarrouk Fayçal, ISSEP Ksar-Said. Cours de statistiques.
- [12] J. P. Javet. Combinatoire et Probabilités. <http://www.gymomath.ch/javmath/polycopie/3Ms>
- [13] R. Rakotomalala. Probabilités et Statistique. <https://eric.univ-lyon2.fr/rizzo/cours/cours/ProbabilitesetStatistique.pdf>.
- [14] <http://www.issep-ks.rnu.tn/fileadmin/templates/Fcad/lesstatistiquesdescriptives.pdf>
- [15] <https://mathsv.univ-lyon1.fr/app/cours/?theme=probachap=5/>
- [16] <https://www.supagro.fr/cnam-lr/statnet/mod1/mod1lx.htm>.