
République Algérienne Démocratique et Populaire Ministère de L'enseignement Supérieur et
de la Recherche Scientifique



Polycopié de cours. Introduction aux statistique descriptive et calcul des probabilités



présentée Par : **Abdesselam Nawel**

Courriels : nawelabedess@gmail.com / n.abdesselam@lagh-univ.dz

2018 - 2019

AVANT PROPOS

Le polycopié a pour but d'initier les étudiants aux principes de base de la statistique et les calculs des probabilités. Le cours vise principalement à introduire et faire méditer les concepts fondamentaux et méthodes élémentaires de la statistique pour permettre un apprentissage autonome ultérieur des méthodes complémentaires. On veut développer le sens critique nécessaire lors de la mise en œuvre et de l'interprétation d'un traitement statistique. Pour cela, on introduira et utilisera un cadre mathématique rigoureux. Nous fournirons autant des exemples et de figures nécessaires afin d'obtenir une meilleure compréhension du cours. La statistique descriptive a pour but d'étudier un phénomène à partir des données. Cette description se fait à travers la présentation des données (la plus synthétique possible), leur représentation graphique et le calcul de résumés numériques.

La place de ce cours dans le futur métier des étudiants :

- Doter les bacheliers d'outils nécessaires pour traiter statistiquement et le plus rigoureusement possible des données relatives à la gestion.
- Fournir à tous qui sont appelés à dialoguer avec les spécialistes de la statistique les connaissances nécessaires à une bonne compréhension et à une utilisation efficace des différentes méthodes.
- Analyse des données (outils scientifiques permettant de résumer un ensemble de données afin de mettre en évidence l'information).
- Simulations (processus stochastique - variable temporelle).
- Prédiction et décisions (probabilités de risque ou d'occurrence).



Table des matières

AVANT PROPOS	2
Symboles et Notations	6
Introduction générale	7
 Chapitre 1. Notions de base et vocabulaire statistique	 11
1. Vocabulaire	11
2. Épreuve statistique	11
3. Concepts de base de la statistique notions de base	12
4. Population	12
5. Variable (ou caractère)	13
6. Observations	15
7. Données	15
8. Étude d'une variable statistique discrète	16
8.1. Effectif partiel - effectif cumulé	17
9. Effectif partiel (fréquence absolue)	17
10. Effectif cumulé	17
11. Fréquence partielle (fréquence relative)	18
12. Fréquence cumulée	19
13. Tableaux Statistiques	20
14. Règle de Sturge et Yule	22
15. Distribution et fonction de répartition	23
16. Fonction de répartition	23
17. Courbe cumulatives croissantes et décroissante	25
18. Introduction	26
19. Cas de variable qualitative	26
20. Représentation circulaire par des secteurs	26
21. Représentation en tuyaux d'orgue	29
22. Diagramme en bandes	29
23. Distribution à caractère quantitatif	29
24. Diagramme en bâtons	30

25.	Histogramme	32
26.	Polygone	33
27.	Étude d'une variable statistique continue	34
28.	Caractère continu	34
28.1.	Classe de valeurs	34
28.2.	Nombre de classes	35
28.3.	Effectif et fréquence d'une classe	35
29.	Travaux Dirigés	36
Chapitre 2. Représentation numérique des données.		43
	Introduction	43
1.	Médiane	43
1.1.	Détermination de la médiane dans le cas d'une variable discrète :	43
2.	Détermination de la médiane dans le cas d'une variable continue.	45
3.	Quartiles	48
4.	Calcul des quartiles dans le cas d'une variable discrète	48
5.	Détermination des quartiles dans le cas d'une variable continue	49
6.	Intervalle inter-quartile.	49
7.	Mode	51
8.	Détermination de le mode dans le cas d'une variable discrète et continue :	51
9.	Détermination de le mode dans le cas d'une variable continue :	53
10.	Moyenne géométrique	54
11.	Moyenne harmonique	54
12.	Moyenne quadratique	55
13.	Étendue	56
14.	Diagramme en boîte ou boîte à moustache	56
15.	Variance	57
16.	Écart type	59
17.	Écart absolu moyen.	60
18.	Coefficient de variation	60
19.	Comparaison des différents paramètres de position :	61
20.	Travaux Dirigés	62
Chapitre 3. Analyse Combinatoire		75
1.	Arrangements	75
2.	Permutations	76
3.	Combinaisons	77

4. Probabilités	79
5. Espace probabilisable	81
6. Événements élémentaires et composés	82
7. Réalisation d'un événement	82
8. Événement incompatibles (disjoints)	82
9. Système complet d'événements	82
10. Algèbre des événements	82
11. Calculs des Probabilités	83
12. Espace probabilisable	83
13. Lien entre la théorie des probabilités et des ensembles	89
14. Construction d'une probabilité et espace probabilisés discrets	89
15. Probabilité conditionnelles, indépendance et probabilité composées	90
16. LES ARBRES DE PROBABILITÉS	92
17. Lois des arbres.	92
18. Travaux Dirigés	92
Bibliographie	103

Symboles et Notations

$\mathbb{N}$ Ensemble des nombres entiers naturels.

$\mathbb{Z}$ Ensemble des nombres entiers relatifs.

$\mathbb{R}$ Ensemble des nombres réels.

$\mathbb{R}^2$ Ensemble des couples de nombres réels.

$Card(\quad)$ Le cardinal.

$:=$ Est défini comme étant (symbole d'affectation).

$\cup$ L'union.

$\cap$ L'intersection.

$|\cdot|$ La valeur absolue.

$\sum_{i=1}^n$ La somme pour i variant de 1 à n .

$v.s$ La variable statistique.

x_i La valeur statistique.

n_i L'effectif.

c_i, c_{i+1} La borne supérieure, la borne inférieure.

C_i Le centre.

a_i L'amplitude.

F_x La fonction de répartition du caractère x .

Me La médiane.

Q_1, Q_2 Les quartiles.

M_0 Le mode.

$\bar{X}$ La moyenne d'une série statistique x .

G La moyenne géométrique.

H La moyenne harmonique.

M_Q La moyenne quadratique.

σ_X L'écart-type de x .

$Var(x)$ La variance de x .

C_v Le coefficient de variation.

e L'écart absolu moyen.

c_i Le centre de la classe.

Q_n^p Arrangements.

Introduction générale

L'histoire de la statistique remonte à une époque très ancienne. Les activités statistiques (dénombrements) ont commencé bien avant la création du mot, l'application de la méthode et de l'analyse statistique. Dans l'autre part, la théorie des probabilités a été originellement motivée par l'analyse des jeux de hasard, mais elle a pris aujourd'hui une place centrale dans la plupart des sciences.

Depuis l'antiquité, les Empereurs, les Rois et les Hommes d'église réalisaient des dénombrements de populations humaines et de terres pour les besoins de la guerre et de l'impôt. Il y a plus de 4 ou 5000 ans, il existait déjà en Chine des descriptions chiffrées de la population et de l'agriculture. Les Égyptiennes de l'époque des Pharaons procédaient au dénombrement de la population ou du bétail.

A Rome, l'empereur Auguste fit procéder à une vaste enquête en dénombrant les soldats, les navires et les revenus publics. Jusqu'au moyen âge, les seules **statistiques** existante étaient les dénombrements faits dans des buts divers : assiettes de l'impôt, répartition des terres, recrutement dans l'armée est effectués avec des méthodes diverses (recensements des personnes, enregistrements de certains actes d'état civil ...).

C'est à partir du XVIII^e siècle, qu'apparaît le mot statistique crée par Achenwall en 1749 à partir du mot STATISTA (politique). Gottfried Achenwall (1719-1772), statisticien et politologue allemand. Il utilisa pour la première fois le terme Statistik dans son (*Staatsverfassung der heutigen vornehmsten europäischen Reiche und Völker im Grundrisse*) c'est à dire [la constitution politique des principaux pays et peuples européens actuels]. Par le terme, il entendait une description complète des caractéristiques sociales, politiques et économiques d'un état. Du simple dénombrement de populations humaines et de terres, la statistique est devenue une science qui a retenu et continue de retenir l'attention, non seulement des empereurs et de rois, mais surtout des personnes de sciences.

L'extension et l'utilisation du calcul des probabilités développé par J. BERNOULLI au 18^e siècle et l'application des études démographiques et sociales ont permis à cette science de connaître un essor considérable. Ainsi au 19^e siècle, de la simple statistique descriptive, elle passe au stade de Statistique Mathématiques. Depuis le 20^e siècle, les travaux de KARL PEARSON (1857-1936), et WILLIAM SEALY GOSSET (1876-1937) ont permis à cette science de connaître un développement considérable et une application vaste et variée. La statistique utilise les techniques et des méthodes de collecte, de présentation, d'étude et, d'analyse des données quantitatives.

La statistique n'est pas uniquement utilisée pour décrire, pour mieux connaître un événement survenu dans le passé mais elle intervient de plus en plus dans les travaux de planification,

dans le choix de prises de décisions et d'actions.

La statistique est l'étude de la collecte de données, leur analyse, leur traitement, l'interprétation des résultats et leur présentation afin de rendre les données compréhensibles par tous. C'est à la fois une science, une méthode et un ensemble de techniques. L'analyse des données est utilisée pour d'écrire les phénomènes étudiés, faire des prévisions et prendre des décisions à leur sujet.

Plus précieusement, l'objet de la méthode statistique est la réduction des données. Une masse de données doit être remplacée par un petit nombre de quantités représentant correctement cette masse, et contenant autant que possible la totalité de l'information pertinente contenue dans les données d'origine.



Sir Ronald Aymler Fisher est un biologiste et statisticien britannique, né le 17 février 1890 à East Finchley (Middlesex, Angleterre) et mort le 29 juillet 1962 à Adelaïde (Australie).

La citation de Sir Ronald Aymler Fisher ci-dessus résume bien l'objet de la méthode statistique.

Avant de commencer, précisons le vocabulaire utilisé dans cette partie des mathématiques.

La statistique est la science qui consiste à réunir des données chiffrées à les analyser et à les commenter.

Sans rentrer dans les détails, Le terme probabilité possède plusieurs sens : venu historiquement du latin *probabilitas*, il désigne l'opposé du concept de certitude, il est également une évaluation du caractère probable d'un événement. aussi la théorie des probabilités débute réellement avec les travaux de Pascal et de Fermat.

La théorie fut ensuite développée par de nombreuses personnes, dont Huygens(1629-1695). J. Bernoulli(1662-1716), de Moivre , D. Bernoulli, Euler, Gauss et Laplace (1749-1827). La théorie moderne des probabilités est fondée sur l'approche axiomatique de Kolmogorov (1959), basée sur la théorie de la mesure de Borel et Lebesgue. Grâce à cette approche, la théorie a alors connu un développement très rapide tout au long du XXème siècle.

En cela, la statistique et la théorie probabilités sont un outil essentiel pour la compréhension et la gestion des phénomènes complexes sans oublier les probabilités. Les données étudiées peuvent être de toute nature, ce qui rend la statistique utile dans tous les champs disciplinaires et explique pourquoi elle est enseignée dans toutes les filières universitaires, de l'économie à la biologie en passant par la psychologie et bien sûr les sciences de l'ingénieur.

On va donner quelques détails des notions de base et vocabulaire concernant la statistique et les calculs des probabilités à travers ce modeste travail.

Notions de base et vocabulaire statistique

La statistique est l'étude de la collecte de données, leur analyse, leur traitement, l'interprétation des résultats et leur présentation afin de rendre les données compréhensibles par tous. C'est à la fois une science, une méthode et un ensemble de techniques. Alors, une étude statistique s'effectue sur un ensemble appelé population dont les éléments sont appelés individus et consiste à observer et étudier un même aspect sur chaque individu, appelé caractère. L'analyse des données est utilisée pour d'écrire les phénomènes étudiés, faire des prévisions et prendre des décisions à leur sujet. En cela, la statistique est un outil essentiel pour la compréhension et la gestion des phénomènes complexes. Les données étudiées peuvent être de toute nature, ce qui rend la statistique utile dans tous les champs disciplinaires et explique pourquoi elle est enseignée dans toutes les filières universitaires, de l'économie à la biologie en passant par la psychologie et bien sûr les sciences de l'ingénieur. La statistique consiste à :

Recueillir des données.

Présenter et résumer ces données.

Tirer des conclusions sur la population étudiée et d'aider à la prise de décision.

En présence de données dépendant du temps, nous essayons de faire de la prévision.

1. Vocabulaire

Les statistiques consistent en diverses méthodes de classement des données tels que les tableaux, les histogrammes et les graphiques, permettant d'organiser un grand nombre de données. Les statistiques se sont développées dans la deuxième moitié du XIXe siècle dans le domaine des sciences humaines (sociologie, économie, anthropologie, ...). Elles se sont dotées d'un vocabulaire particulier.

2. Épreuve statistique

Les statistiques descriptives visent à étudier les caractéristiques d'un ensemble d'observations comme les mesures obtenues lors d'une expérience. L'expérience est l'étape préliminaire à toute étude statistique. Il s'agit de prendre **contact** avec les observations. De manière générale, la méthode statistique est basée sur les concepts suivants : la population, les variables, les observations et les données.

DÉFINITION 1.1. *L'épreuve statistique est une expérience que l'on provoque.*

EXEMPLE 1.1. (*La durée de vie des lampes*)

Imaginons le cas suivant : un fabricant d'ampoules électriques ayant le choix entre 4 types de filaments se propose d'étudier l'influence de la nature du filament sur la durée de vie des ampoules fabriquées. Pour ce faire, il va faire fabriquer 4 échantillons d'ampoules identiques, sauf en ce qui concerne le filament, faire brûler les ampoules jusqu'à extinction, puis comparer les résultats obtenus.

3. Concepts de base de la statistique notions de base

D'une manière générale, la méthode statistique est basée sur quatre concepts : la population, les variables, les observations et les données.

4. Population

En statistique, on travaille sur des populations. Ce terme vient du fait que la démographie, étude des populations humaines, a occupé une place centrale aux débuts de la statistique, notamment au travers des recensements de population. Mais, en statistique, le terme de population s'applique à tout objet statistique étudié, qu'il s'agisse d'étudiants (d'une université ou d'un pays), de ménages ou de n'importe quel autre ensemble sur lequel on fait des observations statistiques. Nous définissons la notion de population.

DÉFINITION 1.2. *On appelle population l'ensemble sur lequel porte notre étude statistique. Cet ensemble est noté Ω .*

EXEMPLE 1.2. *On considère l'ensemble des étudiants de la section A. On s'intéresse au nombre de frères et sœurs de chaque étudiant. Dans ce cas*

$\Omega :=$ *ensemble des étudiants.*

Si l'on s'intéresse maintenant à la circulation automobile dans une ville, la population est alors constituée de l'ensemble des véhicules susceptibles de circuler dans cette ville à une date donnée. Dans ce cas

$\Omega =$ *ensemble des véhicules.*

Individu (unité statistique)

Une population est composée d'individus. Les individus qui composent une population statistique sont appelés unités statistiques.

DÉFINITION 1.3. *On appelle individu tout élément de la population, il est noté ω ; tel que (ω dans Ω).*

REMARQUE 1.1. *L'ensemble peut être un ensemble de personnes, de choses ou d'animaux... L'unité statistique est un objet pour lequel nous sommes intéressés à recueillir de l'information.*

EXEMPLE 1.3. *Dans l'exemple indiqué ci-dessus, un individu est tout étudiant de la section. Si on étudie la production annuelle d'une usine des boîtes de boisson en métal canettes). La population est l'ensemble des boîtes produites durant l'année et une boîte constitue un individu.*

5. Variable (ou caractère)

Chaque individu d'une population est décrit par un ensemble de caractéristiques appelées variable ou caractère. On les représente souvent par des lettres $x, y, \dots$

Les valeurs que peut prendre une variable statistique sont appelées modalités.

Une variable doit donc présenter au minimum deux modalités.

EXEMPLE 1.4. 1/ *La variable **sexe** a deux modalités : masculin ou féminin.*
2/ *les modalités de la variable **age** des ouvriers d'une entreprise peuvent être :*

$$[25, 30[; [30, 40[; [40, 45[\quad \text{et} \quad [45, 60[.$$

a) Variable quantitative :

Une variable est dite **quantitative** si ses différentes valeurs sont mesurables.

Il y a deux types de variables quantitative :

• Variable discrète :

Une variable quantitative est discrète si elle prend un nombre fini ou dénombrable de valeurs.

EXEMPLE 1.5. *La variable **nombre d'enfants par famille**, dont les valeurs sont :*

$$0, 1, 2, 3, \quad \text{ect...}$$

• Variable continue :

Une variable quantitative est dite continue si elle prend toutes ses valeurs dans des intervalles de l'ensemble des nombres réels $\mathbb{R}$.

EXEMPLE 1.6. *La variable (**taille en cm**) observée sur un échantillon de 100 individus est une variable continue, dont les résultats sont les suivants :*

30 individus sont leurs tailles comprises dans l'intervalle $]150 \text{ cm}, 160 \text{ cm}[$ 50 individus sont leurs tailles comprises dans l'intervalle $[160 \text{ cm}, 170 \text{ cm}[$ et 20 individus sont leurs tailles comprises dans l'intervalle $[170 \text{ cm}, 180 \text{ cm}[$

b) Variable qualitative

Une variable est dite qualitative si ses différentes modalités ne sont pas mesurable.

EXEMPLE 1.7. *L'état d'une maison : on peut considérer les modalités suivantes*
Ancienne.

Dégradée.

Nouvelle.

Rénovée.

On peut résumer les catégories de caractères dans la figure suivante :



6. Observations

Une variable donne lieu à plusieurs observations. Si la variable est notée x , alors on notera par $x_1, x_2, x_3, \dots, x_n$, les observations appelées série statistique, où n représente le nombre d'observations. Si les observations portent sur la totalité de la population, on dit que N est la taille (ou effectif) de la population. Si les observations n'ont porté que sur une partie de la population, on dit que n est la taille de l'échantillon de la population observée.

EXEMPLE 1.8. *Soit un échantillon de dix individus, la variable âge observés peut prendre les valeurs suivantes :*

$$x_1 = 28, x_2 = 40, x_3 = 25, x_4 = 30, x_5 = 35, x_6 = 40, x_7 = 28, x_8 = 42, x_9 = 33, x_{10} = 37.$$

Dans ce cas la taille de l'échantillon est $N = 10$.

7. Données

Lorsque les observations portent sur un ou plusieurs variables (on note p variables) et sur un ou plusieurs individus (on note n individus), l'ensemble des valeurs des observations récoltées constituent les données de l'étude ou l'expérience effectuée. On peut les mettre sous forme d'un tableau individu/variable à n ligne et p colonnes :

Individus/Variable	x_1	x_2		x_p
I_1	x_{11}	x_{12}		x_{1p}
I_2	x_{21}	x_{22}		x_{2p}
.	.	.		.
.	.	.		.
I_n	x_{n1}	x_{n2}		x_{np}

Chaque individu I est représenté par un numéro d'ordre $i = 1, 2, \dots, n$, chaque variable x est indexée d'un numéro d'ordre correspondant $j = 1, 2, \dots, p$.

x_{ij} est la valeur prise par la j^{ieme} variable pour le i^{ieme} individu.

EXEMPLE 1.9. *Soit un échantillon de dix individus, sur les quels on observe deux variables x_1 "le poids" et x_2 "l'âge", les résultats de cette étude sont représentés dans le tableau suivant :*

<i>Individus/Variable</i>	$x_1=\text{poids}$	$x_2=\text{âge}$
I_1	40	12
I_2	50	18
I_3	65	25
I_4	75	30
I_5	80	42
I_6	72	37
I_7	78	25
I_8	68	38
I_9	55	19
I_{10}	64	24

En général, la variable quantitative discrète est une variable ne prenant que des valeurs entières (plus rarement décimales). Le nombre de valeurs distinctes d'une telle variable est habituellement assez faible. Citons, par exemple, le nombre des maisons par quartier d'une ville. Une variable quantitative est dite continue lorsque les observations qui lui sont associées ne sont pas des valeurs précises, mais des intervalles. C'est le cas lorsque nous avons un grand nombre d'observations distinctes.

La statistique descriptive a pour objectif de synthétiser l'information contenue dans les jeux de données au moyen de tableaux, figures ou résumés numériques. Les variables statistiques sont analysées différemment selon leur nature (quantitative, qualitative).

8. Étude d'une variable statistique discrète

Le caractère statistique peut prendre un nombre fini raisonnable de valeurs (note, nombre d'enfants, nombre de pièces, ...). Dans ce cas, le caractère statistique étudié est alors appelé un caractère discret. Dans toute la suite, nous considérons la situation suivante :

$$X : \Omega \rightarrow \{x_1, x_2, \dots, x_n\}$$

avec $\text{Card}(\Omega) := N$ est le nombre d'individus dans notre étude. Nous allons utiliser souvent l'exemple ci-dessous pour illustrer les énoncés de ce cours.

EXEMPLE 1.10. Une enquête réalisée dans un village porte sur le nombre d'enfants à charge par famille. On note x le nombre d'enfants, les résultats sont données par ce tableau :

x_i	0	1	2	3	4	5	6
$n_i(\text{Effectif})$	18	32	66	41	32	9	2

Nous avons Ω ensemble des familles.

w une famille.

x nombre d'enfants par famille.

$$X : \omega \rightarrow X(\omega)$$

On lit, à la famille, on associe $X(\omega)$ = le nombre d'enfants de cette famille.

8.1. Effectif partiel - effectif cumulé. On étudie ici un caractère statistique numérique représenté par une suite x_i décrivant la valeur du caractère avec i varie de 1 à k .

9. Effectif partiel (fréquence absolue)

DÉFINITION 1.4. *Pour chaque valeur x_i , on pose par définition*

$$n_i = \text{Card}\{\omega \in \Omega : X(\omega) = x_i\}.$$

n_i : le nombre d'individus qui ont le même

x_i , s'appelle effectif partiel de x_i .

EXEMPLE 1.11. *Dans l'exemple précédent, 66 est le nombre de familles qui ont 2 enfants.*

x_i	...	2	...
n_i (Effectif)	...	66	...

10. Effectif cumulé

DÉFINITION 1.5. *Pour chaque valeur x_i , on pose par définition*

$$N = n_1 + n_2 + \dots + n_k.$$

L'effectif cumulé N d'une valeur est la somme de l'effectif de cette valeur et de tous les effectifs des valeurs qui précèdent.

EXEMPLE 1.12. *Dans l'exemple précédent, 50 est le nombre de familles qui ont un nombre d'enfant inférieur à 1. Nous le regardons dans le tableau suivant :*

x_i	0	1	2	3	4	5	6
n_i (Effectif)	18	50	116	157	189	198	200

Interprétation :

N est le nombre d'individus dont la valeur du caractère est inférieur ou égale à x_i . De ce fait,

l'effectif total est donné par

$$N = \text{card}(\Omega) = \sum_{i=0}^k n_i$$

Dans notre exemple précédent, nous avons $N = 200$.

Dans cette remarque on a :

REMARQUE 1.2. *les fréquence partielle et les fréquence cumulée typiquement se sont les effectifs n_i grands et donc il est intéressant de calculer des grandeurs permettant de résumer la série.*

11. Fréquence partielle (fréquence relative)

DÉFINITION 1.6. *Pour chaque valeur x_i , on pose par définition*

$$f_i := \frac{n_i}{N}$$

f_i s'appelle la fréquence partielle de x_i . La fréquence d'une valeur est le rapport de l'effectif de cette valeur par l'effectif total.

REMARQUE 1.3. *On peut remplacer f_i par $f_i \times 100$ qui représente alors un pourcentage.*

Interprétation :

f_i = est le pourcentage des ω tel que $X(\omega) = x_i$.

EXEMPLE 1.13. *Dans l'exemple précédent, $0,33 := 33\%$ il y a de familles dont le nombre d'enfants égale à 2. Ce pourcentage est calculé de la façon suivante ($N = 200$) :*

x_i	...	2	...
n_i (Effectif)	...	66	...
f_i (Effectif)	...	0,33	...

On a la remarque suivante :

REMARQUE 1.4. *une fréquence est toujours comprise entre 0 et 1. La somme des fréquences est égale à 1.*

Nous pouvons conclure la propriété suivante.

PROPOSITION 1.1. *Soit f_i défini comme précédemment. Alors,*

$$\sum_{i=0}^k f_i = 1$$

DÉMONSTRATION. Rappelons que

$$\sum_{i=0}^k n_i = N$$

Ce qui implique que

$$\sum_{i=0}^k f_i = \sum_{i=0}^k \frac{n_i}{N} = \frac{1}{N} \sum_{i=0}^k n_i = 1.$$

□

12. Fréquence cumulée

DÉFINITION 1.7. Pour chaque valeur x_i , on pose par définition

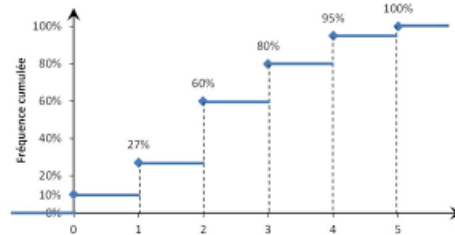
$$F_i = f_1 + f_2 + \dots + f_k.$$

Alors $F_1 = f_1$ et $F_2 = f_1 + f_2, \dots, F_k = f_1 + f_2 + \dots + f_k = 1$

La quantité F_i s'appelle la fréquence cumulée de x_i .

Interprétation :

F_i est le pourcentage des ω tel que la valeur $x(\omega)$ est inférieure ou égale à x_i .



EXEMPLE 1.14. Dans l'exemple précédent, 0.785 représente 78.5 % de familles dont le nombre d'enfants est inférieur ou égale à 3.

Dans un deuxième exemple, nous nous intéressons aux nombres d'erreurs d'assemblage sur un ensemble d'appareils,

Nombre d'erreurs	Nombre d'appareils	Fréquences cumulées
0	101	0.26
1	140	0.61
2	92	0.84
3	42	0.94
4	18	0.99
5	3	1

Nous avons 94 % des appareils qui ont un nombre d'erreurs d'assemblage inférieur ou égale à 3.

13. Tableaux Statistiques

Nous avons vu que les tableaux sont un moyen souvent indispensable, en tous cas très utile, de classification et de présentation des unités d'une population statistique. Dans le paragraphe suivant, nous allons voir comment on traduit ses tableaux en graphique permettant aussi de résumer d'une manière visuelle les données.

Présentation des données sous forme de tableaux.

On distingue les méthodes de représentation d'une variable statistique en fonction de la nature de cette variable (qualitative ou quantitative). Les représentations recommandées et les plus fréquentes sont les tableaux et les diagrammes (graphe).

Au cours d'une étude statistique les données sont recueillies de façon désordonnée. Cette masse d'information ne peut fournir des renseignements que si l'information en question est classée et mise en ordre, ce classement ne peut se faire que dans des tableaux statistiques qui servent de documentations statistiques. Ces tableaux doivent comporter :

- Le titre (indiquant l'objet du tableau).
- L'unité utilisée (Dinars, milliers de tonnes, milliers de personnes, Le pourcentage...).
- Les titres des lignes et des colonnes.
- Les sources qui doivent être citées.

Leur présentation diffère selon la nature des variables.

1.a) Cas d'une variable discrète :

Lorsque la variable statistique est discrète, il arrive souvent qu'une même valeur x_i ($i = 1, \dots, k; k = \text{nombre de modalité}$) soit observée, sur plusieurs individus, on note alors n_i son effectif (fréquence absolue), qui est le nombre d'observations égale à x_i et f_i la fréquence relative correspondante : $f_i = \frac{n_i}{N}, i = 1, 2, \dots, k$. Le tableau statistique se présente alors sous la forme suivante :

x_i	n_i	f_i
.	.	.

Remarque

$$\sum_{i=1}^k n_i = N, \quad \sum_{i=1}^k f_i = 1.$$

Exemple

Une étude sur la répartition des ménages selon le nombre de personnes qui les composent nous donne le tableau suivant :

<i>Nombre de personnes dans un ménages x_i</i>	<i>Nombre de ménages n_i</i>
<i>1</i>	<i>50</i>
<i>2</i>	<i>80</i>
<i>3</i>	<i>90</i>
<i>4</i>	<i>120</i>
<i>5</i>	<i>350</i>
<i>6</i>	<i>160</i>
<i>7 et plus</i>	<i>150</i>
<i>Total</i>	<i>1000</i>

1.b) Cas d'une variable continue :

Lorsque la variable statistique est continue, on regroupe les observations en k — classes d'extrémités $c_0, c_1, \dots, c_k$.

On note pour chaque classe $[c_0, c_1[$ l'effectif n_i et la fréquence relative f_i . Le tableau statistique se présente alors sous la forme suivante :

$[c_i, c_{i+1}[$	n_i	f_i
.	.	.

Par convention la borne supérieur d'une classe est exclue de cette classe, sauf (la dernière classe peut être fermée). On appelle "centre" d'une classe $[c_i, c_{i+1}[$ la valeur $C_i = \frac{c_i + c_{i+1}}{2}$ et "l'amplitude" de cette classe est notée a_i , telle que : $a_i = c_{i+1} - c_i$.

14. Règle de Sturge et Yule

La règle de Sturge est utilisée pour déterminer le nombre de classe à utiliser pour représenter une variable statistique continue, et ainsi la règle de Yule.

Le nombre de classes est égal à l'entier le plus proche de la quantité ; $1 + 3,3 \log n$.

EXEMPLE 1.15. *Par exemple, pour un échantillon de taille $n = 200$, on doit utiliser*

$$1 + 3,3 \log 200 = 1 + 3,3 \times 2,3 = 8,59 \approx 9 \text{ classes.}$$

L'amplitude constante de ces classes sera égale à :

$$a = \frac{x_{\max} - x_{\min}}{\text{Nombre de classe}}$$

Toujours, on commence par calculer Le nombre de classes. Ensuite, on applique les formules (Sturge et Yule).

On peut résumer ces règles dans la conclusion suivante :

Conclusion :

- Nombre de classes K:

Afin de trouver le nombre de classes on utilise

- la formule de Sturge

Le nombre de classes = $1 + (3,3 \log n)$, avec $n=50$

On trouve que le nombre de classes égale à 6,6 ;

On admet 7.

- Formule de Yule

$$K = 2,5 \sqrt[4]{N} = 2,5 \sqrt[4]{50} = 6.64 \approx 7 \text{ classes.}$$

15. Distribution et fonction de répartition

La distribution d'une variable statistique x est l'ensemble des couples $(x_i, f_i), i = 1, \dots, k$ ($k =$ nombre de modalité de la variable étudiée).

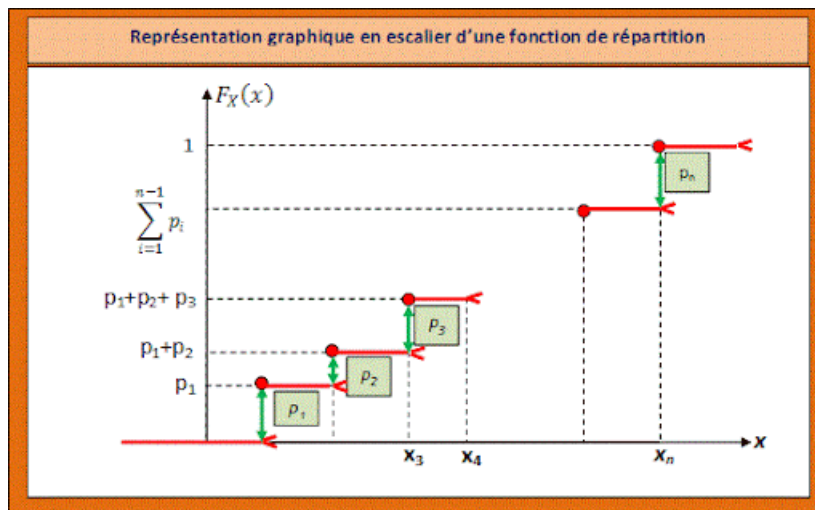
EXEMPLE 1.16. $(0; 0, 2), (1; 0, 3), (2; 0, 4)(3; 0, 1))$ est une distribution statistique représentant la variable "nombre d'enfants par famille".

16. Fonction de répartition

La fonction de répartition $F(x)$, appelée aussi fonction des fréquences cumulées, associé à nombre réel x la proposition d'individus pour lesquels la variable statistique x prend des valeurs inférieures à x , donc :

$$F(x) = \sum_{x_i \leq x} f_i.$$

REMARQUE 1.5. Pour tout $i \in 1, \dots, n$, on a $F(x_i) = F_i$. La courbe de F passe par les points $(x_1, F_1), (x_2, F_2), \dots$ et (x_n, F_n) .



EXEMPLE 1.17. Les taux de glycémie, observés sur un échantillon de 100 patients sont donnés dans le tableau suivant :

Taux de glycémie	$[0, 8, 0, 85[$	$[0, 85, 0, 9[$	$[0, 9, 0, 95[$	$[0, 95, 1[$	$[1, 1, 20[$	Total
Effectifs n_i	10	20	30	30	10	100
Fréquences f_i	0,1	0,2	0,3	0,3	0,1	1
$F(x)$	0	0,1	0,3	0,6	0,9	1

Pour :

$$x = 0,8; F(0,8) = \sum_{x_i \leq 0,8} f_i = 0$$

$$x = 0,85; F(0,85) = \sum_{x_i \leq 0,85} f_i = 0,1$$

$$x = 0,9; F(0,9) = \sum_{x_i \leq 0,9} f_i = f_1 + f_2 = 0,3$$

$$x = 0,95; F(0,95) = \sum_{x_i \leq 0,95} f_i = f_1 + f_2 + f_3 = 0,6$$

$$x = 1; F(1) = \sum_{x_i \leq 1} f_i = f_1 + f_2 + f_3 + f_4 = 0,9$$

$$x = 1,2; F(1,2) = \sum_{x_i \leq 1,2} f_i = f_1 + f_2 + f_3 + f_4 + f_5 = 1$$

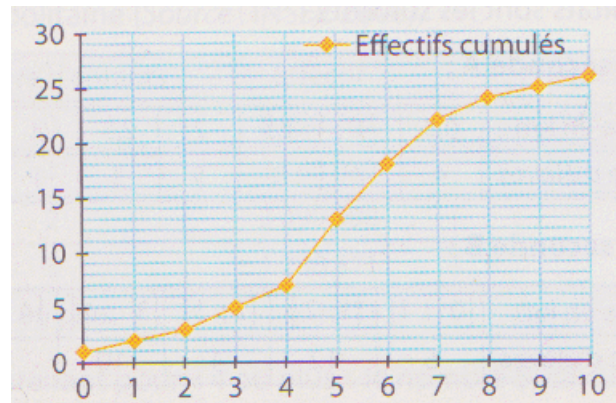
$F(0,95) = 0,6$ par exemple, veut dire que 60 % des patients ont un taux de glycémie inférieur à 0,95.

Dans cette remarque, on a :

REMARQUE 1.6. • $F(x) = 0$, si $x \leq \min x_i$ et $F(x) = 1$, si $x \geq \max x_i$.

• F est une fonction monotone croissante, elle prend ses valeurs dans l'intervalle $[0, 1]$ et présente lorsque x est discrète une discontinuité en chacun des points n_i qui se manifeste par un saut dont la valeur est f_i .

Ainsi la fonction de répartition précédente peut être représentée graphiquement par une courbe continue (courbe cumulative) :



EXEMPLE 1.18. Une enquête faite sur les pointures de 200 filles a fourni les résultats suivants :

x_i	36	37	38	39	40	41	Total
n_i	30	25	70	40	15	20	200
f_i	0,15	0,125	0,35	0,2	0,075	0,1	1
$F(x)$	0	0,15	0,275	0,625	0,825	0,9	1

Comme la variable étudiée est discrète, alors la fonction de répartition va être représentée par une courbe en escaliers :

17. Courbe cumulatives croissantes et décroissante

Pour définir la fonction de répartition, on calcule les fréquences cumulées croissantes, qu'on représente graphiquement par une courbe cumulatives croissantes.

On a la distribution groupée des tailles des 175 étudiants suivantes :

Taille	[149, 5; 159, 5[	[159, 5; 169, 5[	[169, 5; 179, 5[	[179, 5; 189, 5[	[189, 5; 199, 5[	Total
Effectifs n_i	14	32	65	47	17	175
Fréquences f_i	0,08	0,18	0,37	0,27	0,10	1
F_i	0	0,08	0,26	0,63	0,9	1

../Pictures/courbe cumulative nou.png

De la même façon on peut calculer aussi les fréquences cumulées décroissante, à qui correspond une courbe cumulative décroissante. notée :

$$F(x) \downarrow = \sum_{x_i \geq x} f_i$$

Calculons les fréquences cumulées décroissante de la variable étudiée dans l'exemple suivant :

Taux de glycémie	[0, 8, 0, 85[	[0, 85, 0, 9[	[0, 9, 0, 95[	[0, 95, 1[	[1, 1, 20[	Total
Effectifs n_i	10	20	30	30	10	100
Fréquences f_i	0,1	0,2	0,3	0,3	0,1	1
Fréquences $\downarrow$	1	0,9	0,7	0,4	0,1	0
$F(x)$	10	20	30	30	10	100

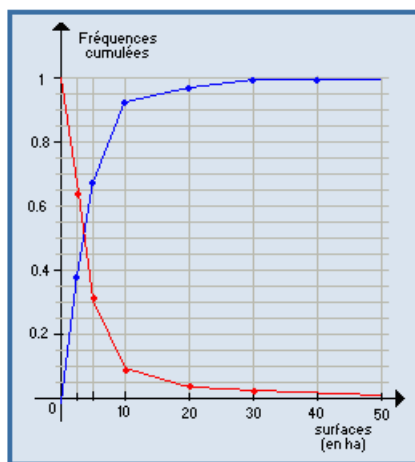
$F(x) \downarrow = 0,7$; c'est à dire 70 % des patients ont un taux de glycémie $\geq 0,9$.

$F(x) \uparrow = 0,3$; c'est à dire 30 % des patients ont un taux de glycémie $\leq 0,9$.

Revenons à l'exemple précédent.

x_i	36	37	38	39	40	41	Total
n_i	30	25	70	40	15	20	200
f_i	0,15	0,125	0,35	0,2	0,075	0,1	1
$F(x) \downarrow$	1	0,85	0,725	0,375	0,175	0,1	0

REMARQUE 1.7. Les courbes cumulatives croissantes et décroissantes peuvent être obtenues aussi à l'aide des effectifs cumulés croissantes (notés $N_i \uparrow$) et les effectifs cumulés décroissantes (notés $N_i \downarrow$)



18. Introduction

Bien qu'un tableau statistique contient l'information prélevée sur l'échantillon mais il est très utile de traduire ces informations par un graphique, afin d'avoir une synthèse visuelle. Suivant la nature de la variable étudiée, on utilise différents types de représentations.

19. Cas de variable qualitative

Une variable qualitative peut être représentée graphiquement de trois manières.

20. Représentation circulaire par des secteurs

Les diagrammes circulaires, ou semi-circulaires, consistent à partager un disque ou un demi-disque, en tranches, ou secteurs, correspondant aux modalités observées et dont la surface est proportionnelle à l'effectif, ou à la fréquence, de la modalité.

DÉFINITION 1.8. Un diagramme circulaire est une représentation graphique de données statistiques sous la forme d'un disque partagé en secteurs.

Elle consiste à représenter sur un cercle chaque modalité par un secteur dont le degré est déterminé à l'aide de la règle de trois de la manière suivante :

Supposons qu'un veut calculer le degré du secteur représentant la modalité i d'effectif partiel n_i alors on peut écrire :

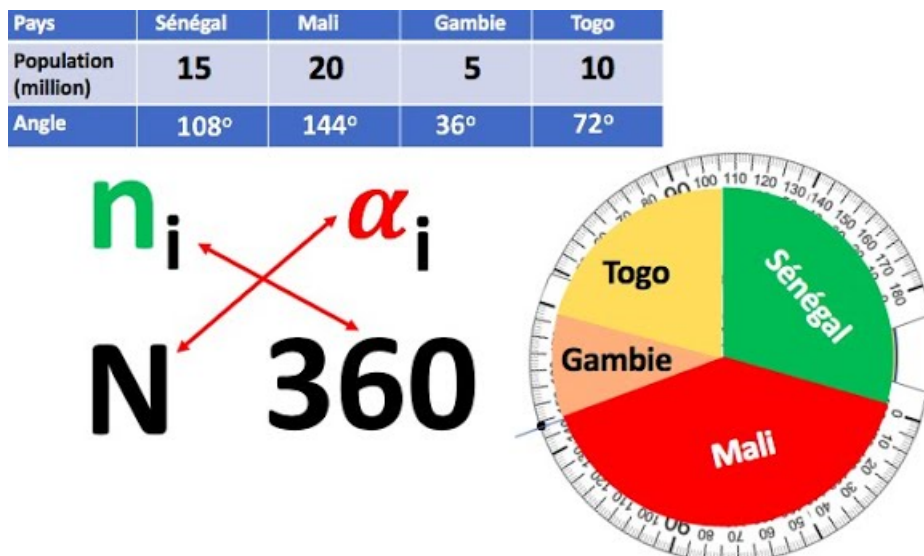
$$N \rightarrow 360 \%$$

$$n_i \rightarrow d_i \%$$

Donc :

$$d_i = \frac{n_i \times 360 \%}{N}$$

On a comme la figure suivante :



EXEMPLE 1.19. On a étudié le transport des gens sur un échantillon de 200 personnes donné, les résultats obtenus sont les suivants :

transport	en voiture	à pied	en métro
Nombre des personnes n_i	70	50	80
f_i	0,35	0,25	0,4

Calculons le degré correspondant aux différents secteurs :

$$d_1 = f_1 \cdot 360 \% = 0,35 \cdot 360 \% = 126 \%$$

$$d_2 = f_2 \cdot 360 \% = 0,25 \cdot 360 \% = 90 \%$$

$$d_i = f_i \cdot 360^\circ = 0,4 \cdot 360^\circ = 144^\circ$$



REMARQUE 1.8. *Un diagramme semi-circulaire se présente sous la forme d'un anneau ou demi-disque.*

EXEMPLE 1.20. *Le tableau suivant donne la répartition selon le groupe sanguin de 60 individus pris au hasard dans une population,*

Groupes sanguins	A	B	AB	O
L'effectif	20	10	25	5

Donner la représentation graphique possible de cette distribution sous la forme d'un anneau.

21. Représentation en tuyaux d'orgue

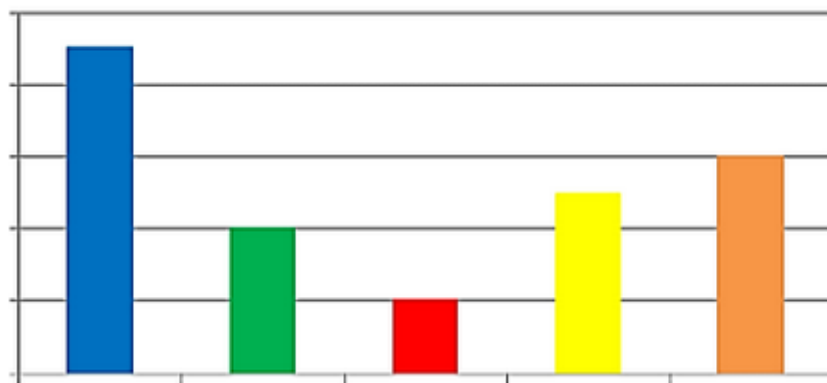
Elle consiste à porter en abscisse les modalités de la variable étudiée ; au dessus des quelles on trace des tuyaux d'orgue de même largeur, et en ordonnée les effectifs ou les fréquences qui vont déterminer leurs hauteur.

Exemple

Une enquête auprès de 500 étudiants de première année SETI (sciences exactes, technologie et informatiques) admis en deuxième année, sur les spécialités préférées a fourni les résultats suivants :

spécialité préférée	Nombre d'étudiants n_i
Informatique (I)	160
Physique (P)	70
Chimie (C)	50
Mathématiques (M)	100
TCT	120

Ces résultats peuvent être par le diagramme en tuyaux d'orgue suivant :



22. Diagramme en bandes

Il consiste à représenter dans une même bande verticale, chaque modalité par une tranche dont la hauteur correspond à sa fréquence. A partir de l'observation d'une variable qualitative, deux diagrammes permettent de représenter cette variable : le diagramme en bandes (dit tuyaux d'orgue) et le diagramme à secteurs angulaires (dit camembert).

23. Distribution à caractère quantitatif

A partir de l'observation d'une variable quantitative discrète, deux diagrammes permettent de représenter cette variable : le diagramme en bâtons et le diagramme cumulé. On va

distinguer les diagramme correspondant à une variable discrète et ceux correspondants à une variable continue.

24. Diagramme en bâtons

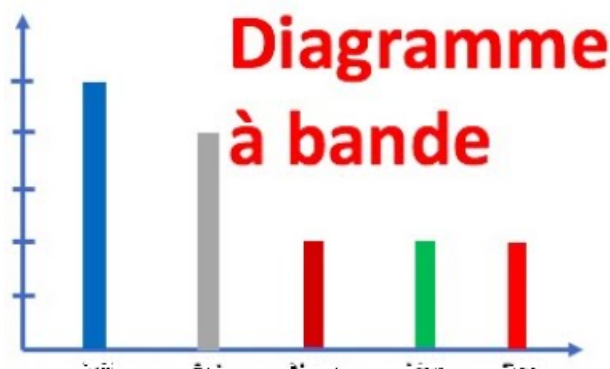
On veut représenter cette répartition sous la forme d'un diagramme en bâtons. À chaque marque correspond un bâton. Les hauteurs des bâtons sont proportionnelles aux effectifs représentés.

DÉFINITION 1.9. Un diagramme en bâtons est une représentation graphique de données statistiques à l'aide de segments. Les valeurs du caractère étudié sont représentées sur l'axe horizontal, les effectifs sur l'axe vertical. À chaque valeur correspond un bâton. Les hauteurs des bâtons sont proportionnelles aux effectifs représentés.

EXEMPLE 1.21. Voici la répartition de 450 collégiens selon la marque de leurs baskets.

Marque	Adidas	Nike	Puma	Autres marques	Total
Effectif	200	100	50	50	50

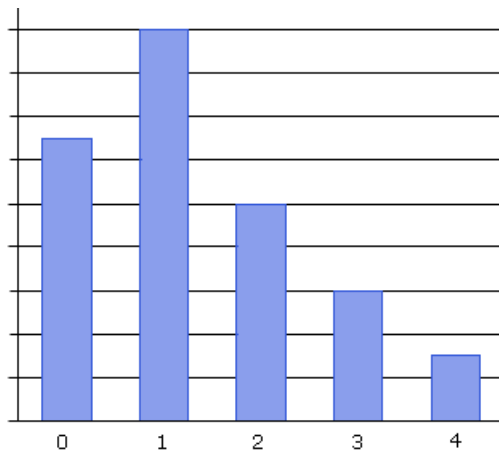
On veut représenter cette répartition sous la forme d'un diagramme en bâtons. À chaque marque correspond un bâton. Les hauteurs des bâtons sont proportionnelles aux effectifs représentés. On obtient le diagramme suivant :



Pour l'illustration, nous prenons un autre exemple. On a un autre exemple.

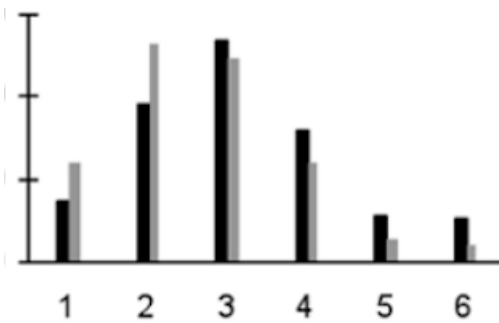
EXEMPLE 1.22. nombre d'enfants par famille. Nous rappelons le tableau statistique associée :

x_i	0	1	2	3	4
n_i	80	100	60	40	30



EXEMPLE 1.23. Une enquête effectuée sur la répartition de 200 foyers selon le nombre de chambre a donné les résultats suivant :

Nombre de chambre x_i	1	2	3	4	5	6	Total
n_i	10	50	60	40	20	20	200



25. Histogramme

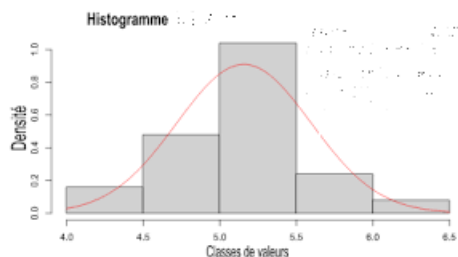
L'histogramme est utilisé pour représenter une variable continue.

On représente sur l'axe des abscisses les différentes classes (supposées de même amplitude) et au dessus de chacune d'elle on trace des rectangles, dont les hauteurs sont directement proportionnelles aux fréquences ou aux effectifs.

EXEMPLE 1.24. Les notes au module de statistique, observées sur un échantillon de 100 étudiants sont données le tableau suivant :

Notes	$[0,4[$	$[4,8[$	$[8,12[$	$[12,16[$	$[16,20[$	Total
n_i (Effectif)	5	20	30	35	10	100

La variable étudiée est continue, donc elle peut être représentée graphiquement au moyen de l'histogramme suivant :



REMARQUE 1.9. On a les remarques suivantes :

- Ce mode de représentation est valable seulement dans le cas où on a des classes d'intervalles d'amplitudes constantes.
- Dans le cas où les amplitudes sont différentes, pour tracer l'histogramme on utilise **La méthode des effectifs rectifiée**, qui consiste à multiplier l'effectif de chaque classe par le rapport entre l'amplitude de la plus petite classe sur l'amplitude de la classe considérée, on obtient ainsi les **effectifs rectifiés** notée n'_i , tel que :

$$n'_i = n_i \times \frac{\text{la plus petite amplitude}}{\text{amplitude de la classe } i}.$$

EXEMPLE 1.25. Les notes au module de biochimie, observées sur un échantillon de 100 étudiants sont données le tableau suivant :

Notes	$[0,4[$	$[4,10[$	$[10,17[$	$[17,20[$	Total
Effectifs n_i	10	25	45	20	100
Amplitudes a_i	4	6	7	3	/
Effectifs rectifiés n'_i	$10 \times \frac{3}{4} = 7$	$25 \times \frac{3}{6} = 13$	$45 \times \frac{3}{7} = 19$	$20 \times \frac{3}{3} = 20$	/

On peut aussi représenter l'histogramme autrement, en utilisant comme hauteur des rectangles les quantités $\frac{f_i}{a_i}$ (ou $\frac{n_i}{a_i}$)

Notes	$[0,4[$	$[4,10[$	$[10,17[$	$[17,20[$	Total
n_i	10	25	45	20	100
f_i	0,1	0,25	0,45	0,2	1
a_i	4	6	7	3	/
$\frac{f_i}{a_i}$	0,025	0,042	0,064	0,066	/

REMARQUE 1.10. La surface de chaque rectangle est égale à la fréquence de la classe correspondante n'_i : $S_i = a_i \times \frac{f_i}{a_i} = f_i$, donc la surface totale de l'histogramme est égal à 1.

26. Polygone

Le polygone est construit à partir du diagramme en bâton ou de l'histogramme.

Il s'agit d'une ligne brisée qui joint les points supérieurs des bâtons ou les points milieux des cotés supérieurs des rectangles successifs de l'histogramme. On commence par l'exemple suivant dans le discret :

Cas discret :

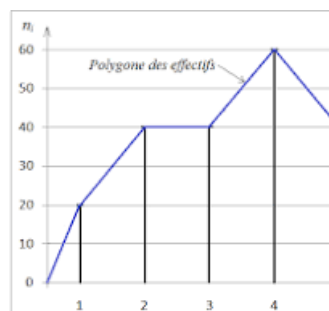
EXEMPLE 1.26. Voici la répartition suivante :

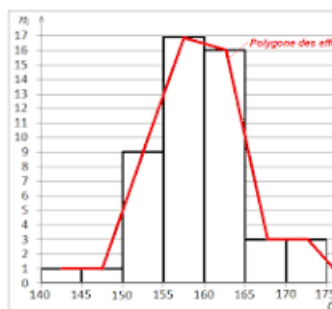
x_i	1	2	3	4	5
Effectif	20	40	40	60	40

On veut représenter cette répartition sous la forme d'un diagramme en bâtons.

Ensuite on trace le polygone, on obtient le polygone des effectifs suivant :

L'exemple suivant dans le cas continu :



Cas continu :**27. Étude d'une variable statistique continue**

Nous rappelons qu'une variable statistique (v.s) quantitative concerne une grandeur mesurable. Ses valeurs sont des nombres exprimant une quantité et sur lesquelles les opérations arithmétiques (addition, multiplication, etc,...) ont un sens. Nous allons dans ce chapitre se focaliser sur la (v.s) quantitative continue.

28. Caractère continu

DÉFINITION 1.10. On appelle v.s continue (ou caractère continu) toute application de et à valeurs réelles et qui prend un nombre **important** de valeurs (Les caractères continus sont ceux qui ont une infinité de modalités).

EXEMPLE 1.27. Soit l'ensemble des nouveaux nés au C.H.U d'une ville pendant les 3 premiers mois de 2019. Nous désignons par X le poids des nouveaux nés. On suppose que $x_{\min} = 2.701$ et $x_{\max} = 5.001$.

REMARQUE 1.11. Comment étudier ce caractère ?

Réponse : Partager les valeurs prises par x en classes de valeurs.

28.1. Classe de valeurs.

DÉFINITION 1.11. On appelle classe de valeurs de x un intervalle de type $[a, b[$ tel que $x \in [a, b[$ si et seulement si $a \leq x(w) < b$, c'est à dire, que les valeurs du caractère sont dans la classe $[a, b[$.

Dès qu'un caractère est identifié en tant que continu, ces modalités $C_k = [L_k, L_{k+1}[$ sont des intervalles avec

○ L_k : borne inférieure.

○ L_{k+1} : borne supérieure.

○ $a_k = L_{k+1} - L_k$: son amplitude, son pas ou sa longueur.

○ $C_k = x_k = \frac{(L_{k+1} + L_k)}{2}$: son centre.

REMARQUE 1.12. *On supposera dans tous les cas étudiés que la distribution à l'intérieur des classes est uniforme. Cette hypothèse permet de justifier le fait qu'on choisisse le centre des classes comme représentant.*

28.2. Nombre de classes. *En combien de classes partageons nous les valeurs ? la réponse n'est pas unique.*
Soit N l'effectif total. Nous pouvons considérer dans ce cours trois réponses à titre d'exemple. Nous rappelons maintenant la définition de l'étendu.
De plus, dans le cas continue nous parlons aussi du pas ou de la longueur de la classe.

28.3. Effectif et fréquence d'une classe.

DÉFINITION 1.12. *La quantité*

$$n_i := \text{Card}\{\omega \in \Omega : x(\omega) \in c_i\}$$

s'appelle effectif partiel de c_i .

DÉFINITION 1.13. *Le nombre*

$$f_i := \frac{n_i}{N}$$

est appelé la fréquence partielle de c_i .

29. Travaux Dirigés

EXERCICE 1.1. *La variable statistique "couleur de maisons d'un quartier" est-elle :*

O qualitative O quantitative

O discrète O continue

La variable statistique "revenu brut" est-elle :

O qualitative O quantitative

O discrète O continue

La variable statistique "nombre de maisons vendues par ville" est-elle :

O qualitative O quantitative

O discrète O continue

Solution :

Pour le premier cas, la variable statistique est qualitative. Pour le deuxième cas, la variable statistique est quantitative continue. Pour le troisième cas, la variable statistique est quantitative discrète.

EXERCICE 1.2. *Parmi ces assertions, préciser celles qui sont vraies, celles qui sont fausses.*

1/ On appelle variable, une caractéristique que l'on étudie.

2/ La tâche de la statistique descriptive est de recueillir des données.

3/ La tâche de la statistique descriptive est de présenter les données sous forme de tableaux, de graphiques et d'indicateurs statistiques.

4/ En Statistique, on classe les variables selon différents types.

5/ Les valeurs des variables sont aussi appelées modalités.

6/ Pour une variable qualitative, chaque individu statistique ne peut avoir qu'une seule modalité.

7/ Pour faire des traitements statistiques, il arrive qu'on transforme une variable quantitative en variable qualitative.

8/ La variable quantitative poids d'automobile peut être reclassé en compacte, intermédiaire et grosse.

9/ En pratique, lorsqu'une variable quantitative discrète prend un grand nombre de valeurs distinctes, on la traite comme continue.

Solution :

Le corrigé en ordre est donné par

1/ VRAI

2/ FAUX

3/ VRAI

4/ VRAI

5/ VRAI

6/ VRAI

7/ VRAI

8/ VRAI

9. VRAI

EXERCICE 1.3. *Proposer des exemples de variable quantitative transformée en variable qualitative. Préciser les modalités de cette dernière.*

Solution :

Les variables quantitatives dans le tableau ci-dessous peuvent être transformées en variable qualitative. Les modalités de cette dernière sont précisées dans la seconde colonne.

<i>Variable quantitative</i>	<i>Modalités envisageables</i>
<i>Hauteur</i>	<i>Petit, Moyen, Grand</i>
<i>Poids</i>	<i>Très léger, Léger, Moyen, Lourd, Très lourd</i>
<i>Rendement</i>	<i>Faible, Moyen, Élevé</i>
<i>Chiffre d'affaire Modéré</i>	<i>Moyen, Important, Très important</i>
<i>Cylindrée</i>	<i>Petite, Moyenne, Grosse</i>

EXERCISE 1.28. *Préciser la nature des caractères suivants :*

Âge, Couleur des yeux, Durée de vie, Sexe , Taille, Revenu annuel , Le poids, Nationalité , Niveau d'instruction, Vitesse, La langue maternelle , Commune de résidence , Groupe sanguin, La pointure des chaussures , Nombre d'enfants à charge.

Solution :

<i>Variable quantitative</i>	<i>Variable qualitative</i>
<i>Âge</i>	<i>Couleur des yeux, Sexe, Grand</i>
<i>Durée de vie</i>	<i>Très léger, Léger, Moyen, Lourd, Très lourd</i>
<i>Taille, Revenu annuel , Le poids,</i>	<i>Nationalité , Niveau d'instruction</i>
<i>Vitesse</i>	<i>Commune de résidence , Groupe sanguin</i>
<i>La pointure des chaussures, Nombre d'enfants à charge</i>	<i>Petite, Moyenne, Grosse</i>

EXERCICE 1.4. *Le tableau suivant donne la répartition selon le groupe sanguin de 40 individus pris au hasard dans une population,*

Groupes sanguins	A	B	AB	O
L'effectif	20	10	n_3	5

- 1/ Déterminer la variable statistique et son type.
- 2/ Déterminer l'effectif des personnes ayant un groupe sanguin AB.
- 3/ Donner toutes les représentations graphiques possibles de cette distribution.

Solution

1/ La population dans cette étude est les 40 personnes. Donc $N = 40$.

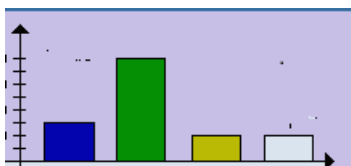
La variable statistique est le groupe sanguin des individus et elle est qualitative.

2/ L'effectif total est égal à 40. Par conséquent,

$$N = 40 = \sum_{i=1}^4 n_i = n_1 + n_2 + n_3 + n_4$$

Alors, $20 + 10 + n_3 + 5 = 40$. Ce qui implique que $n_3 = 5$.

3- Nous avons deux représentations possibles "Tuyaux d'orgue" et "Diagramme en secteur".



EXERCICE 1.5. Le gérant d'un magasin vendant des articles de consommation courante a relevé pour un article particulier qui semble connaître une très forte popularité, le nombre d'articles vendus par jour. Son relevé a porté sur les ventes des mois de Mars et Avril, ce qui correspond à 52 jours de vente. Le relevé des observations se présente comme suit :

7 13 8 10 9 12 10 8 9 10 6 14 7 15 9 11 12 11 12 5 14 11 8 10 14 12 8
5 7 13 12 16 11 9 11 11 12 12 15 14 5 14 9 9 14 13 11 10 11 12 9 15.

- 1/ Quel type est la variable statistique étudiée.
- 2/ Déterminer le tableau statistique en fonction des effectifs, des fréquences, des effectifs cumulés et des fréquences cumulés.
- 3/ Tracer le diagramme des bâtonnés associé à la variable X .
- 4/ Soit F_x la fonction de répartition. Déterminer F_x .
- 5/ Calculer le mode M_o et la moyenne arithmétique $\bar{x}$.
- 6/ Déterminer à partir du tableau puis à partir du graphe, la valeur de la médiane Me .
- 7/ Calculer la variance et l'écart-type.

Solution

1/ La population est les 52 jours et la variable statistique étudiée est le nombre d'articles vendus par jour. Son type est bien évidemment quantitatif discret (nombre).

2/ Le tableau statistique est donné par

x_i	5	6	7	8	9	10
n_i	3	1	3	4	7	5

11	12	13	14	15	16	Σ
8	8	3	6	3	1	$N=52$

3/ L'élaboration du diagramme des bâtonnets de x .

4/ La fonction de répartition est donnée par

$$F_x(x) = \begin{cases} 0, & \text{si } x < 5, \\ \frac{3}{52} & \text{si } 5 \leq x < 6, \\ \frac{4}{52} & \text{si } 6 \leq x < 7. \\ \frac{7}{52} & \text{si } 7 \leq x < 8. \\ \ddots & \ddots \\ 1 & \text{si } x \geq 16. \end{cases}$$

5/ Le mode est la valeur de la variable qui a le plus grand effectif, c'est à dire, $n_i = 8$. Donc, $Mo = 11$ et $Mo = 12$.

La moyenne arithmétique est : $\bar{X} = 10,67$.

6/ La médiane est la valeur de la variable qui divise la population de la série statistique en deux parties égales. Nous avons $Me = 11$.

7/ $Var(x) = 7.64$

EXERCISE 1.29. Pour les sujets d'étude qui suivent, spécifier : l'unité statistique, la variable statistique et son type,

1/ Étude du temps de validité des lampes électriques.

2/ Étude de l'absentéisme des ouvriers, en jours, dans une usine.

3/ Répartition des étudiants d'une promotion selon la mention obtenue sue le diplôme du Bac.

4/ On cherche à modéliser 1 le nombre de collisions impliquant deux voitures sur un ensemble de 100 intersections routières choisies au hasard dans une ville. Les données sont collectées sur une période d'un an et le nombre d'accidents pour chaque intersection est ainsi mesuré.

EXERCISE 1.30. Pour chacune des variables suivantes, préciser si elle est qualitative, quantitative discrète ou quantitative continue,

(a) Revenu annuel. (b) Citoyenneté.

- (c) Distance. (d) Taille.
 (e) Lieu de résidence. (f) Âge.
 (g) Couleur des yeux. (h) Nombre de langues parlées.

EXERCISE 1.31. Quelles formes de présentation de données correspondent ces propriétés ?

- 1/ Il donne une bonne idée des données, mais on lui préfère en général les graphiques.
 2/ Il n'est pas nécessaire de lire des nombres. D'un simple coup d'œil, on a une vision d'ensemble des données.

EXERCICE 1.6. Une enquête menée en 2005 à Batna sur la répartition selon le groupe sanguin de a donné les résultats suivants :

Groupes sanguins	A	B	AB	O
L'effectif	219	123	78	242

- 1/ Déterminer la variable statistique et son type.
 2/ Quel sont les effectifs partiels ?
 3/ Quel est l'effectif total ?
 4/ Donner toutes les représentations graphiques possibles de cette distribution.

EXERCICE 1.7. Pour déterminer le type de logement ($F_2, F_3, \dots$) à construire, on étudie 20 familles selon leur nombre d'enfants. Durant l'expérience, on note les résultats suivants :

1, 3, 5, 5, 3, 2, 4, 4, 7, 0, 2, 4, 3, 7, 0, 5, 4, 2, 3, 2

- 1/ Déterminer, la population, l'unité (individu), la variable statistique et les modalités.
 2/ Déterminer le tableau statistique avec x_i, n_i, f_i et F_i .

Modalité X	0	1	2	3	4	5	6	7
n_i	.	.	.	.	.	.	.	.
N_i	.	.	.	.	.	.	.	.
f_i	.	.	.	.	.	.	.	.
F_i	.	.	.	.	.	.	.	.

EXERCICE 1.8. Au poste de péage, on compte le nombre de voitures se présentant sur une période de 5min. Sur 100 observations de 5min, on obtient les résultats suivants :

<i>Nombre de voitures</i>	1	2	3	4	5	6	7	8	9	10	11	12
<i>Nombre d'observations</i>	2	8	14	20	19	15	9	6	2	3	1	1

1/ Construire la table des fréquences et le diagramme en bâtons en fréquences de la série du nombre de voitures.

2/ Calculer la moyenne et l'écart-type de cette série.

3/ Déterminer la médiane.

EXERCICE 1.9. Dans une petite localité, on a relevé le nombre de pièces par appartement :

<i>Nombre de pièces</i>	1	2	3	4	5	6	7
<i>Nombre d'appartements</i>	48	72	96	64	39	25	3

Le "nombre de pièces par appartement" est à considérer comme une variable aléatoire discrète à valeurs entières.

(A l'interprétation, il faudra préciser que les "demi pièces" ne sont pas comptabilisées).

Déterminer le tableau statistique.

Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.

Calculer la moyenne et l'écart-type de cette série.

Déterminer la médiane.

EXERCICE 1.10. Une machine coupe des barres de 12cm. Mais malheureusement, elle n'est pas bien réglée et les longueurs varient autour de la valeur attendue.

Une étude sur 185 barres donnent les résultats suivants :

<i>Longueurs en cm</i>	11.5	11.6	11.7	11.8	11.9	12	12.1	12.2	12.3
<i>Effectif</i>	3	15	16	16	18	20	25	25	28

1/ Quel type est la variable statistique étudiée.

2/ Déterminer le tableau statistique.

3/ Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.

4/ Calculer la moyenne et l'écart type.

Représentation numérique des données.

Introduction

Dans les cours précédents nous avons vu comment représenter les séries statistiques sous forme de tableaux et de graphiques. Nous allons associer maintenant à chaque distribution divers nombres, appelés caractéristiques qui indiquent le centre de la distribution, son degré de dispersion, son degré de symétrie,...etc. Les indicateurs statistiques de tendance centrale (dits aussi de position) considérés fréquemment sont la médiane, le mode et la moyenne.

On considère plusieurs types de caractéristiques pour une distribution :

- ♣ Les caractéristiques de tendance centrale ou de position.
- ♣ Les caractéristiques de disposition.
- ♣ Les caractéristiques de tendance centrale sont des valeurs qu'on retrouve au centre des observations, elle nous indiquent l'ordre de grandeur des données.

1. Médiane

On appelle médiane la valeur Me de la v.s.x qui vérifie la relation suivante :

$$F_x(Me) < 0.5 \leq F_x(Me+) = F_x(Me).$$

La médiane partage la série statistique en deux groupes de même effectif.

EXEMPLE 2.1. Dans l'exemple précédent, la relation

$$F_x(0) = 0 < 0.5 \leq F_x(0^+) = 0.09$$

n'est pas satisfaite. Donc, la médiane est différente de 0. Par contre, nous avons

$$F_x(2) = 0.25 < 0.5 \leq F_x(2^+) = F(2) = 0.58.$$

Donc, $Me = 2$.

1.1. Détermination de la médiane dans le cas d'une variable discrète : On doit d'abord ordonner la série statistique dans l'ordre croissant :

$x(1) \leq x(2) \leq \dots \leq x(n-1) \leq x(n)$ telle que :

$x(1)$ la plus petite valeur observée.

$x(n)$ la plus grande valeur observée.

Par la suite on peut avoir deux cas :

♦^{1^{er}} cas : Si n est un nombre impair alors on peut écrire $n = 2p + 1$ avec p un entier.

Dans ce cas la médiane est égale à : $Me = x(p + 1)$ où $x(p + 1)$ est l'observation numéro $(p + 1)$ dans l'ordre croissant.

EXEMPLE 2.2. Dans cet exemple on a :

♦ Les notes obtenues par quinze étudiants au module de mathématiques sont les suivantes :

14, 16, 8, 10, 17, 11, 10, 8, 13, 14, 9, 7, 15, 10, 12.

Ordonnons les observations dans l'ordre croissant

$\underbrace{7, 8, 8, 9, 10, 10, 10}_{11}, \underbrace{12, 13, 14, 14, 15, 16, 17}_{10}$

$n = 15$ est un nombre impair alors $n = 2 \times 7 + 1$.

Donc $p = 7 \Rightarrow Me = x(p + 1) = x(7 + 1) = x(8) = 11$

La valeur qui partage les observations en deux parties égales est la médiane $Me = 11$.

EXEMPLE 2.3. On considère la même variable, mais observé sur un échantillon de 121 étudiants :

x_i	6	8	9	10	12	13	15	Total
n_i	10	20	15	30	25	11	10	121

Ordonnons la série statistique dans l'ordre croissant :

$\underbrace{6 \dots 6}_{10} \underbrace{8 \dots 8}_{20} \underbrace{9 \dots 9}_{15} \underbrace{10 \dots 10}_{30} \underbrace{12 \dots 12}_{25} \underbrace{13 \dots 13}_{11} \underbrace{15 \dots 15}_{10}$

$n = 121 \Rightarrow n = 2 \times 60 + 1$, donc $p = 60$, alors $Me = x(p + 1) = x(60 + 1) = x(61)$.

L'observation $x(75) = 10$, ceci implique $Me = x(61) = 10$.

♦^{2^{eme}} cas : Si n est un nombre pair alors, on peut écrire $n = 2p$ avec p un entier.

Dans ce cas la médiane est égale à : $Me = \frac{x(p) + x(p+1)}{2}$ où $x(p)$ est l'observation numéro (p) dans l'ordre croissant et $x(p + 1)$ est l'observation numéro $(p + 1)$ dans l'ordre croissant.

EXEMPLE 2.4. ♦ Les notes obtenues par dix étudiants au module de physique sont les suivantes :

12, 10, 8, 9, 15, 14, 7, 15, 8.

Ordonnons les observations dans l'ordre croissant

7, 8, 8, 9, 10, 11, 14, 15, 15

$n = 10$ est un nombre pair alors $n = 2 \times 5$, avec $p = 5$.

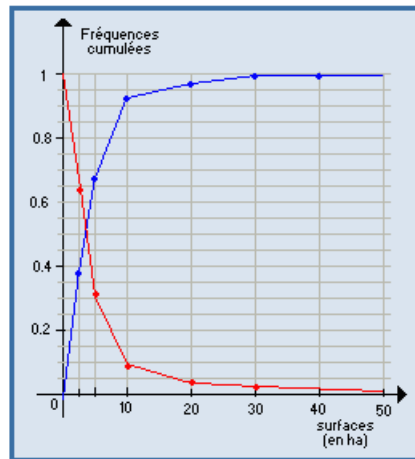
De là, la médiane égale à :

$Me = \frac{x(p)+x(p+1)}{2} = \frac{x(5)+x(5+1)}{2} = \frac{x(5)+x(6)}{2} = \frac{10+11}{2} = 10,5$. Donc 50 % des étudiants observé ont eu des notes inférieure à 10,5.

On peut aussi déterminer la médiane courbe graphiquement à partir de la courbe cumulative croissante ou décroissante.

EXEMPLE 2.5. Une enquête effectuée sur la répartition de 200 foyers selon le nombre de chambre a donné les résultats suivant :

Nombre de chambre x_i	1	2	3	4	5	6	Total
n_i	20	40	70	40	20	10	200
f_i	0,1	0,2	0,35	0,2	0,1	0,05	1
$F_{cc} \uparrow$	0	0,1	0,3	0,65	0,85	0,95	1



2. Détermination de la médiane dans le cas d'une variable continue.

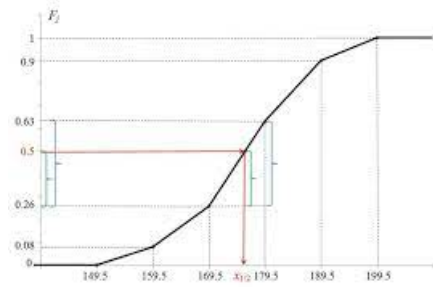
Pour déterminer la médiane d'une variable continue, on doit d'abord la localiser, c'est -à- dire déterminer la classe qui la contient et ceci à partir des fréquences cumulées de la manière suivant :

Si la valeur 0,5 est comprise entre $F(c_1)$ et $F(c_2)$ alors la médiane Me doit se trouver forcément dans la classe $[c_1, c_2[$, c'est -à- dire.

$$F(c_1) \leq 0,5 < F(c_2) \Rightarrow Me \in [c_1, c_2[.$$

Cette classe est appelée classe médiane. Une fois que la médiane est localisée, on utilise la méthode d'interpolation linéaire pour trouver une approche : On a

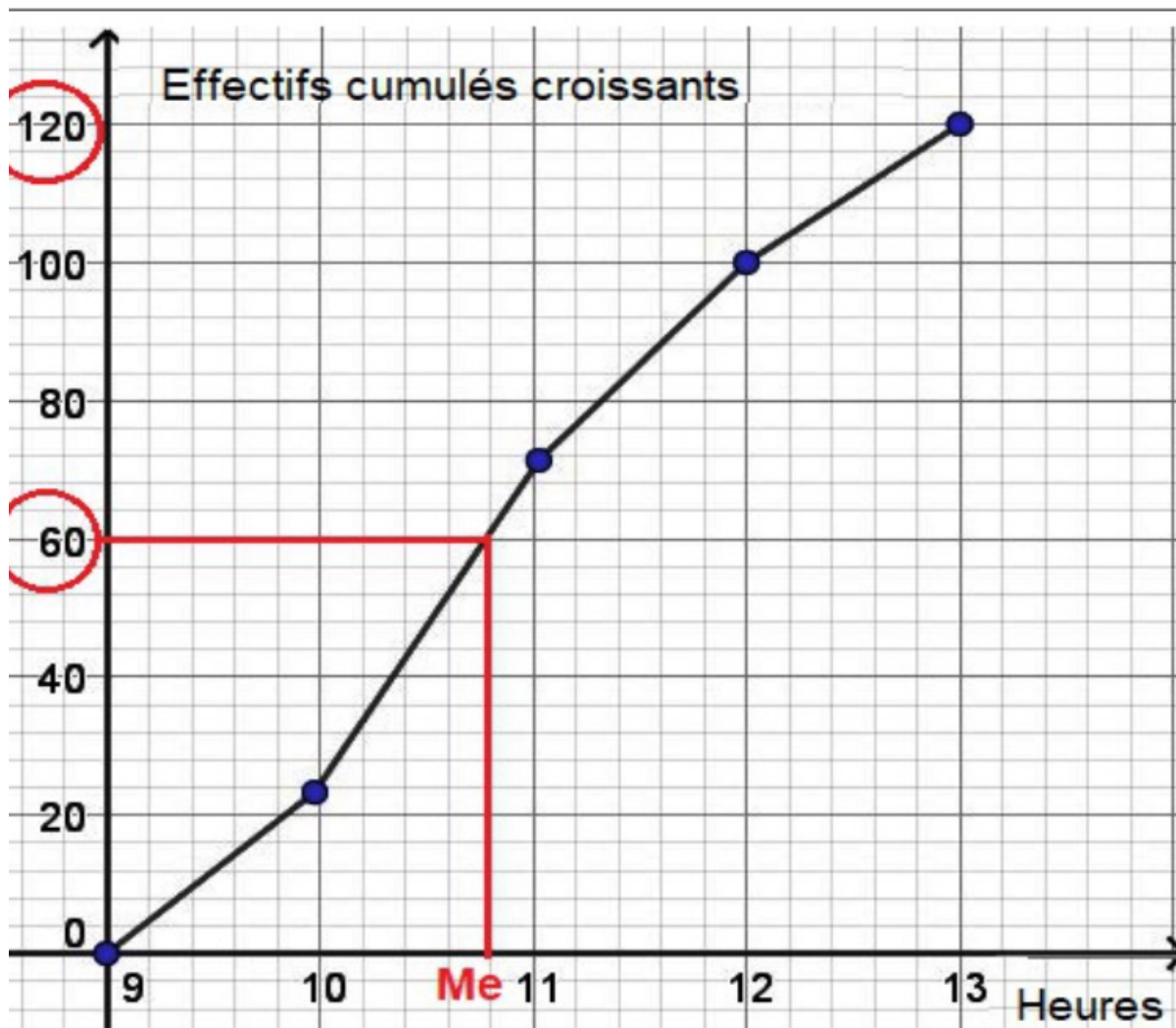
$$\begin{aligned} \alpha_1 = \alpha_2 \Rightarrow \alpha_1 = \alpha_2 &\Rightarrow \frac{0,5 - F(c_1)}{Me - c_1} = \frac{F(c_2) - F(c_1)}{c_2 - c_1} \\ \Rightarrow (c_2 - c_1)(0,5 - F(c_1)) &= (F(c_2) - F(c_1))(Me - c_1) \end{aligned}$$



$$\Rightarrow (Me - c_1) = (c_2 - c_1) \frac{0,5 - F(c_1)}{F(c_2) - F(c_1)}$$

$$Me = c_1 + (c_2 - c_1) \frac{0,5 - F(c_1)}{F(c_2) - F(c_1)}.$$

Pour plus des détails on a :



EXEMPLE 2.6. Un échantillon de 100 mesures (en mm) de pièces de rechange, fabriquées dans une entreprise a donné les résultats suivants :

Mesures	$[40,45[$	$[45,50[$	$[50,55[$	$[55,60[$	$[60,65[$	Total
Effectifs n_i	30	10	20	25	15	100
fréquences f_i	0,3	0,1	0,2	0,25	0,15	1
$f_i \uparrow$	0	0,3	0,4	0,6	0,85	1

On a $F(50) \leq 0.5 < F(55) \Rightarrow Me \in [50, 55[$, donc la classe médiane est la classe $[50, 55[$.

De Là

$$Me = c_1 + (c_2 - c_1) \frac{0,5 - F(c_1)}{F(c_2) - F(c_1)}.$$

$$Me = 50 + (55 - 50) \frac{0,5 - 0,4}{0,6 - 0,4}.$$

$$Me = 50 + 5 \times \frac{0,1}{0,2} = 52,5.$$

D'après la courbe, la valeur $F(Me) = 0,5$ correspond à l'observation 52,5, alors la médiane $Me = 52,5$.

On peut aussi déterminer la médiane graphiquement

EXEMPLE 2.7. Dans cet exemple on a une classe de 25 élèves, on note la taille des élèves dans un tableau en les regroupant par classe a doonné les résultats suivants :

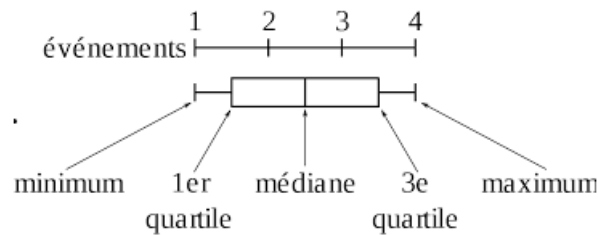
Mesures	$[0,150[$	$[150,160[$	$[160,170[$	$[170,180[$	$[180,190[$
n_i	0	3	11	8	3
$f_i \uparrow$	0	0,12	0,56	0,88	1

D'après la courbe, la valeur $F(Me) = 0,5$ correspond à l'observation 165,2 alors la médiane $Me = 165,2$.

3. Quartiles

Les quartiles sont les valeurs du caractère qui partagent l'effectif total en quatre parties égales.

- Le premier quartile Q_1 , est la plus petite valeur du caractère pour laquelle 25 % des valeurs de la série statistique lui sont inférieures ou égales.
- Le troisième quartile Q_3 est la plus petite valeur du caractère pour laquelle 75 % des valeurs de la série statistique lui sont inférieures ou égales. Il y a donc trois quartiles.



C'est la généralisation de la notion de la médiane, il s'agit des valeurs qui partagent les observations en quatre parties égales :

Le premier quartile noté Q_1 , tel que $F(Q_1) = 0,25$.

Le deuxième quartile noté Q_2 , tel que $F(Q_2) = 0,50$.

Le troisième quartile noté Q_3 , tel que $F(Q_3) = 0,75$.

4. Calcul des quartiles dans le cas d'une variable discrète

Par analogie au calcul de la médiane vu dans la partie précédente, on d'abord ordonner dans l'ordre croissant, deux cas peuvent se présenter :

♦ 1^{er} cas : Si n est un nombre pair alors on peut écrire $n = 2 \times p$ avec p un entier.

Dans ce cas

$$Q_1 = \frac{x(\frac{p}{2}) + x(\frac{p}{2} + 1)}{2}$$

$$Q_2 = Me = \frac{x(p) + x(p + 1)}{2}$$

$$Q_3 = \frac{x(\frac{3p}{2}) + x(\frac{3p}{2} + 1)}{2}$$

♦ 2^{me} cas : Si n est un nombre impair alors on peut écrire $n = 2 \times p + 1$ avec p un entier. Dans ce cas

$$Q_1 = x(\frac{p}{2} + 1)$$

$$Q_2 = Me = x(p + 1)$$

$$Q_3 = x(\frac{3p}{2} + 1).$$

Pour expliquer bien ces notions, on a l'exemple suivant :

EXEMPLE 2.8. Soit la série statistique suivante :

30; 30; 30; 45; 45; 45; 50; 50; 60; 60; 61; 61; 61.

Ici n est impair ($n = 13$), donc la médiane M est la septième valeur, c'est-à-dire 50.
Donc $M = 50$.

On a donc : $\underbrace{30; 30; 30; 45; 45; 45}_{\text{série inférieure}}; \underset{\substack{\uparrow \\ M}}{50}; \underbrace{50; 60; 60; 61; 61; 61}_{\text{série supérieure}}$

- Q_1 est la médiane de la série inférieure :
 $\underbrace{30; 30; 30}_{\text{médiane } 30}; \underbrace{45; 45; 45}_{\text{médiane } 45}$
 Ici, Q_1 est le centre de l'intervalle médian $[30; 45]$.
 Donc $Q_1 = 37,5$.
- Q_3 est la médiane de la série supérieure :
 $\underbrace{50; 60; 60}_{\text{médiane } 60}; \underbrace{61; 61; 61}_{\text{médiane } 61}$
 Ici, Q_3 est le centre de l'intervalle médian $[60; 61]$.
 Donc $Q_3 = 60,5$.

5. Détermination des quartiles dans le cas d'une variable continue

La méthode d'interpolation utilisée pour le calcul de la médiane est appliquée dans ce cas pour déterminer les différents quartiles.

Nous généralisons la notion de la médiane dans la définition suivante.

DÉFINITION 2.1. Pour $i \in \{1, 2, 3\}$ la quantité Q_i tel que $F(Q_i) = \frac{i}{4}$ s'appelle le i ème quartile. Supposons que les différents quartiles sont localisées respectivement dans les classes $]c_1, c_2[$, $]c_2, c_3[$, et $]c_3, c_4[$. alors

$$Q_1 = c_1 + (c_2 - c_1) \frac{0,25 - F(c_1)}{F(c_2) - F(c_1)}.$$

$$Q_2 = Me = c_2 + (c_3 - c_2) \frac{0,5 - F(c_2)}{F(c_3) - F(c_2)}.$$

$$Q_3 = c_3 + (c_4 - c_3) \frac{0,75 - F(c_3)}{F(c_4) - F(c_3)}.$$

EXEMPLE 2.9. Alors on a : $Q_1 = 164$, $Q_2 = 169$, $Q_3 = 176$.

D'après l'exemple précédant, on peut donner la remarque suivante :

REMARQUE 2.1. Les quartiles peuvent être déterminés aussi graphiquement à l'aide de la courbe cumulative et on peut les présenter avec la boîte moustache suivante :

6. Intervalle inter-quartile.

L'intervalle inter-quartile est la différence entre les valeurs du troisième et du premier quartile : $Q_3 - Q_1$. Autrement dit, c'est l'intervalle qui regroupe 50 % des observations, en laissant 25 % à droite et 25 % à gauche.

EXEMPLE 2.10. *Les notes au module de biologie obtenues par 100 étudiants sont données dans le tableau suivant :*

Notes	$[0,5[$	$[5,10[$	$[10,15[$	$[15,20[$	Total
Effectifs n_i	15	20	45	20	100
fréquences f_i	0,15	0,2	0,45	0,2	1
$f_i \uparrow$	0	0,15	0,35	0,8	1

Le premier quartile Q_1 : (localisation d'abord)

$$\underbrace{F(5)}_{=0,15} \leq 0,25 \leq \underbrace{F(10)}_{=0,35} \Rightarrow Q_1 \in [5, 10[, \text{ de là}$$

$$Q_1 = c_1 + (c_2 - c_1) \frac{0,25 - F(c_1)}{F(c_2) - F(c_1)} = 5 + (10 - 5) \frac{0,25 - 0,15}{0,35 - 0,15} = 7,5.$$

C'est -à-dire, 25 % des étudiants ont obtenu des notes à 7,5.

Le deuxième quartile :

$$\underbrace{F(10)}_{=0,35} \leq 0,5 \leq \underbrace{F(15)}_{=0,8} \Rightarrow Q_2 \in [10, 15[, \text{ alors}$$

$$Q_2 = 10 + (15 - 10) \frac{0,5 - 0,35}{0,8 - 0,35} = 11,67.$$

C'est -à-dire, 50 % des étudiants ont obtenu des notes à 11,67.

Le troisième quartile :

$$\underbrace{F(10)}_{=0,35} \leq 0,75 \leq \underbrace{F(15)}_{=0,8} \Rightarrow Q_3 \in [10, 15[, \text{ alors}$$

$$Q_3 = 10 + (15 - 10) \frac{0,75 - 0,35}{0,8 - 0,35} = 14,44.$$

C'est -à-dire, 75 % des étudiants ont obtenu des notes à 14,44.

L'intervalle inter quartile est donc

$$Q_3 - Q_1 = 14,44 - 7,5 = 6,94.$$

REMARQUE 2.2. *Dans cette remarque, on a :*

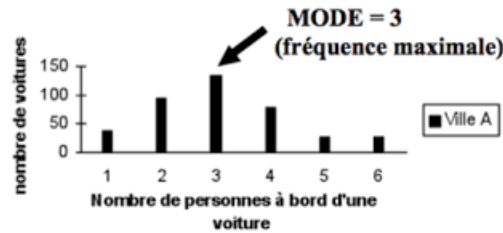
La notion de quartile peut être généralisée aux déciles qui sont au nombre de 9, ils partagent les observations en dix parties égale. Le premier décimale par exemple est D_1 , tel que :

$$F(D_1) = 0,1.$$

7. Mode

8. Détermination de le mode dans le cas d'une variable discrète et continue :

Le mode d'une $v.s$ est la valeur qui a le plus grand effectif partiel (ou la plus grande fréquence partielle) et il est dénoté par M_0 .



Si x est une variable continue, on parle de classe modale.

EXEMPLE 2.11. Une enquête effectuée sur la répartition de 200 foyers selon le nombre de chambre a donné les résultats suivant :

Nombre de chambre	1	2	3	4	5	6	Total
Effectifs	20	40	70	40	20	10	200

Dans l'exemple précédent, le mode est égal à $M_0 = 3$ qui correspond au plus grand effectif 70.

EXEMPLE 2.12. Une enquête sur l'âge des élèves à l'examen du bac a donnée les résultats suivants :

âge(année)	16	17	18	19	20	Total
Effectifs	100	400	200	180	120	1000

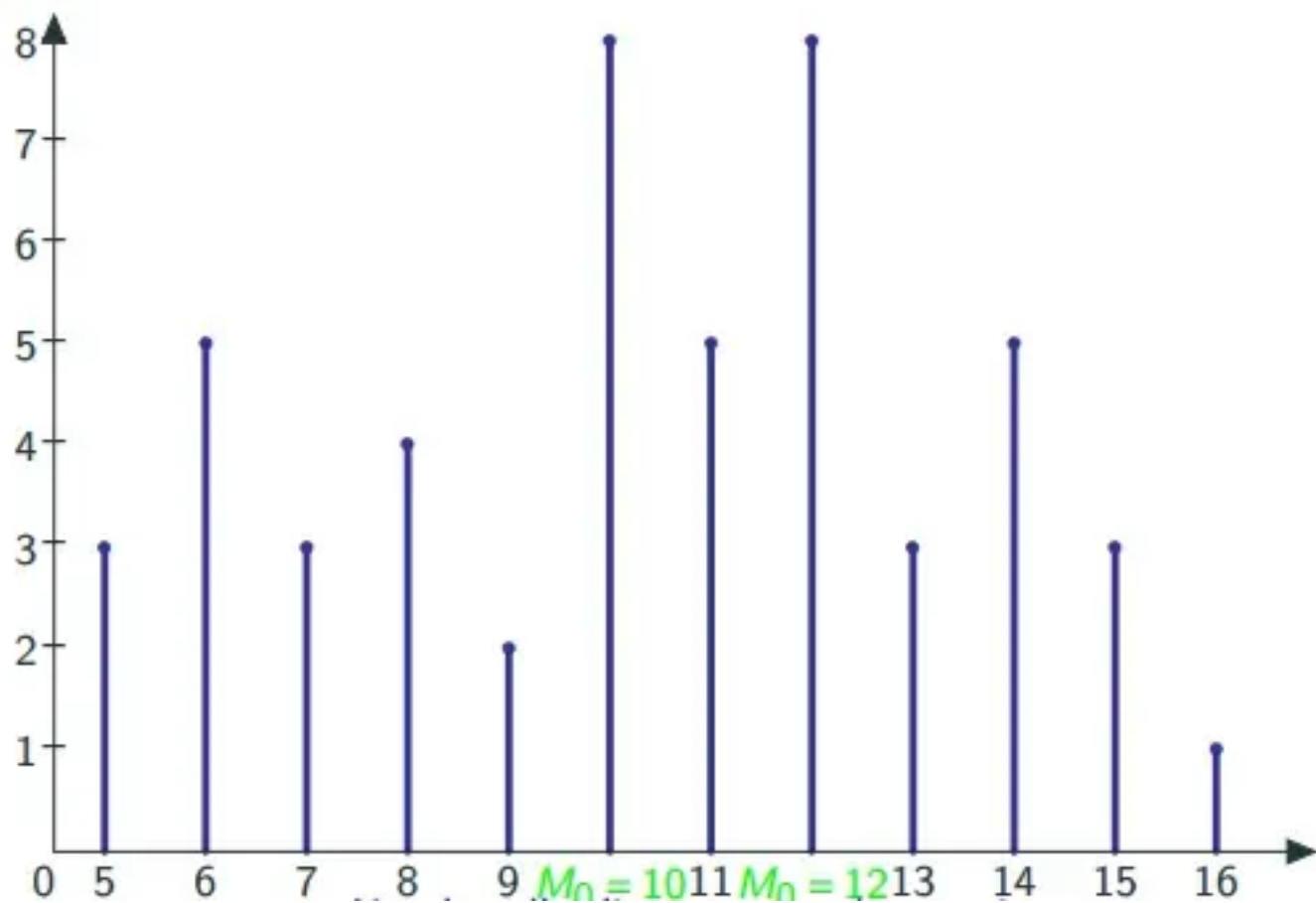
L'effectif la plus élevée est 400, il correspond donc au mode $M_0 = 17$.

REMARQUE 2.3. On a trois remarques importantes :

- Une variable statistique peut avoir plusieurs modes, dans le cas où l'effectif la plus élevée correspond à plusieurs observations.
- On dit aussi qu'une variable ne possède pas de mode si tous les effectifs partiels sont identiques.
- Dans le cas d'une variable continue, le mode ne peut pas être trouvé d'une façon exacte, par contre on peut calculer une valeur approchée une fois qu'on a la classe modale qui est à l'effectif le plus élevé si les classes sont des amplitudes variables on utilise les effectifs rectifié

ou la quantité $\frac{f_i}{a_i}$.

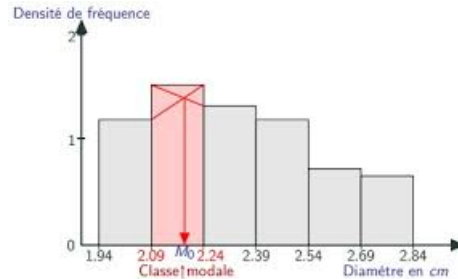
Effectif



9. Détermination de le mode dans le cas d'une variable continue :

Supposons que la classe modale est la classe $[c_1, c_2[$ comprise entre les classes $[c_0, c_1[$ et $[c_2, c_3[$ de même amplitudes.

Une valeur approchée du mode est donnée par la relation suivante :



$$M_0 = c_1 + (c_2 - c_1) \frac{(n_3 - n_1)}{(n_3 - n_1) + (n_3 - n_2)}.$$

EXEMPLE 2.13. Soit le tableau donnant les poids (en kg) de 100 individus :

Poids en kg	$[56, 58[$	$[58, 60[$	$[60, 62[$	$[62, 64[$	$[64, 66[$	Total
Effectifs n_i	5	36	40	15	4	100

On remarque que toutes les classes ont la même amplitude, donc la classe modale est celle qui a l'effectif le plus élevé, qui est la classe $[60, 62[$.

Une valeur approchée du mode est :

$$M_0 = c_1 + (c_2 - c_1) \frac{(n_3 - n_1)}{(n_3 - n_1) + (n_3 - n_2)}.$$

$$M_0 = 60 + (62 - 60) \frac{(40 - 36)}{(40 - 36) + (40 - 15)}.$$

$$M_0 = 60 + (2) \times \left(\frac{4}{4 + 25} \right) = 60,27.$$

10. Moyenne géométrique

La moyenne géométrique simple G d'une série de n variable $x_1, x_2, \dots, x_k$ est la racine n^{ime} du de ces valeurs, avec $\sum_{i=0}^k n_i = n$:

$$G = \sqrt[n]{x_1 \times x_2 \times \dots \times x_k}$$

$$G = x_1 \times x_2 \times \dots \times x_k$$

En prenant le logarithme des deux membre on obtient :

$$G = \frac{1}{n} \sum_{i=0}^k \log x_i$$

$$G = \sum_{i=0}^k \log x_i$$

Si on a une variable statistique ayant k valeurs $x_1 \times x_2 \times \dots \times x_k$ aux quelles correspond respectivement les effectifs $n_1 \times n_2 \times \dots \times n_k$, alors la moyenne géométrique pondérée de ces valeur est :

$$G = \sqrt[n]{x_1^{n_1} \times x_2^{n_2} \times \dots \times x_k^{n_k}}$$

$$G = x_1^{f_1} \times x_2^{f_2} \times \dots \times x_k^{f_k}$$

Ou encore

$$G = \frac{1}{n} \sum_{i=0}^k n_i \log x_i$$

$$G = \sum_{i=0}^k f_i \log x_i$$

EXEMPLE 2.14. *la moyenne géométrique des valeurs 2, 3, 8, 11, 17 est*

$$G = \sqrt[5]{x_1 \times x_2 \times \dots \times x_k}$$

$$G = \sqrt[5]{2 \times 3 \times (8) \times 11 \times 17} = 6,17$$

11. Moyenne harmonique

La moyenne harmonique simple H d'une série de n valeurs $x_1, x_2, \dots, x_k$ est l'inverse de la moyenne arithmétique des inverses des valeurs :

$$H = \frac{1}{\frac{1}{n} \sum_{i=0}^k \frac{1}{x_i}} = \frac{1}{\sum_{i=0}^k \frac{n}{x_i}}$$

Si on a une variable statistique ayant k valeurs $x_1 \times x_2 \times \dots \times x_k$ aux quelles correspond respectivement les effectifs $n_1 \times n_2 \times \dots \times n_k$, alors la moyenne harmonique pondérée de ces valeur est :

$$H = \frac{1}{\frac{1}{n} \sum_{i=0}^k \frac{n_i}{x_i}} = \frac{n}{\sum_{i=0}^k \frac{n_i}{x_i}}.$$

EXEMPLE 2.15. la moyenne harmonique des valeurs 1, 4, 8, 10, 12 est

$$H = \frac{n}{\sum_{i=0}^k \frac{n_i}{x_i}}.$$

$$H = \frac{5}{\frac{1}{1} + \frac{1}{4} + \frac{1}{8} + \frac{1}{10} + \frac{1}{12}} = 3,2$$

12. Moyenne quadratique

La moyenne quadratique simple M_Q d'une série de n valeurs $x_1, x_2, \dots, x_k$, avec $\sum_{i=0}^k n_i = n$ est définie par :

$$M_Q = \sqrt{\bar{x^2}} = \sqrt{\frac{1}{n} \sum_{i=0}^k x_i^2}$$

Si on a une variable statistique ayant k valeurs $x_1 \times x_2 \times \dots \times x_k$ aux quelles correspond respectivement les effectifs $n_1 \times n_2 \times \dots \times n_k$, alors la moyenne quadratique pondérée de ces valeur est :

$$M_Q = \sqrt{\frac{1}{n} \sum_{i=0}^k n_i \times \bar{x}_i^2}$$

avec $\sum_{i=0}^k n_i = n$.

EXEMPLE 2.16. la moyenne quadratique des valeurs 7, 8, 10, 11, 14 est

$$M_Q = \sqrt{\frac{1}{n} \sum_{i=0}^k n_i \times \bar{x}_i^2}$$

$$M_Q = \sqrt{\frac{1}{5}[(7)^2 + (8)^2 + (10)^2 + (11)^2 + (14)^2]} = 10,29.$$

REMARQUE 2.4. **1-** Dans cette remarque, on donne un exemple de calculer la moyenne dans le cas continue.

Notes	[0 ; 5[	[5 ; 8[	[8 ; 12[	[12 ; 15[	[15 ; 20]
Effectif	10	8	12	11	9
Centre de la classe (x_i)	2,5	6,5	10	13,5	17,5

$$\begin{aligned}\bar{x} &= [(2,5 \times 10) + (6,5 \times 8) + (10 \times 12) + (13,5 \times 11) + (17,5 \times 9)] / 50 \\ &= 503/50 \\ &= 10,06\end{aligned}$$

2-Dans cette remarque, on donne la relation entre les moyennes :

♣ Les moyennes géométrique et harmonique sont utilisées

dans certains cas exceptionnels et la moyenne quadratique est très utilisée en physique.

♣ On a la relation suivante entre ces différentes moyennes.

$$H \leq G \leq \bar{X} \leq M_Q.$$

13. Étendue

La différence entre la plus grande valeur et la plus petite valeur du caractère, donnée par la quantité

$$e = x_{\max} - x_{\min},$$

s'appelle l'étendue de la *v.s.x.* Le calcul de l'étendue est très simple. Il donne une première idée de la dispersion des observations.

EXEMPLE 2.17. Soit la série statistique suivante : 3, 6, 8, 9, 10, 14, 15, 6, 7, 11, 2, 5, 9.
l'étendue de cette série d'observations est : $15 - 2 = 13$.

DÉFINITION 2.2. Le nombre

$$e = x_{\max} - x_{\min}$$

s'appelle l'étendu de x . Dans ce cas, on peut définir le pas par

$$a_i := \frac{\text{étendu}}{\text{nombre de classes}} = \frac{x_{\max} - x_{\min}}{k}$$

14. Diagramme en boîte ou boîte à moustache

Les cinq nombres $\min, Q_1, Q_2 = Me, Q_3, \max$ permettent de résumer une série statistique quantitative. Une façon possible de représenter graphiquement ce résumé est d'utiliser un diagramme en boîte ou boîte à moustaches. La boîte à moustaches a été inventée en 1977 par **John Tukey**.



REMARQUE 2.5. *John Wilder Tukey était un mathématicien et statisticien américain, né le 16 juin 1915 à New Bedford dans le Massachussetts et décédé le 26 juillet 2000 à New Brunswick dans le New Jersey.*

La boîte à moustaches est aussi quelques fois appelée diagramme de Tukey ou boîte à pattes ou box plot.



Dessiné au-dessus d'un axe, un diagramme en boîte est constitué :

- D'une boîte délimitée par les premiers et troisièmes quartiles et partagée par la médiane.
- De deux moustaches qui relient les quartiles aux valeurs extrêmes de la série. La hauteur de la boîte est arbitraire.
- L'objectif ou la superposition de boîtes à moustaches peut être utile pour comparer plusieurs séries associées à un même caractère sur des populations différentes.

15. Variance

On appelle variance de cette série statistique, le nombre

$$Var(x) = \frac{1}{N} \sum_{i=0}^k n_i (\bar{X} - x_i)^2$$

$$Var(x) = \sum_{i=0}^k f_i (\bar{X} - x_i)^2$$

On dit que la variance est la moyenne des carrés des écarts à la moyenne $\bar{X}$. Les écarts à la moyenne sont les $(\bar{X} - x_i)$, les carrés des écarts à la moyenne sont donc les $(\bar{X} - x_i)^2$. En faisant la moyenne de ces écarts, on trouve la variance.

En statistiques et en principe des probabilités, le théorème de König-Huygens est une identité remarquable reliant la variance et la moyenne. Le théorème suivant (Théorème de König-Huygens) donne une identité remarquable reliant la variance et la moyenne, parfois plus pratique dans le calcul de la variance.

THÉORÈME 2.18. *Soit (x_i, n_i) une série statistique de moyenne $\bar{X}$ et de variance $Var(x)$. Alors,*

$$Var(x) = \sum_{i=0}^k f_i x_i^2 - \bar{X}^2$$

$$Var(x) = \frac{1}{N} \sum_{i=0}^k n_i x_i^2 - \bar{X}^2$$

DÉMONSTRATION. Par définition, nous avons

$$Var(x) = \sum_{i=0}^k f_i (\bar{X} - x_i)^2 = \frac{1}{N} \sum_{i=0}^k n_i (\bar{X} - x_i)^2 = \frac{\sum_{i=0}^k n_i (\bar{X} - x_i)^2}{\sum_{i=0}^k n_i}$$

Donc

$$Var(x) = \frac{\sum_{i=0}^k n_i (\bar{X} - x_i)^2}{\sum_{i=0}^k n_i} = \frac{\sum_{i=0}^k n_i (\bar{X}^2 - 2\bar{X}x_i + x_i^2)}{\sum_{i=0}^k n_i}$$

Par égalité, nous avons

$$Var(x) = \frac{\sum_{i=0}^k n_i \bar{X}^2}{\sum_{i=0}^k n_i} - \frac{\sum_{i=0}^k 2n_i \bar{X}x_i}{\sum_{i=0}^k n_i} + \frac{\sum_{i=0}^k n_i x_i^2}{\sum_{i=0}^k n_i}$$

Ce qui implique que

$$Var(x) = \bar{X}^2 - 2 \frac{\sum_{i=0}^k \bar{X}x_i}{\sum_{i=0}^k n_i} - 2\bar{X}^2 \bar{X}^2$$

□

REMARQUE 2.6. • *En pratique, il est préférable d'utiliser la formule du théorème précédent, car elle rend les calculs plus facile.*

- *Dans le cas d'une variable continue la variance est la quantité*

$$\text{Var}(x) = \sum_{i=1}^k f_i(\bar{X} - c_i)^2,$$

où c_i = centre de la classe i .

EXEMPLE 2.19. *Soit la variable statistique dont la distribution est donnée par le tableau suivant :*

x_i	1	2	3	4	5	Total
n_i	10	25	40	15	10	100

Calculons d'abord la moyenne $\bar{X}$:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^k n_i \times x_i$$

$$\bar{X} = \frac{1}{100}[(10 \times 1) + (25 \times 2) + (40 \times 3) + (15 \times 4) + (10 \times 5)] = 2,9.$$

La variance $\text{Var}(x)$ est donc égale à :

$$\text{Var}(x) = \frac{1}{N} \sum_{i=0}^k n_i x_i^2 - \bar{X}^2$$

$$\text{Var}(x) = \frac{1}{100}[10(1)^2 + 25(2)^2 + 40(3)^2 + 15(4)^2 + 10(5)^2] - 2,9^2$$

$$\text{Var}(x) = 1,19$$

16. Écart type

La quantité $\sigma_X = \sqrt{\text{Var}(x)}$ s'appelle l'écart type de la variable statistique x .

REMARQUE 2.7. *Le paramètre σ_X mesure la distance moyenne entre $\bar{X}$ et les valeurs de x . Il sert à mesurer la dispersion d'une série statistique autour de sa moyenne.*

Plus il est petit, plus les caractères sont concentrés autour de la moyenne (on dit que la série est homogène).

Plus il est grand, plus les caractères sont dispersés autour de la moyenne (on dit que la série est hétérogène).

EXEMPLE 2.20. *L'écart type de la distribution étudiée dans l'exemple précédent est :*

$$\sigma_X = \sqrt{\text{Var}(x)} = \sqrt{1,19} = 1,09.$$

17. Écart absolu moyen.

Soit la variable statistique x , ayant k modalités $x_1, x_2, \dots, x_k$, aux quelles correspondent respectivement les effectifs $n_1, n_2, \dots, n_k$.

L'écart absolu moyen de la variable statistique x la moyenne arithmétique des valeurs absolues des écarts à la moyenne arithmétique :

$$e = \frac{1}{N} \sum_{i=1}^N n_i |x_i - \bar{X}|$$

EXEMPLE 2.21. On donne la série des observations $x_i : 5, 12, 13, 14, 2, 7, 18, 4, 10, 9$ la moyenne arithmétique de cette variable statistique est :

$$\bar{X} = \frac{1}{10} \times (5 + 12 + 13 + 14 + 2 + 7 + 18 + 4 + 10 + 9) = 9,4.$$

L'écart absolu moyen est donc égale à :

$$e = \frac{1}{10} [|5-9,4| + |12-9,4| + |13-9,4| + |14-9,4| + |2-9,4| + |7-9,4| + |18-9,4| + |10-9,4| + |9-9,4|] = 4.$$

18. Coefficient de variation

L'écart type et la moyenne s'expriment dans la même unité que la variable statistique, mais dans certain cas on peut être ramené à comparer les dispersions de distribution qu'elles ne sont pas exprimées dans la même unité.

Pour cela on utilise le coefficient de variation qui défini par le rapport de l'écart type à la moyenne :

$$C_v = \frac{\sigma_X}{\bar{X}}$$

REMARQUE 2.8. Le coefficient de variation est souvent exprimé sous la forme d'un pourcentage.

EXEMPLE 2.22. La distribution des salaires observés dans deux entreprises ; une du sud et l'autre du nord est donnée par les tableaux suivants :

Salaires mensuels (DA)	[20000,30000[	[30000,35000[	[35000,40000[	[40000,50000[	Total
Effectifs n_i	30	20	40	10	100
Centre c_i	25000	32500	37500	45000	/

$$\bar{X}_1 = \frac{1}{N} \sum_{i=1}^k n_i \times c_i$$

$$\bar{X}_1 = \frac{1}{100} \sum_{i=1}^4 [30 \times 25000 + 20 \times 32500 + 40 \times 37500 + 10 \times 45000] = 33500.$$

$$Var(x_1) = \sum_{i=0}^N f_i (\bar{X} - c_i)^2 = 141500000 \Rightarrow \sigma_{X1} = 11895,38.$$

Distribution des salaires dans l'usine du nord :

<i>Salaires mensuels (DA)</i>	<i>[10000,15000[</i>	<i>[15000,20000[</i>	<i>[20000,25000[</i>	<i>[25000,40000[</i>	<i>Total</i>
<i>Effectifs n_i</i>	<i>40</i>	<i>30</i>	<i>20</i>	<i>10</i>	<i>100</i>
<i>Centre c_i</i>	<i>12500</i>	<i>17500</i>	<i>22500</i>	<i>32500</i>	<i>/</i>

$$\bar{X}_2 = \frac{1}{N} \sum_{i=1}^k n_i \times c_i$$

$$\bar{X}_2 = \frac{1}{100} \sum_{i=1}^4 [40 \times 12500 + 30 \times 17500 + 20 \times 22500 + 10 \times 32500] = 18000.$$

$$Var(x_2) = \sum_{i=0}^N f_i (\bar{X} - c_i)^2 = 37250000 \Rightarrow \sigma_{X_2} = 6103,28.$$

Les coefficients de variations sont :

$$C_{v1} = \frac{\sigma_X}{\bar{X}_1} = \frac{11895,38}{33500} = 0,355.$$

$$C_{v2} = \frac{\sigma_X}{\bar{X}_2} = \frac{6103,28}{18000} = 0,339.$$

La dispersion relative des salaires est presque la même dans les deux entreprises.

19. Comparaison des différents paramètres de position :

I La médiane est plus sensible aux fluctuations d'échantillonnage, elle l'est moins à des valeurs aberrantes. Toutefois, elle se prête moins bien à des calculs algébriques.

II Le mode est représentatif de la valeur du caractère le plus courant, le plus typique, mais il peut présenter une certaine ambiguïté. La comparaison des trois permet de se faire une idée plus complète de la distribution.

III La moyenne arithmétique est peu sensible aux fluctuations d'échantillonnage. Elle se prête bien aux comparaisons. Des valeurs aberrantes peuvent toutefois la modifier sensiblement.

Finalement Si les trois sont à peu près égales alors la série statistique est à peu près symétrique).

20. Travaux Dirigés

EXERCICE 2.1. Les 33 élèves d'une classe ont obtenu les notes suivantes lors d'un devoir :

Note	2	4	5	8	10	11	12	14	15	18	20
Effectif	1	2	1	4	2	7	6	3	4	2	1

- 1) Déterminer l'étendue et le mode de cette série.
- 2) Calculer la moyenne de cette série.
- 3) Construire un tableau donnant les effectifs cumulés, les fréquences et les fréquences cumulées.
- 4) Déterminer la médiane de cette série.
- 5) Quel est le nombre d'élèves ayant une note strictement inférieure à 8 ?
- 6) Quel est le pourcentage d'élèves ayant une note supérieure ou égale à 10 ?

Solution :

1) L'étendue de cette série est la différence entre les valeurs extrêmes de la série. Elle vaut ici $20 - 2 = 18$.

Le mode de cette série est la valeur du caractère correspondant à l'effectif maximum. Il vaut ici 11

2) La moyenne de cette série statistique est égale à arrondi au centième.

3) Les fréquences sont égales au quotient entre les effectifs et l'effectif total.

Note	2	4	5	8	10	11	12	14	15	18	20
Effectif	1	2	1	4	2	7	6	3	4	2	1
Effectif cumulés croissants	1	3	4	8	10	17	23	26	30	32	33
Fréquences	1	2	1	4	2	7	6	3	4	2	1
Fréquences cumulées croissantes	1	2	1	4	2	7	6	3	4	2	1

(Remarque : la dernière ligne peut être obtenue par quotient des effectifs cumulés et de l'effectif total)

4) La médiane d'une série ordonnée de 33 valeurs est égale à 17^{me} valeur.

D'après le tableau dressé en question (2),

10 élèves ont une note inférieure ou égale à 10.

Tandis que 17 élèves ont une note inférieure ou égale à 11.

La note du 17^{me} élève se situe donc parmi les 7 notes égales à 11.

La médiane de cette série statistique est donc égale à 11.

EXERCICE 2.2. Répartition du nombre de supermarchés en France suivant la surface en m^2 :

Surface	$[400; 800[$	$[800; 1000[$	$[1000; 2500]$
Effectif	2613	928	3379

- 1) Déterminer la surface moyenne $\bar{X}$ d'après ce regroupement par classe.
- 2) Sachant que la surface totale de vente est de $6739000 m^2$, calculer la surface moyenne d'un supermarché. Comparer avec la valeur obtenue à la question (1).

Solution :

- 1) Pour déterminer la surface moyenne $\bar{X}$, il faut considérer le milieu de chaque intervalle. On obtient le tableau :

Surface	600	900	1750
Effectif	2613	928	3379

On calcule $\bar{X} = \frac{600 \times 2613 + 900 \times 928 + 1750 \times 3379}{2613 + 928 + 3379} \approx 1201,77 m^2$.

- 2) Il y a au total $2613 + 928 + 3379 = 6920$ supermarchés.

Si la surface totale de vente est de $6739000 m^2$, la surface moyenne d'un supermarché est égale à $\frac{6739000}{6920} = 973,84$.

Cette dernière valeur est bien inférieure à celle obtenue à la question (1).

EXERCICE 2.3. Une loterie a été organisée avec des gains en argent liquide. Tous les billets n'ont pas été vendus. Le tableau ci-dessous résume les gains effectivement perçus par les joueurs :

Gain (en euros)	100	200	300	400	500	600	700	800	900	1000
Effectif	2	1	1	3	2	2	3	5	0	1

Partie A : Analyse de la série statistique

- 1) Combien y a-t-il de gagnants à cette loterie ? (personne n'a gagné plus d'une fois)
- 2) Quel a été le gain moyen parmi les gagnants ?
- 3) a) Quelle est la médiane de cette série statistique ? Quels sont les quartiles ?
- b) Déterminer l'écart inter-quartile.
- 4) Faire un diagramme en boîte à moustaches de la série.
- 5) Calculer l'écart type de la série.

Partie B : Augmentation des gains

L'association qui organise la loterie envisage une augmentation des gains.

6) La première hypothèse envisagée consiste à augmenter tous les gains de 217 euros.

Dans ce cas, comment varient :

a) La moyenne ? b) L'écart type ? c) La médiane ?

7) La deuxième hypothèse envisagée consiste à multiplier tous les gains par 1,2.

Dans ce cas, comment varient :

a) La moyenne ? b) L'écart type ? c) La médiane ?

Solution :

Partie A : Analyse de la série statistique

1) Le nombre de gagnants à cette loterie est égal à $2 + 1 + 1 + 3 + 2 + 2 + 3 + 5 + 0 + 1 = 20$.

Il y a 20 gagnants à cette loterie.

2) Le gain moyen parmi les gagnants vaut :

$$\bar{X} = \frac{2 \times 100 + 200 + 300 + 3 \times 400 + 2 \times 500 + 2 \times 600 + 3 \times 700 + 5 \times 800 + 1000}{20} = \frac{11200}{20} = 560 \text{ euros.}$$

3) a) La médiane d'une série statistique ordonnée de 20 valeurs est égale à la demi somme entre la 10^{me} et la 11^{me} valeur.

Puisque les 10^{me} et 11^{me} valeur valent toutes les deux 600 euros, on en déduit que la médiane de cette série statistique vaut donc 600 euros.

Le premier quartile Q_1 est la plus petite valeur de la série telle qu'au moins 25 % des valeurs de la série lui sont inférieures ou égales. 25 % des 20 valeurs de la série représentent 5 valeurs. 7 valeurs sur 20 sont inférieures ou égales à 400 euros, donc $Q_1 = 400$.

Le troisième quartile Q_3 est la plus petite valeur de la série telle qu'au moins 75 % des valeurs de la série lui sont inférieures ou égales. 75 % des 20 valeurs de la série représentent 15 valeurs. 19 valeurs sur 20 sont inférieures ou égales à 800 euros, donc $Q_3 = 800$

b) L'écart inter-quartile vaut $Q_3 - Q_1 = 800 - 400 = 400$.

4) Le diagramme en boîte à moustaches de la série est donné ci-dessous :

5) D'après la formule de la variance, on a :

$$\text{Var}(x) = \frac{1}{N} \sum_{i=0}^k n_i (\bar{X} - x_i)^2$$

$$\text{Var}(x) = \frac{2 \times 100^2 + 200^2 + 300^2 + 3 \times 400^2 + 2 \times 500^2 + 2 \times 600^2 + 3 \times 700^2 + 5 \times 800^2 + 1000^2}{20} - 560^2 = 62400.$$

On en déduit que l'écart-type de la série vaut $\sigma_X = \sqrt{\text{Var}(x)} = \sqrt{62400} \approx 250$

Partie B : Augmentation des gains

6) Si l'association augmente tous les gains de 217 euros, sans changer les effectifs :

a) La moyenne s'en trouve également augmentée de 217 euros, et devient donc égale à $560 + 217 = 777$ euros.

b) L'écart type est inchangé car la dispersion des notes autour de la moyenne n'est pas modifiée

par cette augmentation.

c) La médiane s'en trouve également augmentée de 217 euros, et devient donc égale à $600 + 217 = 817$ euros.

7) Si l'association multiplie tous les gains par 1,2, sans changer les effectifs :

a) La moyenne s'en trouve également multipliée par 1,2 et devient donc égale à

$$560 \times 1,2 = 672$$

b) L'écart s'en trouve également multipliée par 1,2 et devient donc environ égal à

$$250 \times 1,2 = 300$$

c) La médiane s'en trouve également multipliée par 1,2, et devient donc égale à

$$600 \times 1,2 = 720.$$

EXERCICE 2.4. On effectue un contrôle de la qualité pendant 100 heures de travail sur deux machines produisant des pièces mécaniques destinées à la fabrication de grues.

Certaines pièces présentent un défaut qui les rend inutilisables.

On a relevé le nombre de pièces inutilisables constatées durant chaque heure :

Machine A :

Nombre de pièces inutilisables	0	1	2	3	4	5	6	7
Nombres d'heures	13	42	38	2	2	1	1	1

Machine B :

Nombre de pièces inutilisables	0	1	2	3	4	5
Nombres d'heures	35	40	1	1	10	13

1) a) Calculer le nombre moyen $\bar{X}_A$ de pièces inutilisables pendant les 100 heures étudiées pour la machine A. Calculer ensuite la variance $\text{Var}A$.

b) Calculer le nombre moyen $\bar{X}_B$ de pièces inutilisables pendant les 100 heures étudiées pour la machine B. Calculer ensuite la variance $\text{Var}B$.

2) a) Déterminer la médiane, puis l'écart inter-quartile dans le cas de la machine A. Calculer l'étendue.

b) Déterminer la médiane, puis l'écart inter-quartile dans le cas de la machine B. Calculer l'étendue.

3) a) Parmi la moyenne, l'écart type, la médiane, l'écart inter-quartile ou l'étendue, quels sont les paramètres qui mesurent la dispersion ?

b) Quel(s) paramètre(s) semble(nt) le(s) plus intéressant(s) à exploiter pour comparer ces deux machines ? Justifier.

EXERCICE 2.5. Une machine coupe des barres de 12cm. Mais malheureusement, elle n'est pas bien réglée et les longueurs varient autour de la valeur attendue. Une étude sur 185 barres donnent les résultats suivants :

Longueurs en cm	11.5	11.6	11.7	11.8	11.9	12	12.1	12.2	12.3
Effectif	3	15	16	16	18	20	25	25	28

- 1/ Quel type est la variable statistique étudiée.
- 2/ Déterminer le tableau statistique.
- 3/ Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.
- 4/ Calculer la moyenne et l'écart type.

EXERCICE 2.6. Dans une petite localité, on a relevé le nombre de pièces par appartement :

Nombre de pièces	1	2	3	4	5	6	7
Nombre d'appartements	48	72	96	64	39	25	3

Le nombre de pièces par appartement est à considérer comme une variable aléatoire discrète à valeurs entières.

(A l'interprétation, il faudra préciser que les demi pièces ne sont pas comptabilisées).

- 1/ Déterminer le tableau statistique.
- 2/ Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.
- Calculer la moyenne et l'écart-type de cette série.
- 3/ Déterminer la médiane.

EXERCICE 2.7. Au poste de péage, on compte le nombre de voitures se présentant sur une période de 5min. Sur 100 observations de 5min, on obtient les résultats suivants :

Nombre de voitures	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'observations	2	8	14	20	19	15	9	6	2	3	1	1

- 1/ Construire la table des fréquences et le diagramme en bâtons en fréquences de la série du nombre de voitures.
- 2/ Calculer la moyenne et l'écart-type de cette série.
- 3/ Déterminer la médiane.

EXERCICE 2.8. Dans une petite localité, on a relevé le nombre de pièces par appartement :

Nombre de pièces	1	2	3	4	5	6	7
Nombre d'appartements	48	72	96	64	39	25	3

Le nombre de pièces par appartement est à considérer comme une variable aléatoire discrète à valeurs entières. (A l'interprétation, il faudra préciser que les demi pièces ne sont pas comptabilisées).

1/ Déterminer le tableau statistique.

2/ Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.

3/ Calculer la moyenne et l'écart-type de cette série.

4/ Déterminer la médiane.

EXERCICE 2.9. Une machine coupe des barres de 12cm. Mais malheureusement, elle n'est pas bien réglée et les longueurs varient autour de la valeur attendue. Une étude sur 185 barres donnent les résultats suivants :

Longueurs en cm	11.5	11.6	11.7	11.8	11.9	12	12.1	12.2	12.3
Effectif	3	15	16	16	18	20	25	25	28

1/ Quel type est la variable statistique étudiée.

2/ Déterminer le tableau statistique.

3/ Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.

4/ Calculer la moyenne et l'écart type.

EXERCICE 2.10. Un enquêteur a noté le prix en euro d'une même marchandise dans dix points de vente différents :

14, 2; 13, 8; 14, 2; 13, 9; 14; 14, 1; 13, 8; 14, 3; 15, 2; 13, 5.

1/ Donne un prix médian de cette série.

2/ Calcule le prix moyen.

3/ Calcule l'étendue de cette série.

EXERCICE 2.11. complétez le tableau suivant :

Classes	$[10,20[$	$[20,30[$	$[30,40[$	$[40,60[$	$[60,80[$	Total
Centres des classes	.	.	.	.	.	.
fréquences f_i	0,08	0,21	.	.	.	.
$f_i \uparrow$	.	.	0,55	0,86	.	.
$f_i \downarrow$	.	.	.	.	.	.

EXERCICE 2.12. On a relevé le nombre de garçons dans deux cents familles de quatre enfants. On a obtenu le tableau suivants :

x_i	0	1	2	3	4	Total
n_i	8	42	67	70	13	200

- 1/ Définir le caractère étudié et donner sa nature.
- 2/ Que représentent x_i, n_i .
- 3/ Représenter graphiquement ces données.
- 4/ Calculer les fréquences et les fréquences cumulées.
- 5/ Représenter la courbe cumulative.
- 6/ Tracer le polygone des effectifs.

EXERCICE 2.13. On a relevé les nombres d'allumettes contenues respectivement dans 20 boîtes, lors d'un contrôle dans une usine de fabrication. Les résultats sont les suivants :

40, 42, 32, 38, 40, 48, 30, 38, 36, 40, 34, 40, 34, 40, 38, 40, 42, 44, 36, 42.

- 1/ Déterminer la moyenne, l'écart-type, le mode et la médiane cette série statistique ainsi que Q_1 et Q_3 :
- 2/ Ranger ces résultats en classes d'intervalle 4 allumettes.
- 3/ Tracer l'histogramme de cette distribution et la courbe des fréquences cumulées.
- 4/ Calculer la moyenne, la médiane de cette distribution (les comparer avec celles obtenues en 1) ainsi que l'écart-type.

EXERCICE 2.14. La distribution suivante représente la taille (exprimée en centimètres) de 40 étudiants d'une classe :

175 165 180 152 161 152 155 158 173 154

178 179 160 178 180 167 172 152 177 180

175 154 157 159 169 169 170 167 153 169

171 172 158 172 162 157 158 163 168 164

En utilisant la règle de **STURGES** ranger ces observations dans un tableau statistique.

1/ Déterminer la moyenne, l'écart-type, le mode et la médiane cette série statistique ainsi que Q_1 et Q_3 .

EXERCICE 2.15. Les notes attribuées, au cours d'une épreuve orale, à 200 candidats se distribuent selon les données suivantes :

Notes (x_i)	5	6	7	8	9	10	11	12	13
Nombres d'observations (n_i)	8	15	28	35	45	36	20	11	2

- 1/ Construire le diagramme en bâtons et la courbe cumulative associés à cette distribution.
- 2/ Déterminer le mode, la moyenne arithmétique, l'écart type.
- 3/ Déterminer les moyennes harmonique, géométrique et quadratique et la médiane.

EXERCICE 2.16. L'observation des salaires horaires, exprimés en euros, dans une entreprise de 100 salariés donne les résultats suivants :

Salaires	[56 ; 58[	[58 ; 60[	[60 ; 62[	[62 ; 64[	[64 ; 66[	Total
Effectifs	5	36	40	15	4	100

- 1/ Construire l'histogramme et la courbe cumulative.
- 2/ Déterminer : la classe modale, le mode, la moyenne, la médiane.
- 3/ Déterminer : écart inter- quartile, étendue, écart-type.

EXERCICE 2.17. Le tableau suivant donne la distribution de 53500 des exploitations agricoles en fonction de la surface agricole utile en hectares :

Surface agricole(hectares)	[0 ; 5[	[5 ; 10[	[10 ; 20[	[20 ; 50[	[50 ; 100[
Nombre d'exploitations n_i	9000	11000	24500	6000	3000

- 1/ Représente graphiquement cette série statistique.
- 2/ Déterminer la classe modale.
- 3/ Calculer la moyenne des exploitations.
- 4/ Quel est le pourcentage d'exploitations dont la surface agricole utile :
 - Est inférieure à 10 hectares.
 - Est comprise dans l'intervalle [5; 20[.
- 5/ Calculer l'écart type absolu moyen.

EXERCICE 2.18. Une machine coupe des barres de 12cm. Mais malheureusement, elle n'est pas bien réglée et les longueurs varient autour de la valeur attendue. Une étude sur 185 barres

donnent les résultats suivants :

Longueurs en cm	11.5	11.6	11.7	11.8	11.9	12	12.1	12.2	12.3
Effectif	3	15	16	16	18	20	25	25	28

- 1 Quel type est la variable statistique étudiée.
- 2 Déterminer le tableau statistique.
- 3 Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.
- 4 Calculer la moyenne et l'écart type.

EXERCICE 2.19. Le magasin Supertanker fait la liste des capacités des disques durs en Go des ordinateurs qu'il propose à la vente. Les résultats sont résumés dans le tableau suivant :

Capacité en Go	10	20	50	80	160	250	320	500	800	1000	1150
Effectif	2	4	5	12	10	7	2	4	1	2	1

- 1) Combien d'ordinateurs sont proposés à la vente dans ce magasin ?
- 2) a) Déterminer la médiane Me de cette série.
b) Quel est le pourcentage d'ordinateurs dont la capacité en Go est inférieure à Me ?
- 3) a) Déterminer le premier quartile Q_1 et le troisième quartile Q_3 .
b) Déterminer l'écart inter-quartile.
- 4) Faire un diagramme en boîte à moustaches sur lequel on précisera les valeurs extrêmes des capacités en Go, la médiane et les quartiles Q_1 et Q_3 .
- 5) a) L'écart type vaut environ 260Go. La moyenne vaut 225Go.
Un autre magasin, HyperTech avec la même capacité moyenne de 225Go pour les ordinateurs vendus présente un écart type de 90Go.
Que peut-on en déduire pour les ordinateurs vendus dans le magasin HyperTech ?
b) Rappeler le lien entre écart type et variance.
Donner, sans faire les calculs, une formule permettant de calculer la variance.
- 6) De manière à simplifier l'étude du stock, on regroupe les ordinateurs en 4 catégories :
"moins de 80Go", "Entre 80 et 250Go", "entre 250 et 500Go", "Plus de 500Go"
a) Recopier et compléter le tableau suivant en étant attentif aux bornes des différents intervalles :

Capacité en Go	$]0;80]$	$]80;250]$	$]250;500]$	$]500;1150]$
Effectif	23	.	.	.

b) On considère, pour calculer la moyenne, le centre de chaque classe, c'est-à-dire qu'on remplace l'intervalle par la valeur située au centre de l'intervalle.

On remplace ainsi $]0; 80]$ par 40.

Déterminer la moyenne obtenue par ce procédé. Pourquoi ne trouve-t-on pas 225Go ?

EXERCICE 2.20. Lors d'un contrôle de police sur l'autoroute, un agent a relevé les vitesses suivantes (arrondies à l'entier inférieur ou égal) :

117 134 130 113 127 125 98 110 124 122 126 101
 106 121 121 104 124 117 109 128 134 146 111 139
 123 124 130 123 120 133 111 143 145 111 110 119
 114 104 126 99 140 105 119 134 128 119 137 109
 122 130 92 104 113 130 120 84 166 138 129 119

a/ Groupez ces données par classes : $[80 - 90[$, $[90 - 100[$, etc.

b/ Dessinez le diagramme à secteurs correspondant.

c/ Calculez le mode, la médiane et l'intervalle inter-quartile.

EXERCICE 2.21. I) Soit la répartition des salaires journaliers des employés d'une usine **A** :

Salaires (102DA)	$[5, 6[$	$[6, 7[$	$[7, 8.5 [$	$[8.5, 9[$	$[9, 10[$
Nombre d'employés	120	190	240	100	50

- 1/ Déterminer la population statistique, le caractère étudié et sa nature.
- 2/ Tracer l'histogramme des effectifs et le polygone des effectifs.
- 3/ Déterminer la valeur du mode (par le calcul).
- 4/ Déterminer la valeur de la médiane graphiquement et par le calcul, interpréter le résultat.
- 5/ Déterminer l'écart inter quartile (intervalle inter-quartile).
- 6/ Quel est le nombre d'employés qui perçoivent un salaire compris entre 600 et 800 DA par jour.

II) Soit la répartition des salaires journaliers des 620 employés d'une usine **B** :

Salaires (102DA)	$[5, 6[$	$[6, 7[$	$[7, ?[$	$[?, 9[$	$[9, 10[$
Nombre d'employés	100	80	240	160	?
Centre de classe	.	.	7,25	.	.

- 1/ Compléter le tableau.
- 2/ Évaluer la dispersion relative des salaires de chacune des usines. En déduire laquelle des

usines possède les salaires les plus homogènes (les moins dispersés).

III) Le syndicat des employés de l'usine **A** a demandé une augmentation des salaires de 16 pour chaque employé. Quel sera le montant du salaire journalier moyen dans le cas de cette augmentation.

EXERCICE 2.22. La production d'un bien **Y** était de 1820 unités en 2005.

1/ Sachant que la production du bien **Y** a connu une augmentation de 3 entre 2004 et 2005, calculer la production de 2004.

2/ La production du bien **Y** a connu les taux de croissance successifs suivants :

Périodes	2 ans	1 ans	1 ans	3 ans
Taux de croissance	3	6	-3	2

Calculer le taux de croissance annuel moyen.

EXERCICE 2.23. Choisissez la (les) bonne (s) réponse (s) :

1) Quand les amplitudes sont inégales, pour dessiner l'histogramme :

1/ On calcule l'étendu.

2/ On corrige les effectifs.

3/ On calcule les effectifs cumulés.

4/ On ne corrige pas les effectifs.

2) Pour calculer la moyenne arithmétique dans le cas d'un caractère quantitatif continu :

1/ On calcule les effectifs corrigés.

2/ On calcule les centres de classe.

3/ On calcule les effectifs cumulés.

4/ On calcule l'étendu.

3) L'intervalle (l'écart) inter quartile :

1/ Contient 50 des observations.

2/ Est égal à $Q_3 - Q_1$.

3/ Contient 20 des observations.

4/ Est égal à $Q_1 - Q_3$.

4) La représentation graphique correspondant au cas quantitatif continu est :

1/ L'histogramme.

2/ Le diagramme en bâtons.

3/ La courbe en escalier.

4/ La courbe cumulative.

EXERCICE 2.24. Dans une bibliothèque, on a relevé le nombre de livres prêtés par mois durant l'année 2007.

Mois	J	F	M	A	M	J	J	A	S	O	N	D
Nombre de livres prêtés	1124	1236	1146	1136	1086	987	840	620	1027	1220	1128	994

- a/ Calcule le nombre total de livres prêtés en 2007.
- b/ Calcule le nombre moyen de livres prêtés par mois durant cette année.
- c/ Détermine une médiane de cette série statistique.
Donne une interprétation de la valeur obtenue.
- d/ Détermine les valeurs des premier et troisième quartiles de cette série statistique. Donne une interprétation des valeurs obtenues.
- e/ Détermine l'étendue de cette série statistique.

EXERCICE 2.25. Le diagramme en barres donne les résultats obtenus à un contrôle de mathématiques par les élèves d'une classe.

- a/ Calcule la moyenne de la classe à ce contrôle.
- b/ Détermine une note médiane. c/ Détermine les valeurs des premier et troisième quartiles de cette série de notes.

Analyse Combinatoire

1. Arrangements

Soit un ensemble de n éléments tous distincts $\{a, b, \dots, s\}$.

DÉFINITION 3.1. On appelle **arrangements avec répétition** de p éléments choisis parmi n , toute disposition ordonnée de p éléments parmi les n avec répétition d'un ou plusieurs éléments. Le nombre d'arrangements de p éléments parmi n , noté Q_n^p est égal à :

$$Q_n^p = \underbrace{n \times \dots \times n}_{p \text{ fois}}$$

EXEMPLE 3.1. 1/ Un numéro d'appel est composé de 8 chiffres, ce numéro est une disposition ordonnée avec répétition de 8 éléments parmi 10 $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. donc un numéro est un arrangement avec répétition.

Le nombre de numéros possible est $= Q_{10}^8 = 10^8 = 100.10^6$.

2/ Combien de mots différents de 4 lettres au plus, qu'on peut former avec un alphabet de (02) lettres ?

Un mot de p lettres choisi parmi deux est un arrangement avec répétition de p parmi $n = 2$ $\{a, b\}$.

- Nombre de mots d'une lettre $= Q_1^2 = 2 = (a)(b)$.
- Nombre de mots deux (02) lettres $= Q_2^2 = 2^2 = (a, a)(a, b), (b, a)(b, b)$.
- Nombre de mots deux (03) lettres $= Q_2^3 = 8$.
- Nombre de mots deux (04) lettres $= Q_2^4 = 16$.

Donc au total il y a $2 + 4 + 8 + 16 = 30$ mots formés d'au plus 4 lettres parmi 2.

REMARQUE 3.1. Le nombre d'éléments p d'un arrangement avec répétition peut être $\geq n$ mais ce n'est pas le cas d'un arrangement sans répétition.

DÉFINITION 3.2. On appelle **arrangements sans répétition** (ou simplement arrangement de p éléments choisis parmi n , une disposition ordonnée sans répétition de p éléments choisis parmi les n .

REMARQUE 3.2. • L'ensemble des arrangements sans répétition est inclus dans l'ensemble des arrangements avec répétition.

- Si A_n^p désigne le nombre d'arrangement sans répétition de p éléments parmi n alors

$$A_n^p \leq Q_n^p.$$

$$\begin{aligned} A_r^n &= \frac{n!}{(n-r)!} \quad ; 1 \leq p \leq n \\ &= n(n-1)\dots(n-(p-1)). \end{aligned}$$

EXEMPLE 3.2. 1/ Quel est le nombre de tiercés dans l'ordre, d'une course de 10 chevaux. L'ordre est important et on ne peut pas avoir de répétition, donc il s'agit d'arrangement sans répétition de trois éléments choisis parmi 10, leur est égale à

$$A_r^n = \frac{n!}{(n-r)!} = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 10 \times 9 \times 8 = 720.$$

2/ Trois joueurs lancent un dé à 6 faces.

Combien, y a-t-il de résultat possibles ?

Combien, y a-t-il de résultat où les points obtenus sont tous distincts ?

Le nombre de résultats possibles est le nombre d'arrangements avec répétition de 3 parmi 6, $Q_6^3 = 6^3 = 216$.

Le nombre de résultats différents est le nombre d'arrangements sans répétition de 3 parmi 6; $A_6^3 = \frac{6!}{3!} = 120$

2. Permutations

DÉFINITION 3.3. Soit un ensemble de n éléments $\{a, b, \dots, s\}$. on appelle **Permutation sans répétition** (ou simplement permutation) de ces n élément toutes disposition ordonnée de l'ensemble de ces n éléments.

Chacun de ces éléments figure dans un ordre déterminé, une et une seule fois. Donc une permutation est un arrangement sans répétition de n éléments choisis parmi n .

Le nombre de permutation possible de n éléments est $P_n = n! = \frac{n!}{(0-0)!}$

EXEMPLE 3.3. Trouver le nombre de configurations possible à l'arrivée d'une course de 10 chevaux.

Une arrivée est une permutation de 10 chevaux. Le nombre d'arrivée possibles est donc

$$P_{10} = 10! = 3628000.$$

Soit un ensemble de n éléments formés de k groupes différents, d'éléments identiques

$$\underbrace{(a \times \dots \times a)}_{n_1}, \underbrace{(b \times \dots \times b)}_{n_2}, \dots, \underbrace{(s \times \dots \times s)}_{n_k} \text{ tels que } n_1 + n_2 + \dots + n_k = n$$

DÉFINITION 3.4. On appelle *permutation avec répétition* de ces n éléments, une transposition ordonnée de l'ensemble de ces n éléments et dans chaque permutation avec répétition

La lettre a figure n_1 fois

La lettre b figure n_2 fois...

La lettre s figure n_k fois

Le nombre de ces permutations avec répétition est noté $P_n(n_1, n_2, \dots, n_k)$ cas où $k = 2$.

On a deux groupes $(\underbrace{a \times \dots \times a}_{n_1}), (\underbrace{b \times \dots \times b}_{n_2}), n_1 + n_2 = n$

Soit par exemple $n = 4, n_1 = n_2 = 2$ donc (a, b, b, a) est une permutation avec répétition.

Pour obtenir le nombre de permutations possibles, il faut diviser le nombre de permutations sans répétition des n_1 éléments, soit $n!$ par le nombre de permutations sans répétition des n_1 éléments a entre eux.

Puis par le nombre des permutations sans répétition des n_2 ...

Par le nombre des permutations sans répétition des n_k éléments s entre eux.

D'où $P_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1!n_2!\dots n_k!}$

EXEMPLE 3.4. 1) Trouver le nombre de permutations du nombre 11222.

$$11222 = \{ \underbrace{(1, 1)}_{n_1=2}, \underbrace{(2, 2, 2)}_{n_2=3} \} \Rightarrow P_5(2, 3) = \frac{5!}{2!3!} = 10$$

11222	21122	22121
12122	21212	22211
12212	21221	
12221	22112	

2) On jet 12 dés et on appelle **résultat** une suite ordonnée de 12 points obtenus.

Combien y a-t-il de résultats au total et parmi eux combien y a-t-il de résultats où chaque face est obtenue 2 fois.

Un résultat est un arrangement avec répétition des 12 éléments choisis parmi 6, leur nombre est $Q_6^{12} = 6^{12}$.

Un résultat où chaque face est obtenue deux fois est une permutation des 12 éléments scindés en 6 groupes d'effectifs $n_1 = n_2 \dots = n_6 = 2$ leur nombre est

$$P_{12}(2, 2, 2, 2, 2, 2) = \frac{12!}{2!2!2!2!2!2!} \text{ résultats.}$$

3. Combinaisons

Soit un ensemble de n éléments distincts $\{a, b, \dots, s\}$.

DÉFINITION 3.5. On appelle **Combinaison sans répétition** (ou plus simplement combinaison) de p éléments choisis parmi n , une disposition non ordonnée sans répétition de p éléments parmi n , avec $1 \leq p \leq n$.

On remarque qu'une combinaison engendre $P!$ arrangement sans répétition, il y a donc $P!$ fois plus d'arrangement que de combinaison. Donc si C_n^p désigne le nombre de combinaison de p parmi n alors

$$p!C_n^p = A_n^p = \frac{n!}{(n-p)!}, \text{ d'où } C_n^p = \frac{n!}{p!(n-p)!}, p \leq n.$$

REMARQUE 3.3.

- $C_n^0 = C_n^n = 1.$
- $C_n^p = C_n^{n-p}.$
- $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}.$

EXEMPLE 3.5. Trouver le nombre de tiercés dans le désordre, dans une course de 10 chevaux.

Un tiercé dans le désordre est une combinaison de 3 éléments choisis parmi 10 leur nombre est $C_{10}^3 = \frac{10!}{3!(10-3)!} = 120.$

DÉFINITION 3.6. On appelle **Combinaison avec répétition** de p éléments choisis parmi n , une disposition non ordonnée avec répétition éventuellement, de p éléments parmi les n éléments.

EXEMPLE 3.6. $\{a, b, c\}, n = 3, p = 2.$

On a 6 combinaisons avec répétition

$$(a, a), (b, b), (c, c), (a, b), (a, c), (b, c).$$

REMARQUE 3.4. P peut être supérieur à n alors que, dans le cas des combinaisons sans répétition, p est au plus égal à n .

Le nombre des combinaisons avec répétition de p éléments choisis parmi n est donné par la formule suivante : $K_n^p = C_{n+p-1}^p$

4. Probabilités

1. Introduction Le calcul des probabilités est la base essentielle pour toute approche de la théorie des probabilité et des statistiques. C'est une suite naturelle de la statistique descriptive dans laquelle on introduit de façon empirique les notions intuitive du langage de probabilité. La théorie des probabilités est considérée comme une science dont l'objet consiste à étudier des phénomènes qui dépendent du hasard.

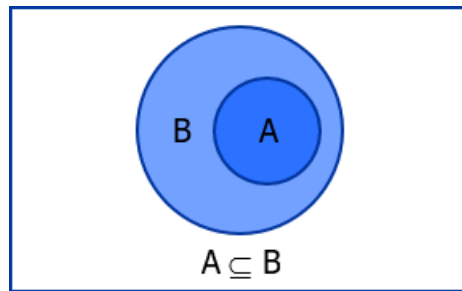
2. Rappels de résultats mathématiques.

Principaux résultats de la théorie des ensembles.

• **Opération d'inclusion :**

Soit un ensemble E , on dit que A est inclus dans E (que A est un sens ensemble de E) si tous les éléments de A appartiennent à E .

On note : $A \subset E$ Si $\forall x \in A, x \in E \Rightarrow A \subset E$



EXEMPLE 3.7. Soit E l'ensemble des élèves du lycée Mustafa Ben Boulaid et $A =$ l'ensemble des élèves de la classe de troisième année du lycée Mustafa Ben Boulaid.

A est un sous-ensemble de E .

C'est $\forall$ l'élève de la troisième année du lycée Mustafa Ben Boulaid, l'élève in au lycée Mustafa Ben Boulaid.

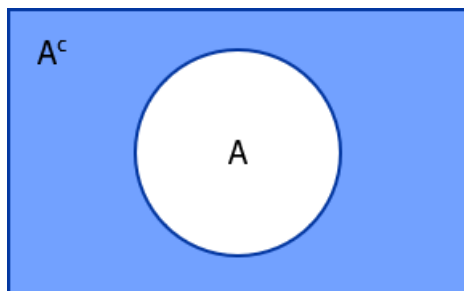
• **Opération de complémentarité : C_E^A**

Soit A un sous ensemble de E $A \subset E$. L'ensemble des éléments qui appartiennent à E sans appartenir à A s'appelle le complémentaire de A dans E , noté par C_E^A et on a :

$$\forall x \in C_E^A \Rightarrow x \notin A$$

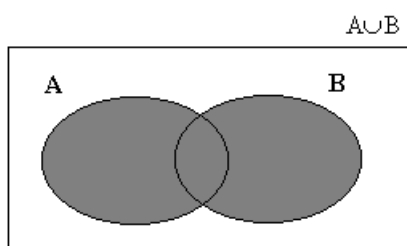
REMARQUE 3.5.

- $A \cup C_E^A = E$.
- $A \cap C_E^A = \emptyset$.
- C_E^A est noté aussi par $\bar{A}$.



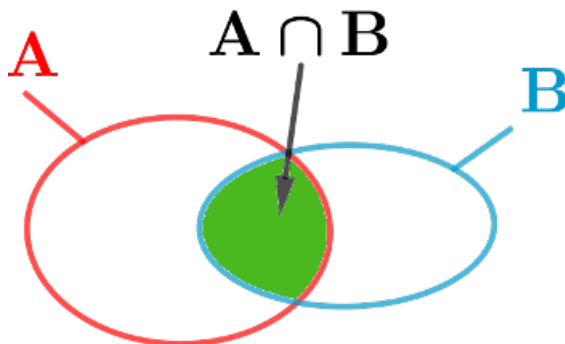
• **Opération d'union $\cup$:**

Soient A et B deux sous ensembles de E , l'ensemble des éléments appartenant à A ou à B l'union de A et B , notée $A \cup B : \forall x \in A \cup B \Rightarrow x \in A \text{ ou } x \in B$



• **Opération d'intersection $\cap$:**

Soient A et B deux sous ensembles de E , l'ensemble des éléments appartenant à A ou à B la réunion de A et B , notée $A \cap B : \forall x \in A \cap B \Rightarrow x \in A \text{ et } x \in B$.



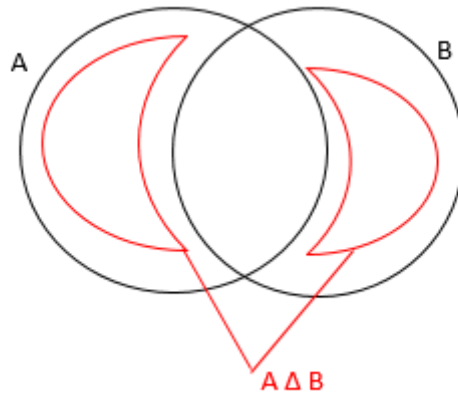
EXEMPLE 3.8. Soient $A = \{ 2, 4, 6 \}$. et $B = \{ 1, 3, 5 \}$. Alors $A \cap B = \emptyset$.

EXEMPLE 3.9. Soient A et B deux ensembles, tel que :
 $A = \{ 2, 3, 4, 6 \}$. et $B = \{ 1, 3, 5 \}$. Alors
 $A \cap B = \{ 1, 3, 5 \}$.

• **Opération de différence symétrique Δ :**

Soient A et B deux sous ensembles de E , l'ensemble des éléments appartenant à $A \cup B$ et qui

n'appartiennent pas à $A \cap B$ est la différence symétrique de A et B , notée $A \Delta B$



EXEMPLE 3.10. Dans l'exemple suivant, on a :

Soient $A = \{ 2, 4, 6 \}$. et $B = \{ 1, 3, 6 \}$. Alors

$$A \cup B = \{ 1, 2, 3, 4, 6 \}.$$

$$A \cap B = \{ 6 \}.$$

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

5. Espace probabilisable

Expérience aléatoire : Une expérience est qualifiée d'aléatoire si on ne peut pas prévoir à l'avance son résultat et répétée dans des conditions identiques, elle peut donner lieu à des résultats différents.

REMARQUE 3.6. Le résultat de l'épreuve dépend d'un ensemble de conditions dans lesquelles s'effectue cette épreuve. Ces conditions doivent être pratiquement identiques, lorsqu'on répète l'expérience afin s'assurer de la compatibilité des résultats.

L'ensemble de tous les résultats possibles d'une épreuve est dit espace fondamental ou espace des événements (appelé aussi référentiel) noté Ω .

EXEMPLE 3.11. • Le jet d'un dé peut donner l'ensemble fondamental suivant :

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

• Une pièce de monnaie :

$$\Omega = \{ \text{pile}, \text{face} \}.$$

6. Événements élémentaires et composés

- Un événements élémentaires est un sous ensemble de Ω , réduit à un seul élément.
- Un événement composé est un sous ensemble des événements élémentaires.

EXEMPLE 3.12. Pour le jet d'un dé.

1/ Avoir le chiffre 6 est un événement élémentaire.

2/ Avoir un chiffre pair est un événement composé de trois événements élémentaires : 2, 4 et 6.

7. Réalisation d'un événement

Considérons une expérience aléatoire d'ensemble fondamental Ω . Soit un événement quelconque associé à cette épreuve et considérons un résultat noté ω . Lorsque le résultat $\omega \in A$, on dira que l'événement A s'est réalisé. Dans le cas contraire $\omega \notin A$, on dira que l'événement A ne s'est pas réalisé.

8. Événement incompatibles (disjoints)

Deux événements A et B sont dits **incompatibles** ou **disjoints** si la réalisation de l'un exclue celle de l'autre, c'est à dire $A \cap B = \emptyset$.

EXEMPLE 3.13. On lance un dé et soient les événements :

A avoir un chiffre pair = $\{2, 4, 6\}$.

A avoir un chiffre impair = $\{1, 3, 5\}$.

$A \cup B = \emptyset$ donc A et B sont incompatibles.

9. Système complet d'événements

Soient $A_1, A_2, \dots, A_n$ une familles d'événements. On dit que cette famille forme un **système complet** d'événements si :

- $A_i \cap A_j = \emptyset, \quad \forall i \neq j.$
- $\bigcup_{i=1}^n A_i = \Omega.$

10. Algèbre des événements

On appelle **Algèbre** ou **tribu** sur un espace fondamental Ω , qu'on notera $\mathfrak{M}$, une famille d'événements vérifiant :

- $\Omega \in \mathfrak{M}$
- Si $A \in \mathfrak{M}$ alors $\bar{A} \in \mathfrak{M}$ où $\bar{A}$ est l'événement contraire de A .
- Si $A_i \in \mathfrak{M}, i \in I$ fini, alors $\cup_{i \in I} A_i \in \mathfrak{M}$.

Propriétés :

- $\emptyset \in \mathfrak{M}$.
- Si $A_i \in \mathfrak{M}$, $i \in I$ fini, alors $\cap_{i \in I} A_i \in \mathfrak{M}$.

EXEMPLE 3.14. Dans l'exemple suivant on a :

- La famille $\mathfrak{M} = \{\emptyset, \Omega\}$ est une algèbre sur Ω , c'est la plus petite algèbre sur Ω .)
- La famille $\mathfrak{M} = \{\emptyset, \Omega, A, \bar{A}\}$ est une algèbre sur Ω , est une algèbre de Bernoulli.

11. Calculs des Probabilités

On se limite dans ce cours à étudier les univers dénombrables. La probabilité d'un événement est une valeur numérique qui représente la proportion de fois où l'événement va se réaliser, quand on répète l'expérience dans des conditions identiques.

On peut déduire de cette définition qu'une probabilité doit être entre 0 et 1 et que la probabilité d'un événement est la somme des probabilités de chacun des événements élémentaires qui le constituent. Enfin, la somme des probabilités de tous les éléments de Ω est 1.

Important : rappelons qu'un événement n'est rien d'autre qu'une partie de Ω .

Une probabilité associe à chaque événement un nombre entre 0 et 1. Il s'agit donc d'une application de l'ensemble des parties de Ω , noté $P(\Omega)$, dans $[0, 1]$.

EXEMPLE 3.15. Soit $\Omega = \{0, 1, 2\}$. Construisons $P(\Omega)$.

$$P(\Omega) = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, \{0, 1, 2\}, \Omega\}$$

12. Espace probabilisable

Le couple $(\Omega, \mathfrak{M})$ constitué d'un ensemble fondamental Ω et une algèbre $\mathfrak{M}$ est appelé **espace probabilisable**. Les éléments de $\mathfrak{M}$ sont appelés **événements** et ceux de Ω **événements élémentaires**.

Concept de probabilité

Définir une probabilité c'est associer à chaque événement A , un poids $p(A)$ représentant la chance qu'il a de se réaliser.

a) Les différentes approches :

Il existe trois approches conceptuelles différentes des calculs de probabilités.

• L'approche classique à priori :

Consiste à déterminer tous les résultats possible, le nombre de résultats favorables n à la réalisation de l'événement A souhaité :

$$P(A) = \frac{n}{N} = \frac{\text{Nombre de cas favorable}}{\text{Nombre de cas possible}}$$

EXEMPLE 3.16. On jette une pièce de monnaie, la probabilité de tomber sur face est :

$$P(\text{face}) = \frac{1}{2}$$

Cette approche ne nécessite pas une expérience est parfois désignés d'approche à priori.

• ***L'approche à posteriori :***

Cette approche se base sur les données fournies par un certain nombre d'observations ou d'expériences antérieures.

EXEMPLE 3.17. *A priori la probabilité qu'une femme enceinte attendre un garçon est $\frac{1}{2}$. Or on constate, sur un million de naissances, qu'il y a 515000 garçons et 485000 filles.*

La probabilité à posteriori (après l'expérience) d'avoir un garçon est égal à 0,515.

Aussi bien, l'approche à priori à posteriori la probabilité, elle donnent des valeurs quantifiées ou objectives, souvent sous forme de taux de réalisation de l'événement considéré qui sera vérifiable à long terme.

b) Définition théorique :

On appelle **probabilité** sur l'espace probabilisable $(\Omega, \mathfrak{M})$, toute application p définie de $(\Omega, \mathfrak{M})$, dans l'intervalle $[0, 1]$ et vérifiant :

- $0 \leq P(\Omega) \leq 1$.

- Pour tout ensemble fini d'événements incompatibles $A_1, A_2, \dots, A_n$ on a :

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

En définissant une probabilité P sur $(\Omega, \mathfrak{M})$, on obtient un espace **probabilisé** ou **espace de probabilité** $(\Omega, \mathfrak{M}, P)$.

Dans le cas d'une σ -algèbre, la deuxième condition sera remplacé par l'axiome de σ -additivité :

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

- **Conséquences de l'axiome d'additivité :**

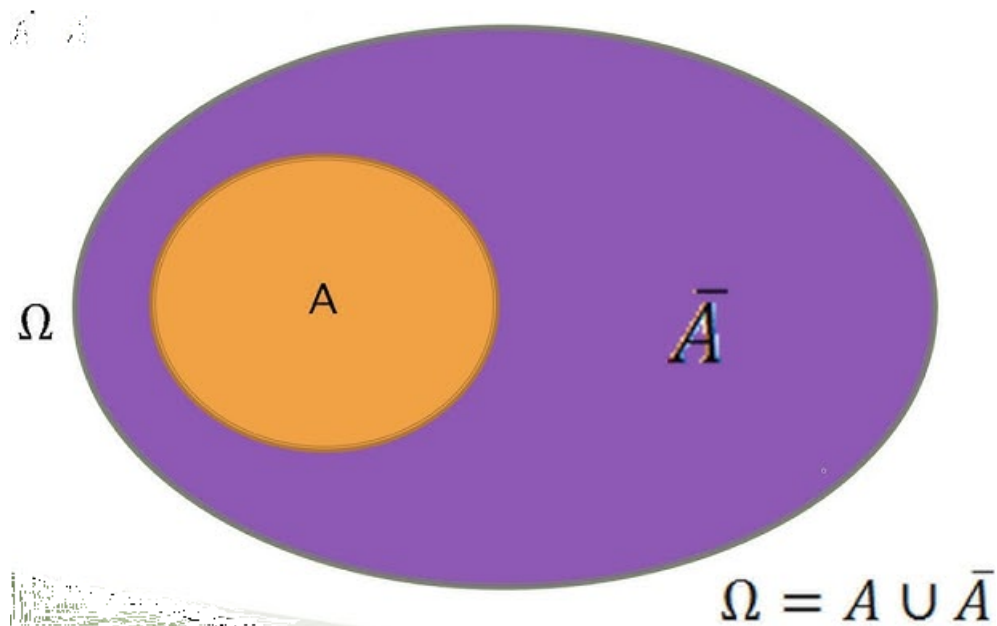
Soit $(\Omega, \mathfrak{M}, P)$. un espace de probabilité, et A, B, C trois événements quelconques de $\mathfrak{M}$.

- Probabilité de l'événement contraire :

$$P(\bar{A}) = 1 - P(A)$$

En effet, on a : $A \cup \bar{A} = \Omega$ et $A \cap \bar{A} = \emptyset$.

Donc $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.

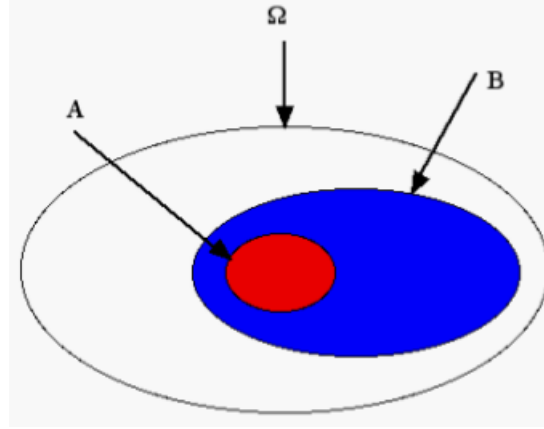


• **Probabilité de deux événements liés par l'inclusion :**

Si $A \subset B$ alors $p(A) \leq p(B)$.

Nous avons $B = A \cup (B \cap A^c)$ et $A \cap (B \cap A^c) = \emptyset$, donc

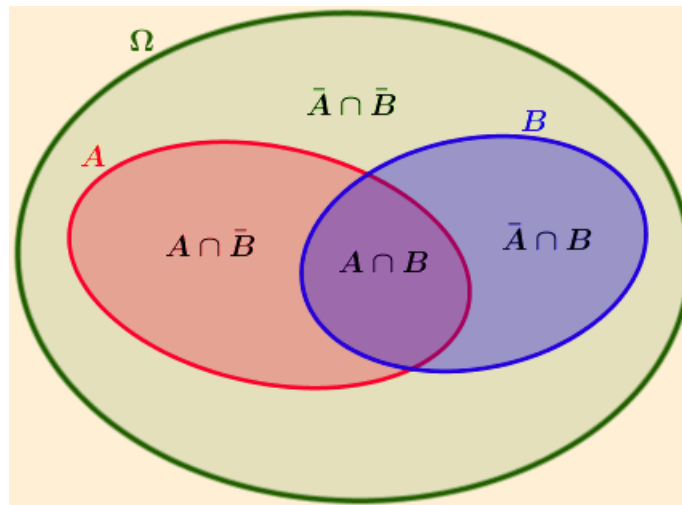
$P(B) = P(A) + P(B \cap A^c)$ avec $P(B \cap A^c) > 0$.



• **Probabilité de la différence de deux événements :**

$$P(A \cap B^c) = P(A) - P(A \cap B)$$

On a : $A = (A \cap B^c) \cup (A \cap B)$ et $(A \cap B^c) \cap (A \cap B) = \emptyset$



• **Probabilité de la réunion de deux événements :**

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

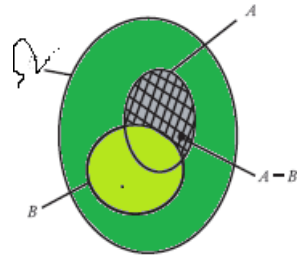
Car $A \cup B = A \cup (B \cap A^c)$ et

$$P(A \cup B) = P(A) + P(B \cap A^c)$$

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

Dans le cas de n événements on a la formule suivante (dite formule de Poincaré) :

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} S_k^n \text{ (avec)}$$



$$S_k^n = \sum_{1 \leq i_1, \dots, i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$$

- *Inégalité de Boole :*

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

En général on a la définition suivante :

DÉFINITION 3.7. Une probabilité est une application sur $P(\Omega)$, l'ensemble des parties de Ω , telle que :

- $0 \leq P(\Omega) \leq 1$, pour tout événement $A \subset \Omega$.
- $P(A) = \sum_{\omega \in A} P(\omega)$, pour tout événement A .
- $P(\Omega) = \sum_{\omega \in \Omega} P(\omega) = 1$.

REMARQUE 3.7. Que signifie un événement A a pour probabilité ?

0.95 : A va très probablement se produire.

0.03 : A a très peu de chance d'être réalisé.

4.0 : incorrect.

-2 : incorrect.

0.4 : A va se produire dans un peu moins de la moitié des essais.

0.5 : une chance sur deux.

0 : aucune chance que A soit réalisé.

De la définition, on peut facilement déduire la proposition suivante, fort utile pour faire quelques calculs :

PROPOSITION 3.1. Soient A et B deux événements.

1/ Si A et B sont incompatibles, $P(A \cup B) = P(A) + P(B)$.

2/ $P(A^c) = 1 - P(A)$.

3/ $P(\emptyset) = 0$.

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

DÉMONSTRATION. 1/ immédiat d'après le second point de la définition d'une probabilité.

2/ Comme A et A^c sont incompatibles, $1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$.

3/ $P(\emptyset) = 1 - P(\emptyset^c) = 1 - P(\Omega) = 0$.

4/ La technique est très souvent la même pour calculer la probabilité d'une réunion d'ensembles : on écrit cette réunion comme une union d'ensembles incompatibles, puis on utilise le 1). Ici, on écrit $A \cup B = A \cup (B \cap A^c)$ et on obtient :

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

Puis on écrit $B = (A \cap B) \cup (B \cap A^c)$ pour déduire $P(B) = P(A \cap B) + P(B \cap A^c)$.

En rassemblant ces deux égalités, on obtient la proposition.

Signalons une définition plus générale de probabilité, valable pour des espaces des possibles non dénombrables.

13. Lien entre la théorie des probabilités et des ensembles

Les résultats de la théorie des ensembles peuvent être interprétés sous forme d'événements en théorie des probabilités. Il existe un vocabulaire propre aux événements, différent du vocabulaire ensembliste.

notations	vocabulaire ensembliste	vocabulaire probabiliste
Ω	ensemble plein	événement certain
$\emptyset$	ensemble vide	événement impossible
ω	élément de Ω	événement élémentaire
A	sous - ensemble de Ω	événement
$\omega \in A$	ω appartient à A	réalise A
$A \subset B$	A inclus dans B	A implique B
$A \cup B$	union de A et B	A ou B
$A \cap B$	intersection de A et B	A et B
A^c ou $\bar{A}$	complémentaire de A	événement contraire de A
$A \cup B = \emptyset$	A et B disjoints	A et B incompatibles

14. Construction d'une probabilité et espace probabilisés discrets

• Les événements $\{\omega\}_{\omega \in \Omega}$ sont dits équiprobables s'ils ont tous la même probabilité, soit k . On suppose bien sûr que Ω est fini de cardinal n , muni de la tribu $\mathfrak{P}(\Omega)$. On a alors :

$$1 = P(\Omega) = P(\bigcup_{i=1}^n \omega_i) = \sum_{i=1}^n P(\omega_i) = n.k \Rightarrow \frac{1}{n}.$$

Si un événement $A \in \mathfrak{P}$ est constitué de n_A éléments alors :

$$P(A) = P(\bigcup_{\omega \in A} \{\omega\}) = \sum_{\omega \in A} P(\{\omega\}) = \frac{n_A}{n} = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

$$P(A) = \frac{\text{nombre de cas favorables à la réalisation de } A}{\text{nombre de cas possibles}}$$

$$P(A) = \frac{|A|}{|\Omega|}.$$

EXEMPLE 3.18. Soit $(\Omega, \mathfrak{P}(\Omega), P)$ un espace probabilité et supposons que les événements élémentaires $\{\omega_i\}$ sont équiprobables de probabilité P , $\text{card}(\Omega) = n$,

$$P(\Omega) = P(\bigcup_{i=1}^n \{\omega_i\}) = \sum_{i=1}^n P(\{\omega_i\}) = n.P = 1 = \frac{1}{n}.$$

$$P(A) = P(\{\omega_1, \omega_2, \omega_3\}) = P(\{\omega_1\}) + P(\{\omega_2\}) + P(\{\omega_3\}) = \frac{3}{n}.$$

REMARQUE 3.8. Cette probabilité P s'appelle probabilité uniforme sur $(\Omega, \mathfrak{P}(\Omega))$.

• Soit Ω un ensemble infini dénombrable muni de la tribu $\mathfrak{P}(\Omega)$ et p une probabilité sur $(\Omega, \mathfrak{P}(\Omega), P)$.

Posons $P(\{\omega\}) = P(\omega) = P_\omega$, $\omega \in \Omega$.

On a $\sum_{\omega \in \Omega} P_\omega = p(\bigcup_{\omega \in \Omega} \{\omega\}) = P(\Omega) = 1$ et pour tout

$$A \in \mathfrak{P}, \quad p(A) = p\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in \Omega} P_A.$$

Réciproquement la donnée d'une suite de nombre positifs $(P_\omega)_{\omega \in \Omega}$ tel que $\sum_{\omega \in \Omega} P_\omega$ et $P(\{\omega\}) = P_\omega$ définit une probabilité sur $(\Omega, \mathfrak{P}(\Omega))$.

Donc toute probabilité P sur un ensemble infini dénombrable $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$ muni de la tribu de ses parties $\mathfrak{P}(\Omega)$ est définie par une suite de nombre positifs P_n tels que $P(\omega_n) = P_n$ à condition que la série $\sum P_n$ ait pour somme 1.

Un tel espace probabilité $(\Omega, \mathfrak{P}(\Omega))$ est dit **Espace discret** et P est dite **probabilité discrète**.

REMARQUE 3.9. Il n'est ce pendant plus possible de construire sur Ω ensemble **infini dénombrable** une probabilité rendant les événements élémentaires équiprobables du fait qu'une série à terme positifs tous égaux entre eux ne peut converger vers 1.

15. Probabilité conditionnelles, indépendance et probabilité composées

1) Probabilité conditionnelles :

Étant donné l'espace probabilité $(\Omega, \mathfrak{P}, P)$ et un événement B de probabilité non nulle, on appelle : **probabilité conditionnelle** de A sachant B , soit $P(A/B)$, le rapport

$$P(A/B) = \frac{P(A \cap B)}{P(B) \neq 0}.$$

L'application $A \rightarrow p(A/B)$ définie sur $(\Omega, \mathfrak{P})$ est une probabilité.

2) Indépendance :

DÉFINITION 3.8. On dit que l'événement A est indépendant de l'événement B si $p(A/B) = p(A)$, on alors $P(A \cap B) = P(A).P(B) \Rightarrow P(B/A) = P(B)$

En d'autre terme B est indépendant de A .

Ainsi on peut écrire, A et B sont indépendants deux à deux $\Leftrightarrow P(A \cap B) = P(A).P(B)$.

3) Indépendance mutuelle :

DÉFINITION 3.9. Soient $A_1, A_2, \dots, A_n$ des éléments de $\mathfrak{P}$ ils sont dits **mutuellement indépendance** si pour toute 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 3.19. On lance deux dés, soient :

L'événement A " Avoir (6,6) "

L'événement B " Avoir un couple de nombre pairs ".

$$P(A) = \frac{1}{36} ; P(B) = \frac{9}{36}.$$

On lance les deux dés, on apprend que B s'est réalisé, on s'intéresse alors à l'événement A sachant que B est réalisé, que l'on note A/B .

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{36}}{\frac{9}{36}} = \frac{1}{9}.$$

$$\text{Car } P(A \cap B) = P(\{6, 6\}) = \frac{1}{36}$$

Soient les événements

A " le premier dé obtient un point impair. "

B " le deuxième dé obtient un point impair. "

C " la somme des points est impaire. "

$$\text{Donc on a } P(A) = \frac{18}{36}; P(B) = \frac{18}{36}; P(C) = \frac{18}{36} = \frac{1}{2}$$

$$P(A \cap B) = \frac{9}{36} = \frac{1}{4} = P(A).P(B)$$

$$P(A \cap C) = \frac{9}{36} = \frac{1}{4} = P(A).P(C)$$

$$P(B \cap C) = \frac{9}{36} = \frac{1}{4} = P(B).P(C)$$

A , A et C sont deux à deux indépendants, mais ils ne sont pas mutuellement indépendants, car $P(A \cap B \cap C) = 0$.

4) Probabilités composées :

PROPOSITION 3.2.

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

Résultat que l'on généralise à n événements $A_1, \dots, A_n$:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1).P(A_2/A_1).P(A_3/A_1 \cap A_2) \dots P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

5) Formules de Bayes :

1er formule de Bayes :

$$P(B/A) = \frac{P(A/B)P(B)}{P(A)} \quad (1)$$

DÉFINITION 3.10. $A_1, A_2, \dots, A_n$ forment un **système complet d'événements** si $A_1, A_2, \dots, A_n$ forment une partition de Ω , c'est-à-dire :

$$(1) \quad \begin{cases} A_i \cap A_j = \emptyset \\ \bigcup_{i=1}^n A_i = \Omega \end{cases}$$

• **2 ème formule de Bayes :**

Soit $(B_i)_{1 \leq i \leq n}$ un système complet d'événements et A un événement de probabilité non nulle, de la formule (1) on déduit la deuxième formule de Bayes (dite aussi **probabilité des cause**).

$$P(B_i/A) = \frac{P(A/B_i)P(B_i)}{\sum_{j=1}^n P(A/B_j)P(B_j)}, \quad 1 \leq i \leq n.$$

16. LES ARBRES DE PROBABILITÉS

Dans cette partie, on va donner et suivre des chemins pour construire un arbre de probabilité, c'est-à-dire une succession de nœuds consécutifs reliés par une branche encore appelée arête ou bien arc.

Dans un arbre, on distingue des arbres élémentaires. Un arbre élémentaire ne contient qu'un seul nœud, duquel partent un certain nombre de branches que l'on peut pondérer. La pondération d'une branche représente la probabilité de l'issue matérialisée par la branche considérée. L'arbre, plus généralement, est une concaténation d'arbres élémentaires.

17. Lois des arbres.

Deux lois fondamentales régissent les arbres en probabilité.

Loi du nœud :

la somme des pondérations portées par les branches issues d'un même nœud, est égale à 1.

Loi du chemin :

la pondération d'un chemin est égale au produit des pondérations des branches composant ce chemin.

18. Travaux Dirigés

EXERCICE 3.1. A et B sont des événements d'un univers Ω , tels que :

$$p(\bar{A}) = 3/5, P(B) = 1/4 \text{ et } p(A \cap B) = 1/8.$$

Déterminez les probabilités des événements :

$$A \cup B, A, \bar{B}, \bar{A} \cap \bar{B}, \bar{A} \cup \bar{B}, \text{ et, } A \cup \bar{B}.$$

Solution :

En utilisant les propriétés suivantes :

$$P(\bar{A}) = 1 - P(A), \quad P(A \cup B) = P(A) + P(B) - P(A \cap B), \quad P(\bar{A} \cup \bar{B}) = 1 - P(A \cap B), \quad P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B),$$

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

Alors, on a :

$p(A \cup B) = 21/40$, $p(A) = 2/5$, $p(\bar{B}) = 3/4$, $p(\bar{A} \cap \bar{B}) = 19/40$, $p(\bar{A} \cup \bar{B}) = 7/8$, $p(A \cap \bar{B}) = 11/40$.

EXERCICE 3.2. Une urne contient 8 boules noires et 4 boules rouges. On tire successivement et sans remise 3 boules de l'urne.

On considère les événements suivants :

A : " Deux boules exactement sont rouges. "

B : "La première boule tirée est noire. "

a) Calculez $p(A)$, $p(B)$ et $p(\bar{A} \cap \bar{B})$.

b) Déduisez $p(A \cup B)$, $p(A \cap \bar{B})$, $p(\bar{A} \cap \bar{B})$ et $p(A \cup B)$.

c) Calculez les probabilités conditionnelles $p(B|A)$ et $p(A|\bar{B})$.

d) Les événements A et B sont-ils indépendants ? (Justifiez la réponse).

f) Les événements A et B sont-ils incompatibles ? (Justifiez la réponse).

Solution :

a) $p(A) = p((N; R; R)) + p((R; N; R)) + p((R; R; N)) = \frac{12}{55} = 21,82 \%$

$p(B) = \frac{8}{12} = \frac{2}{3} = 66,67 \%$, $p(A \cap B) = p(NRR) = \frac{4}{55} = 7,27 \%$

b) $p(A \cup B) = p(A) + p(B) - p(A \cap B) = 81,21 \%$, $p(A \cap \bar{B}) = P(A) - p(A \cap B) = 14,55 \%$

$p(A \cap \bar{B}) = 1 \setminus p(A \cup B) = 18,79 \%$, $p(\bar{A} \cup \bar{B}) = 1 \setminus p(A \cap B) = 92,73 \%$

c) $p(B|A) = 1/3 = 33,33 \%$, $p(A|\bar{B}) = \frac{p(A \cap \bar{B})}{1 - p(B)} = 43,65 \%$

e) A et B sont dépendants car $p(A).p(B) \neq p(A \cap B)$.

d) A et B ne sont pas incompatibles car $A \cap B \neq \emptyset$.

EXERCICE 3.3. Une pièce bien équilibrée est lancée 6 fois :

a) Quelle est la probabilité d'avoir exactement deux "piles" ?

b) Quelle est la probabilité d'avoir au moins 4 "piles" ?

c) Quelle est la probabilité d'avoir au moins 1 "pile" ?

Solution :

EXERCICE 3.4. Dans chacune de situations décrites ci-dessous, énoncer l'événement contraire de l'événement donné.

1) Dans une classe, on choisit deux élèves au hasard.

A : " Les deux élèves sont des filles ".

2) Dans un groupe de suisses et de belges, on discute avec une personne.

B : " La personne est un homme belge ".

3) *Au restaurant, Luc prend un plat et un dessert.*

C : " Luc prend une viande et une glace ".

4) *A une loterie, Élise achète 3 billets.*

D : " L'un des billets au moins est gagnant ".

E : " Deux billets au maximum sont gagnants ".

Solution :

1) *L'évènement est " au moins un des deux élèves est un garçon ".*

2) *L'évènement est " La personne est soit une femme, soit un suisse ".*

3) *L'évènement est " Luc ne prend pas de viande ou ne prend pas de glace ".*

4) *L'évènement est " aucun billet n'est gagnant ".*

5) *L'évènement est " les trois billets sont gagnants ".*

EXERCICE 3.5. *Une urne contient des boules blanches, noires et rouges. On tire une boule de l'urne. On note :*

A : " Tirer une boule blanche ".

B : " Tirer une boule ni blanche ni rouge ".

C : " Tirer une boule noire ou une boule rouge ".

1) *A et B sont-ils incompatibles ?*

2) *B et C sont-ils incompatibles ?*

3) *Traduire par une phrase ne comportant pas de négation A et B .*

Solution :

1) *A et B sont incompatibles car une boule ne peut être simultanément blanche et non blanche.*

2) *B et C ne sont pas incompatibles car le tirage d'une boule noire les réalise simultanément.*

3) *L'évènement est " tirer une boule noire ou rouge ".*

4) *L'évènement est " tirer une boule blanche ou rouge ".*

EXERCICE 3.6. *Lors d'un jet de deux dés cubiques, on s'intéresse aux événements suivants :*

A : " La somme obtenue est au moins égale à 5 ".

B : " La somme obtenue est au plus égale à 5 ".

C : " La somme obtenue est strictement inférieure à 3 ".

1) *A et B sont-ils contraires ?*

2) *B et C sont-ils incompatibles ?*

3) *Traduire par une phrase C .*

4) *A et C sont-ils incompatibles ?*

Solution :

1) *A et B ne sont pas contraires car une somme égale à 5 les réalise simultanément.*

2) *B et C sont incompatibles car la somme ne peut être simultanément strictement supérieure*

à 5 (événement) et strictement inférieure à 3 (événement C).

3) L'événement est " La somme est supérieure ou égale à 3 ".

4) A et ne sont pas incompatibles car ils sont simultanément réalisés par une somme supérieure ou égale à 5.

EXERCICE 3.7. On jette une pièce de monnaie 3 fois de suite.

1) Donner la liste de tous les résultats possibles en notant P pour Pile et F pour Face (exemple : PPF).

2) Donner la probabilité des événements suivants :

A " le tirage ne comporte que des Piles ".

B " le tirage comporte au moins une fois Face ".

Solution :

1) On peut lister

$$W = \{ PPP; PPF; PFP; PFF; FPP; FPF; FFP; FFF \}.$$

D'où $\text{Card}(W) = 8$.

2) Les tirages étant équiprobables, on a $P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{1}{8}$ (seul le tirage PPP convient).

Enfin, on remarque que $B = \bar{A}$ donc $P(B) = P(\bar{A}) = 1 - P(A) = 1 - \frac{1}{8} = \frac{7}{8}$.

EXERCICE 3.8. Dans une assemblée de 250 personnes, on ne remarque que les hommes portant la cravate ou ayant les yeux bleus. Il y a 120 hommes qui portent la cravate, 85 hommes qui ont les yeux bleus, dont 50 portent la cravate.

On discute avec une personne choisie au hasard dans cette assemblée.

1) Quelle est la probabilité que ce soit un homme portant la cravate.

2) Quelle est la probabilité que ce soit un homme aux yeux bleus et portant la cravate.

3) Quelle est la probabilité que ce soit un homme aux yeux bleus ou portant la cravate.

4) Quelle est la probabilité de discuter avec une personne qui n'est ni un homme aux yeux bleus, ni un homme portant la cravate ?

Solution :

Le tableau suivant permet de dénombrer les différentes catégories :

	Cravate (événement C)	Pas de Cravate (événement)	Total
Yeux Bleus (événement B)	50	35	85
Yeux non bleus (événement)	70	95	165
Total	120	130	250

On note Ω l'univers des possibles, ensemble des 250 personnes.

Ainsi $\text{Card}(\Omega) = 250$. Il y a équiprobabilité des choix de personnes. Ainsi

$$1) P(C) = \frac{\text{Card}(C)}{\text{Card}(\Omega)} = \frac{120}{250} = \frac{12}{25}.$$

$$2) P(B \cap C) = \frac{\text{Card}(B \cap C)}{\text{Card}(\Omega)} = \frac{50}{250} = \frac{1}{5}.$$

$$3) P(B \cup C) = P(B) + P(C) - P(B \cap C) = \frac{85}{250} + \frac{120}{250} - \frac{50}{250} = \frac{155}{250} = \frac{31}{50}.$$

$$(On\ pouvait\ aussi\ directement\ écrire\ P(B \cup C) = \frac{\text{Card}(B \cup C)}{\text{Card}(\Omega)} = \frac{50+70+35}{250} = \frac{155}{250} = \frac{31}{50}.$$

$$3) P(\bar{B} \cap \bar{C}) = P(\overline{B \cup C}) = 1 - P(B \cup C) = 1 - \frac{31}{50} = \frac{19}{50}.$$

EXERCICE 3.9. Lors d'un référendum, deux questions étaient posées.

65 % des personnes ont répondu "oui" à la première question, 51 % ont répondu "oui" à la seconde question, et 46 % ont répondu "oui" aux deux questions.

1) Quelle est la probabilité qu'une personne ait répondu "oui" à l'une ou l'autre des questions ?

2) Quelle est la probabilité qu'une personne ait répondu "non" aux deux questions ?

Solution :

Si on note A l'événement "la personne a répondu oui à la première question" et B l'événement "la personne a répondu oui à la deuxième question", l'énoncé nous fournit

$$P(A) = 0,65, P(B) = 0,51 \text{ et } P(A \cap B) = 0,46.$$

$$1) \text{ On calcule } P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,65 + 0,51 - 0,46 = 0,7.$$

$$2) \text{ On calcule } P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 0,7 = 0,3.$$

EXERCICE 3.10. Dans un magasin d'électroménager, on s'intéresse au comportement d'un acheteur potentiel d'un téléviseur et d'un magnétoscope. La probabilité pour qu'il achète un téléviseur est de 0,6.

La probabilité pour qu'il achète un magnétoscope quand il a acheté un téléviseur est de 0,4.

La probabilité pour qu'il achète un magnétoscope quand il n'a pas acheté de téléviseur est de 0,2.

1) Quelle est la probabilité pour qu'il achète un téléviseur et un magnétoscope ?

2) Quelle est la probabilité pour qu'il achète un magnétoscope ?

3) Le client achète un magnétoscope. Quelle est la probabilité qu'il achète un téléviseur ?

Solution

On note T l'événement "le client achète un téléviseur" et M l'événement "le client achète un magnétoscope." L'énoncé fournit $P(T) = 0,6$ (donc $P(\bar{T}) = 1 - P(T) = 1 - 0,6 = 0,4$)

1) En appliquant la formule de définition d'une probabilité conditionnelle, dans sa "version multiplicative",

$$P_T(M) = \frac{P(T \cap M)}{P(T)} \Leftrightarrow P(T \cap M) = P_T(M) \times P(T) = 0,6 \times 0,4 = 0,24.$$

2) En appliquant la formule des probabilités totales,

$$P(M) = P(T \cap M) + P(\bar{T} \cap M) \Rightarrow P(M) = P(T) \times P_T(M) + P(\bar{T}) \times P_{\bar{T}}(M)$$

$$\Rightarrow P(M) = (0, 6) \times (0, 4) + (0, 4) \times (0, 2) = 0, 32.$$

$$3) \text{ On demande } p_M(T) = \frac{P(M \cap T)}{P(M)} = \frac{0,6 \times 0,4}{0,32} = 0, 75.$$

$$4) \text{ Puisque } P(M) = 0, 32, \text{ on a } P(\bar{M}) = 1 - 0, 32 = 0, 68.$$

$$\text{Puisque } P_M(T) = 0, 75. \text{ on a } P_M(\bar{T}) = 1 - 0, 75 = 0, 25.$$

On calcule de la même manière qu'à la question 3)

$$P_{\bar{M}}(T) = \frac{P(\bar{M} \cap T)}{P(\bar{M})} = \frac{0,6 \times 0,6}{0,68} = \frac{0,36}{0,68} = \frac{9}{17}.$$

$$\text{donc } P_{\bar{M}}(\bar{T}) = 1 - \frac{9}{17} = \frac{8}{17}.$$

EXERCICE 3.11. Justifiez vos réponses avec un arbre, si possible, et précisez s'il s'agit d'une combinaison, d'une permutation ou d'un arrangement.

Le code de ma carte de crédit est composé de 4 chiffres, mais j'ai oublié le code.

a) Dans un premier temps, j'aimerais me souvenir des chiffres présents dans mon code.

Combien y a-t-il de possibilités (dans le cas avec 4 chiffres différents) ?

La mémoire me revient partiellement et il me semble que le code de 4 chiffres comporte peut-être les chiffres : 1, 3, 4, 5, 7 ou 8.

b) Combien y a-t-il de codes possibles (sans répétition de ces chiffres) ?

c) Combien y a-t-il de codes possibles (avec répétition) ?

Finalement, je me souviens que le code comporte seulement les chiffres : 1, 4, 7 et 8 et il n'y a pas de répétition de chiffres.

d) Combien y a-t-il de codes possibles ?

e) Quelle est la probabilité de faire le bon code ?

Conclusion : Que faut-il faire pour augmenter la sécurité de votre carte.

Solution :

$$a) C_{10}^6 = 210.$$

$$b) A_4^6 = 360.$$

$$c) A_6^4 = 1296.$$

$$d) P_4 = 4! = 24.$$

$$e) \text{ Probabilités d'obtenir le bon code } = \frac{1}{24}.$$

EXERCICE 3.12. Dans les élections des comités pédagogiques, 5 étudiants délégués et 6 étudiants comme adjoints se présentent.

Combien de comités comprenant délégués et adjoints qu'on peut former ?

EXERCICE 3.13. Qu'elle est la capacité (théorique) d'un réseau téléphonique sachant qu'un numéro est constitué de 6 chiffres.

EXERCICE 3.14. *La probabilité de fermeture du relai i des circuits décrits ci-dessous est p_i .*

Tous les relais fonctionnent indépendamment. Dans chacun des cas suivants, quelle est la probabilité pour que le courant passe entre A et B ?

- a) A et B sont séparés par n relais reliés en série.*
- b) A et B sont séparés par n relais reliés en parallèle.*

EXERCICE 3.15. *Avec quelle probabilité une famille de 3 enfants comporte-t-elle au moins un garçon ?*

EXERCICE 3.16. *Trois joueurs A , B , C lancent chaque un dé à 6 faces.*

- 1) Combien a-t-il de résultats possibles ?*
- 2) Combien a-t-il de résultats où les points obtenus sont tous différents.*

EXERCICE 3.17. *Quel est le nombre d'ordre différents dont un médecin peut recevoir 7 personnes arrivées au même temps.*

EXERCICE 3.18. *Une urne contient 9 boules numérotées de 1 à 9, on prélève 4 boules l'une après l'autre.*

Quel est le nombre de chiffres qu'on peut former avec les boules prélevées.

N.B :*les tirages se font sans remise.*

EXERCICE 3.19. *Une assemblée veut élire un bureau de 3 membres, il y a 5 candidats.*

- 1) Combien de bureau peut-on former ?*
- 2) Si parmi les cinq candidats il y a 2 femmes, combien de bureaux comportant exactement 2 femmes peut-on former ?*

EXERCICE 3.20. *On dispose d'un lot de pièces dont trois sont bonnes 3 sont défectueuses.*

- 1) Combien peut-on réaliser d'échantillon différents de 3 pièces ?*
- 2) Combien parmi ces échantillons, ceux qui contiennent 3 pièces bonnes ?*
- 3) Combien d'échantillons qui contiennent au moins une pièce bonne ?*

EXERCICE 3.21. *Un hôpital veut engager 4 médecins dans 4 spécialités différentes. Six médecins se présentent.*

Combien de choix s'offrent au responsable de l'embauche dans les cas suivants :

- 1) Les six médecins sont polyvalents (pouvant occuper un des 4 postes).*
- 2) Un seul est polyvalent pour les 4 postes, les 5 autres le sont seulement dans 3 branches, les même pour les cinq.*

EXERCICE 3.22. *Un groupe composé de $2n$ filles et garçon se vont séparer en deux sous groupes de même effectif. Calculer la probabilité p que chaque sous groupe soit composé d'autant de filles que de garçons.*

EXERCICE 3.23. *Parmi les fonctions suivantes les quelle définissent une probabilité sur $\Omega = \{a, b, c\}$?*

1) $P(a) = \frac{1}{4}; P(b) = \frac{1}{3}; P(c) = \frac{1}{2}.$

2) $P(a) = \frac{1}{6}; P(b) = \frac{1}{3}; P(c) = \frac{1}{2}.$

3) $P(a) = \frac{2}{3}; P(b) = \frac{-1}{3}; P(c) = \frac{2}{3}.$

4) $P(a) = 0; P(b) = \frac{1}{3}; P(a) = \frac{2}{3}.$

EXERCICE 3.24. *Soient A et B deux événements tels que :*
 $P(A) = \frac{3}{8}; P(B) = \frac{1}{2}; P(A \cap B) = \frac{1}{4}.$

Calculer $P(A \cup B); P(\bar{A}); P(\bar{B}); P(\bar{A} \cap \bar{B}); P(\bar{A} \cup \bar{B})$

EXERCICE 3.25. *Une urne contient 10 boules dont 4 rouge et 6 noires, on prèle 3 boules de cette urne. En distinguant le cas d'un tirage simultané et de tirage successifs, calculer la probabilité pour que :*

- 1) *Les trois boules soient de même couleur.*
- 2) *Deux boules exactement sont de couleur rouge.*
- 3) *Au moins une des boules est de couleur rouge.*

EXERCICE 3.26. *On jette simultanément une pièce de monnaie et un dé.*

- 1) *Décrire l'ensemble fondamental.*
- 2) *Soit A l'événement **Obtenir pile** et B l'événement **Obtenir un nombre < 3** .*
 - *Décrire $\bar{A}$ et $\bar{B}$.*
 - *Donner la composition de A , B et $A \cap B$*
- 3) *Calculer $P(A)$, $P(B)$, $P(A \cap B)$, $P(A/B)$.*
- 4) *A et B sont ils indépendants ?*

EXERCICE 3.27. *Une urne contient 4 rouge et 6 blanches et 10 boules noires, on tire 3 boules à la fois. Calculer la probabilité des événements suivants :*

- 1) *Toutes les boules tirées sont rouges.*
- 2) *Toutes les boules sont noires.*
- 3) *Deux boules sont rouges et une blanche.*
- 4) *Une boule de chaque couleur.*
- 5) *Au moins une boule noire.*

EXERCICE 3.28. *Proposer un univers pour les expériences aléatoires suivantes et dénombrer les résultats possibles :*

1/ On lance un dé.

2/ On lance 2 dés.

3/ On tire trois cartes dans un jeu.

4/ On place les 5 lettres qui forment " proba " au hasard sur une règlette de Scrabble.

5/ On place les 6 lettres qui forment " erreur " au hasard sur une règlette de Scrabble.

EXERCICE 3.29. Soit P une probabilité sur un ensemble et les deux événements A et B .

On suppose que

$$P(A \cup B) = \frac{7}{8}; P(A \cap B) = \frac{1}{4}; P(A) = \frac{3}{8}.$$

Calculer $P(B)$, $P(A \cap B^c)$, $P(B \cap A^c)$.

EXERCICE 3.30. Supposons que les faces d'un dé sont truquées de telle manière que les numéros impairs ont chacun la même chance d'apparaître, chance qui est deux fois plus grande que pour chacun des numéros pairs. On jette le dé. Quelle est la probabilité d'obtenir un nombre supérieur ou égal à 4 ?

EXERCICE 3.31. Un parking contient douze places alignées. Huit voitures s'y sont garées au hasard, et l'on observe que les quatre places libres se suivent. Est-ce surprenant ?

EXERCICE 3.32. La probabilité qu'un objet fabriqué à la chaîne ait un défaut est de 0,1. Trouver la probabilité que, dans un lot de 100 objets, il y ait au moins un objet défectueux. Quelle est la probabilité qu'il y ait, dans un tel lot, exactement un objet défectueux ?

EXERCICE 3.33. Soient M_1 , M_2 , M_3 trois personnes. La première M_1 dispose d'une information codée sous forme + ou -. Elle la transmet à la deuxième personne M_2 . Puis M_2 la transmet à M_3 .

Malheureusement, à chaque fois que l'information est transmise, il y a une probabilité P que l'information soit changer en son contraire. En tenant compte du fait que deux changements rétablissent la vérité, quelle est la probabilité pour que M_3 ait le bon message ?

Et si M_3 transmet l'information dont il dispose à une quatrième personne M_4 , quelle est la probabilité pour que M_4 ait la bonne information ?

Lorsque $P = 0,2$, quelle est la valeur numérique de cette probabilité ?

EXERCICE 3.34. Un nouveau vaccin a été testé sur 12500 personnes, 75 d'entre elles, dont 35 femmes enceintes, ont eu des réactions secondaires nécessitant une hospitalisation. Parmi les 12500 personnes testées, 680 personnes sont des femmes enceintes.

1/ Quelle est la probabilité pour une femme enceinte, d'avoir une réaction secondaire si elle reçoit un vaccin ?

2/ Quelle est la probabilité pour une personne non enceinte, d'avoir une réaction secondaire ?

EXERCICE 3.35. Dans une usine, la machine A fabrique 60 % des pièces, dont 2 % sont défectueuses.

La machine B fabrique 30 % des pièces, dont 3 % sont défectueuses.

La machine C fabrique 10 % des pièces, dont 4 % sont défectueuses.

1/ On tire une pièce au hasard dans la fabrication. Quelle est la probabilité qu'elle soit défectueuse ?

2/ On tire une pièce au hasard dans la fabrication. Elle est défectueuse. Quelle est la probabilité qu'elle ait été fabriquée par la machine A ? par la machine B ? par la machine C ?

EXERCICE 3.36. Deux chausseurs de bus se relaient sur la même ligne. Lors d'une grève, le premier a 60 % de chances de faire grève et le second 80 %. Pendant la prochaine grève, quelle est la probabilité pour qu'un seul des deux chausseurs fasse grève ?

EXERCICE 3.37. Une loterie comporte 500 billets dont deux seulement sont gagnants. Combien doit-on acheter de billets pour que la probabilité d'avoir au moins un billet gagnant soit supérieure ou égale à 0.5 ?

EXERCICE 3.38. 1/ Dans une classe de 36 élèves, quelle est la probabilité pour que deux élèves au moins soient nés le même jour ? (on considérera que l'année compte 365 jours, et que toutes les dates d'anniversaires sont indépendantes et équiprobables).

2/ Généraliser ce résultat pour une classe de n élèves.

Tracer le résultat obtenu.

EXERCICE 3.39. Deux événements A et B disjoints ($A \cap B = \emptyset$) et de probabilités non nulles peuvent-ils être indépendants ?

EXERCICE 3.40. Une urne contient 10 bulletins indiscernables au toucher, de 3 sortes : 4 sont marqués "oui", 3 sont marqués "non" et 3 sont marqués "blanc". Un joueur tire simultanément deux bulletins de l'urne.

Quelle est la probabilité qu'il obtienne un tirage de deux bulletins de sortes différentes.

EXERCICE 3.41. Un sac contient 5 jetons verts (numérotés de 1 à 5) et 4 jetons rouges (numérotés de 1 à 4).

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 au hasard 3 jetons du sac. Reprendre alors les questions a), b), c) et d).

EXERCICE 3.42. Une urne contient 3 pièces équilibrées. Deux d'entre elles sont normales : elles possèdent un côté "Pile" et un côté "Face". La troisième est truquée et possède deux côtés "Face".

On prend une pièce au hasard dans l'urne et on effectue de manière indépendante des lancers successifs de cette pièce. On considère les événements suivants :

B : la pièce prise est normale. $\bar{B}$: la pièce prise est truquée.

P : on obtient "Pile" au premier lancer. F_n : on obtient "Face" pour les n premiers lancers.

a) Quelle est la probabilité de l'événement B ?

b) Quelle est la probabilité de l'événement P sachant que B est réalisé ?

2) Calculer la probabilité de l'événement $P \cap B$, puis de l'événement $P \cap \bar{B}$.

En déduire la probabilité de l'événement P . 3) Calculer la probabilité de l'événement $F_n \cap B$ puis de l'événement $F_n \cap \bar{B}$.

En déduire la probabilité de l'événement F_n .

EXERCICE 3.43. On lance un dé à six faces équilibré.

a/ Quelle est la probabilité d'obtenir un nombre pair ?

b/ Quelle est la probabilité d'obtenir un multiple de 3 ?

EXERCICE 3.44. Une urne contient trois boules rouges, quatre boules noires et deux boules jaunes indiscernables au toucher.

On tire successivement et sans remise deux boules dans l'urne.

a Quelle est la probabilité d'obtenir deux boules rouges ?

b Quelle est la probabilité d'obtenir deux boules de même couleur ?

EXERCICE 3.45. Dans un jeu, on doit tourner deux roues. La première roue donne une lettre : A, B, C ou D avec la même probabilité. La deuxième roue donne un chiffre entre 1 et 6 avec la même probabilité.

Si, après avoir tourné les roues, les aiguilles se trouvent comme sur le schéma, on note $(A, 1)$ le résultat obtenu.

a Quelle est la probabilité du résultat $(B, 2)$?

b Quelle est la probabilité d'obtenir C et un chiffre impair ?

Bibliographie

- [1] J. Blard-Laborderie, L'essentiel des outils de statistique descriptive pour aborder des études en sciences humaines et sociales, 2015.
- [2] G. Calot, Cours de statistique descriptive, Dunod, 1969.
- [3] M. Tenenhaus, statistique : Méthodes pour décrire, expliquer et prévoir, Dunod, 2006.
- [4] J.-J. Dreesbeke, Éléments de statistiques, Ellipses, 2001.
- [5] L. Leboucher and M.-J. Voisin, Introduction à la statistique descriptive, 2013.
- [6] F. Mazerolle, Statistique descriptive, 2009.
- [7] B. Oukacha and M. Benmessaoud, Statistique descriptive et calcul des probabilités, 2013.
- [8] J.-Y. Ouvrard, Probabilités : Tome 1, 2001.
- [9] J. Vaillant, Éléments de Statistique descriptive, 2015.