La statistique au collège

Par:
Sérigne Touba Sall: directeur de l'IREMPT
Djibril Ndiaye: conseiller pédagogique à l'IREMPT
Mory Dieng: conseiller pédagogique à l'IREMPT

En hommage à:
Marcel Diouf et **Farba Faye**

TABLE DES MATIÈRES

AVANT PROPOS

Ce manuel est le fruit d'une fructueuse collaboration entre le moyen et le supérieur, ce qui illustre à plus d'un titre la pertinence de la mission de l'IREMPT, consistant à fédérer les différentes strates du système éducatif, allant du préscolaire au supérieur.
Le fait qu'il soit exclusivement consacré à la statistique est une innovation de taille, dans la mesure où cette discipline a été toujours considérée, à tort ou à raison, comme un appendice des activités numériques. Or, l'importance de la statistique dans les activités humaines et sa force de persuasion dans les discours lui confèrent, de fait, un statut d'outil incontournable pour la recherche, la gouvernance et la communication.
Pour s'en convaincre, il suffit de suivre les rapports des professionnels de la santé et les politiques sur la situation de la pandémie au plan national et international, rapports qui sont émaillés de statistiques. Pour dire que l'analphabète de demain, c'est celui-là qui ne comprend pas le langage de la statistique car, rares sont les domaines qui ne font pas appel à cette discipline pour les besoins d'un éclairage.
Au Sénégal, malheureusement certains enseignants rencontrent beaucoup de difficultés dans l'enseignement de la statistique, surtout ceux qui n'ont pas eu la chance de l'étudier au niveau de l'enseignement moyen.

En effet, les recherches menées par les étudiants de la FASTEF dans le cadre de leurs mémoires de fin de formation, révèlent que la statistique est placée en fin d'année, par la majorité des enseignants, dans leur répartition annuel du programme, ce qui pourrait nourrir le soupçon d'une stratégie de contournement de la difficulté inhérente à cette branche. Pour pallier ces difficultés, on ne trouve pas mieux qu'un fascicule de statistique conforme au programme et rédigé par des spécialistes du domaine qui ont une longue expérience dans l'enseignement des mathématiques.

Par ailleurs, la structuration des contenus en chapitres sous forme de modules dont la cohérence et la cohésion internes facilitent leur exploitation par les utilisateurs, favorise davantage la lisibilité du document. Dans le même ordre d'idées, l'organisation pédagogique des chapitres en différentes parties à savoir : un résumé de cours succinct qui découle d'activités de découverte, des exercices types d'applications des notions acquises et des exercices de consolidation bien étudiés, font de cette ressource un support didactique pour les enseignants et les élèves, doublé de document de référence pour les formateurs.

Actuellement, qui parle de statistique parle d'utilisation des technologies de l'information et de la communication, plus particulièrement de l'informatique. C'est pourquoi ce document consacre une partie importante à l'initiation à l'utilisation du logiciel libre (open source) R pour la capacitation progressive des usagers au traitement des données statistiques liées aux connaissances statistiques acquises.

Cet ouvrage, comme les autres en instance, est une bonne illustration de la mission de l'IREMPT consistant à mettre en synergie les acteurs du cycle fondamental, du secondaire

et du supérieur pour l'amélioration de la qualité des apprentissages.
Nous invitons, dès lors, les pouvoirs publics, à travers le ministère de l'éducation et l'Université à encourager et à accompagner davantage des initiatives de cette nature pour une meilleure cohérence des actions du système éducatif.

Mangary Ka,
Maître de conférences assimilé,
Inspecteur de l'enseignement élémentaire à la retraite.

INTRODUCTION

Dans sa quête de bonheur et d'équilibre l'homme est appelé à programmer ses activités afin d'en maîtriser l'impact sur son environnement social, culturel et économique. Dès lors la statistique devient une activité sociale privilégiée.

Le mot statistique est attribué au professeur ACHENWALL[1] qui l'aurait tiré du mot allemand "**STAATSKUNDE**" qui signifie "**Science des états**".

La statistique est une science qui a pour objet la collecte de données, leur organisation, leur analyse et leur interprétation rigoureuse en vue d'une prise de décision. L'activité statistique renvoie aussi aux méthodes de traitement, d'analyse et d'exploitation des données.

L'activité statistique remonte à plus de 2000 ans avant Jésus Christ notamment en Egypte pharaonique et en Chine.

On distingue deux méthodes d'analyse statistique :

- La statistique descriptive qui décrit les faits à partir d'observations ; elle est déductive.
- La statistique inférentielle qui consiste à étendre à un groupe les résultats tirés d'une partie de ce même groupe. Elle teste des hypothèses, émet des conjectures et fait des prédictions. Cette démarche est rigoureusement validée par des notions mathématiques et doit éviter une mauvaise interprétation.

Les domaines d'application de la statistique sont divers et très vastes : la Gestion, la Finance, la Médecine, la Météo, les Sciences sociales...

C'est en 1992 avec la réforme des programmes sénégalais, initiée deux ans plus tôt que la statistique qui n'était abordée auparavant qu'en classe de première fut introduite au premier cycle notamment en quatrième et en troisième.

Notons que l'activité statistique au collège est exclusivement descriptive. Elle procède par :

- une enquête pour la collecte des données ;
- un dépouillement pour ordonner ces données, les classer et les présenter dans des tableaux et graphiques ;
- une exploitation et une interprétation des résultats.

1. Gottfried ACHENWALL (1719-1772) était un économiste allemand, professeur d'économie de statistique, de politique, de droit à l'université de Gottingen

Programme officiel de la classe de quatrième

Contenus	Commentaires	Compétences exigibles
1) **Exemples et vocabulaire** : population, individu, échantillon, caractère qualitatif, caractère quantitatif, variables	- Introduire le vocabulaire à partir d'exemples de la vie courante. - On étudiera uniquement les caractères quantitatifs discrets.	Connaître le vocabulaire suivant : population, individu, échantillon, caractère quantitatif, variable, valeur du caractère (modalité), effectif, mode, moyenne, fréquence, pourcentage.
2) Classement des données statistiques a) Séries statistiques brutes. b) Séries statistiques ordonnées : effectif, mode, moyenne, fréquence et pourcentage.	- Des activités d'enquêtes au niveau de la classe (notes, âge, taille des élèves,...) fourniront des séries statistiques qui pourront être exploitées dans la suite du chapitre. On a l'habitude d'ordonner les séries dans l'ordre croissant. - Il faudra attirer l'attention des élèves sur l'intérêt de ces différentes notions dans la vie courante.	- Ordonner une série statistique. - Établir le tableau des effectifs. - Déterminer le mode d'une série statistique. - Calculer la fréquence et le pourcentage d'une valeur du caractère et la moyenne d'une série statistique.
3) Représentations Diagramme en bâtons, diagramme à bandes, diagramme circulaire, diagramme semi-circulaire.	- On s'intéressera surtout à l'aspect comparatif de ces différents diagrammes. - L'interprétation consiste à donner un avis argumenté à partir des résultats obtenus.	- Représenter une série statistique par un diagramme en bâtons, par un diagramme à bandes, par un diagramme circulaire, par un diagramme semi-circulaire. - Déterminer à l'aide d'un diagramme les effectifs d'une série statistique. - Interpréter des données statistiques.

Programme officiel de la classe de troisième

L'objectif est d'amener les élèves à interpréter les résultats obtenus. La finalité étant d'amener les élèves à avoir une attitude critique devant les informations statistiques reçues dans la vie courante.

Contenus	Commentaires	Compétences exigibles
1) **Exemples et vocabulaire**. - amplitude d'une classe, - centre de classe. 2) **Classement et représentation des données statistiques** a) Distribution groupée en classe d'égale amplitude. b) Histogramme	On introduira ces termes à partir d'exemples concrets. Des activités d'enquêtes menées par les élèves seront exploitées pour les histogrammes et les diagrammes cumulatifs.	Regrouper en classes une série brute. Déterminer les tableaux des effectifs et des fréquences cumulées croissantes ou décroissantes. Construire un histogramme. Interpréter un graphique représentant une série statistique.
c) Diagramme des effectifs cumulés et des fréquences cumulées.	• Les diagrammes cumulatifs se feront aussi avec des exemples à caractère discret. • On utilisera les diagrammes des effectifs cumulés dans des activités pratiques.	• Construire un diagramme cumulatif.
3) Paramètres de position Classe modale, médiane (détermination par le calcul et par le graphique), moyenne	• On recherchera des antécédents de $\frac{N}{2}$, $\frac{N}{4}$, $\frac{3N}{4}$, sur les représentations graphiques, N étant l'effectif total.	• Déterminer la moyenne, la classe modale. • Déterminer, graphiquement et par le calcul , la médiane.

CHAPITRE 1

TERMINOLOGIE OU VOCABULAIRE STATISTIQUE

1.1 Étude statistique

Faire de la statistique consiste à étudier un aspect commun, une caractéristique commune aux éléments d'un groupe de personnes, d'animaux, d'objets, de choses, de phénomènes... Cette étude se fera à travers des observations sur les éléments composant ce groupe.
Par exemple une étude statistique peut porter sur :

— la production céréalière du Sénégal en 2018,

— les notes d'un devoir surveillé des élèves d'une classe de $3^{\text{ième}}$.

1.2 Population statistique

Définition
On appelle population l'ensemble sur lequel porte l'étude statistique.
Exemples
Une population d'élèves, de malades, de voitures, de vendeurs de journaux, de mots d'un texte.

1.3 Individu ou unité statistique

Définition
Tout élément de la population étudiée est appelé individu ou unité statistique.
Exemples :
- Dans une population d'élèves chaque élève est un individu.
- Dans une population de voitures chaque voiture est un individu.
- Dans une population de collèges d'enseignement chaque collège d'enseignement est un

individu.
- Dans une population de mots d'un texte chaque mot est un individu.

1.4 Échantillon de population

Lorsque pour des raisons diverses on ne peut atteindre tous les individus d'une population, on se contente d'une partie de celle-ci appelée échantillon.

Définition
On appelle échantillon toute partie de la population.

Remarque :
Un échantillon doit être représentatif de la population dans sa composition.

Exemples :
- L'ensemble des élèves d'un établissement qui participent aux activités de l'U.A.S.S.U (Union des Associations Sportives, Scolaires et Universitaires) constitue un échantillon des licenciés de cette organisation.
- Dans une fabrique de piles, l'ensemble des piles qui sont testées pour déterminer leur durée de fonctionnement constitue un échantillon des piles fabriquées.

1.5 Variable ou caractère statistique

Définition
On appelle variable ou caractère l'aspect ou la caractéristique commune aux individus de la population que l'on souhaite étudier.

Exemples
- Dans une population d'élèves, une étude statistique peut sur la taille ou le genre.
- Dans une population d'œuvres littéraires une étude statistique peut porter sur le genre littéraire (roman, poésie, etc).

1.6 Modalités d'une variable ou d'un caractère

Définition
On appelle modalité d'un caractère étudié toute valeur prise par ce caractère.

Exemples
- Homme et femme sont les modalités du caractère genre.
- Arabe, espagnol, sciences physiques, sont des modalités pour le caractère option au collège.

— $1,78\,m$; $1,80\,m$ sont des modalités pour la variable taille.
— 2 enfants, 5 enfants sont des modalités pour le caractère nombre d'enfants par famille.

1.7 Nature d'un caractère ou d'une variable statistique

1.7.1 Variable quantitative

Définition
Une variable est dite quantitative lorsqu'elle est mesurable (c'est à dire obtenue par un décompte, une mesure, un repérage...). Ses valeurs sont des réels.

Exemples :
— Dans une population d'élèves la taille en mètres, le poids en kg, l'âge en années, le nombre d'absences au premier semestre sont des caractères quantitatifs.

— La durée de fonctionnement en heures des piles électriques fabriquées dans une usine est une variable quantitative.

1.7.2 Variable qualitative

Définition
Une variable est dite qualitative lorsqu'elle n'est pas mesurable.

Exemples
— Dans une population de travailleurs d'une entreprise, le genre, la profession, la religion, le plat préféré sont des caractères qualitatifs.
— Dans l'ensemble des ouvrages d'une bibliothèque, le genre littéraire est un caractère qualitatif.
— Dans l'ensemble des mots d'un texte, la nature grammaticale (nom, verbe, article, pronom...) est un caractère qualitatif.

1.8 Types de variables quantitatives

1.8.1 Variable quantitative discrète

Une variable quantitative est dite discrète lorsqu'elle prend des valeurs isolées (en général, des valeurs entières).

Exemples
— Dans l'ensemble des familles d'un quartier le nombre d'enfants par famille est un caractère quantitatif discret.

— Dans une population d'élèves le nombre d'absences au premier semestre est un caractère quantitatif discret.

Remarque :
Une variable quantitative est considérée comme discrète lorsque le nombre de modalités est suffisamment faible pour être présenté dans un tableau statistique de taille raisonnable.

1.8.2 Les variables quantitatives continues

Une variable quantitative est dite continue lorsqu'elle est susceptible de prendre n'importe quelle valeur d'un intervalle.

Exemples :
Dans une population d'élèves, la Taille, le Poids et la Moyenne au premier semestre sont des caractères quantitatifs continus.

Remarque :
Une variable quantitative est considérée comme continue lorsque le nombre de modalités est élevé si bien qu'il n'est pas raisonnable de les présenter toutes dans un tableau statistique. Un regroupement des valeurs observées en intervalles s'avère alors nécessaire.

1.9 Types de variables qualitatives

1.9.1 Les variables qualitatives nominales

Une variable qualitative est dite nominale lorsqu'il n'est pas possible d'ordonner ses valeurs.

Exemples :
Dans une population d'élèves, l'ethnie et le plat préféré sont des variables qualitatives nominales.

1.9.2 les variables qualitatives ordinales

Une variable qualitative est dite ordinale lorsqu'il est possible d'ordonner ses valeurs.

Exemples :
Dans une population d'élèves le niveau d'instruction, l'année de naissance et le niveau de difficulté des exercices d'une série.

1.10 Exercices sur le chapitre 1

Exercice 1

Une enquête menée au sein d'un groupe d'élèves et portant sur le trajet "maison-école" mesuré en km a permis d'obtenir le relevé ci-dessous :

$$2,5;\ 3.5;\ 4;\ 2;\ 3;\ 4;\ 2;\ 5;\ 2,5;\ 4;\ 2;\ 4;\ 2;\ 3;\ 2;\ 3,5;\ 2;\ 2,5;\ 4.$$

1) Indique la population.
2) Indique le caractère (ou variable statistique).
3) Précise la nature du caractère.
4) Cite trois modalités du caractère.

Exercice 2

Une enquête portant sur le taux d'audience (le nombre de téléspectateurs) de quelques stations de télévision de la place lors du défilé d'un " 04 Avril" [1] a permis d'avoir les résultats suivants :

Station de télévision	E	A	C	B	F	D
Nombre de téléspectateurs	120	200	75	342	203	60

1) Quels sont les individus ?
2) Indique le caractère étudié.
3) Précise la nature de ce caractère.
4) Quelles sont les modalités du caractère ?

Exercice 3

Une étude statistique porte sur les notes d'anglais, ci-dessous indiquées, obtenues par les 13 filles d'une classe de $4^{ième}$:

$$15 - 11 - 2 - 9 - 11 - 7 - 9 - 15 - 11 - 10 - 13 - 8 - 14.$$

1) Quelle est la population sur laquelle porte cette étude ?
2) Quel est le caractère étudié ?
3) Le caractère est-il qualitatif ou quantitatif ? S'il est quantitatif, est-il continu ou discret ? Justifie.

1. Date de la célébration de l'indépendance du Sénégal

Exercice 4

Une étude ethnographique portant sur les 28 professeurs d'un collège d'enseignement moyen (CEM) a révélé :
8 professeurs d'ethnie poular,
7 professeurs d'ethnie bambara,
9 professeurs d'ethnie sérère,
4 professeurs d'ethnie wolof.
Mets une croix (X) devant les affirmations justes.
1) La population est :
- L'ensemble des travailleurs du CEM.
- L'ensemble des professeurs de Maths du CEM.
- L'ensemble des professeurs du CEM.

2)
- Tout élève du CEM est un individu.
- Les professeurs de mathématiques sont des individus.
- Le principal du CEM est une unité statistique.

3)
- Le caractère étudié est qualitatif.
- Le caractère étudié est quantitatif discret.
- Le caractère étudié n'est pas qualitatif.
- Le caractère étudié est quantitatif continu.

Propositions de solutions

Exercice 1
1) La population étudiée est l'ensemble des élèves du groupe.
2) Le caractère étudié est la distance "maison-école".
3) Le caractère est quantitatif car la distance est une grandeur quantifiable.
4) Je cite trois modalités du caractère : 4 ; $3,5$ et 5.
Exercice 2
1) Les individus sont les téléspectateurs sondés.
2) Le caractère étudié est la station télé suivie.
3) Le caractère étudié est qualitatif car la station la plus suivie n'est pas une grandeur quantifiable.
4) Les modalités du caractère étudié sont les stations télé indiquées dans le tableau.
Exercice 3
1) La population étudiée est l'ensemble des filles d'une classe de quatrième.
2) Le caractère étudié est la note d'anglais.
3) Le caractère est quantitatif. Il est discret car il prend des valeurs isolées.
Exercice 4
1) La population est :
- L'ensemble des travailleurs du CEM.
- L'ensemble des professeurs de Maths du CEM.
- L'ensemble des professeurs du CEM ($\times$).

2)
- Tout élève du CEM est un individu.
- Les professeurs de mathématiques sont des individus ($\times$).
- Le principal du CEM est une unité statistique.

3)
- Le caractère étudié est qualitatif ($\times$).
- Le caractère étudié est quantitatif discret.
- Le caractère étudié n'est pas qualitatif.
- Le caractère étudié est quantitatif continu.

CHAPITRE 2

TRAITEMENT DE DONNÉES STATISTIQUES

2.1 Recueil de données

L'étude d'un caractère au sein d'une population statistique clairement identifiée commence par le recueil des données ou informations statistiques.
Pour chaque individu de la population une seule valeur de la variable est relevée. Les méthodes utilisées sont : l'enquête, le sondage...

Exemple 1

Chacun des 40 sociétaires d'une équipe de football a été sélectionné plusieurs fois.
Voici le relevé brut du nombre de sélections par joueur :
5; 7; 17; 5; 12; 12; 12; 5; 7; 12; 8; 12; 12; 12; 13; 5; 12;
12; 13; 12; 5; 6; 7; 12; 5; 17; 17; 12; 12; 12; 12; 7; 17; 8;
5; 8; 17; 8; 8; 8.

Exemple 2

Voici le relevé des tailles (en centimètres) des 36 professeurs d'un établissement d'enseignement moyen :

155; 164; 187; 185; 176; 158; 157; 158; 185; 175; 175; 180;

185; 190; 180; 165; 175; 180; 158; 185; 172; 180; 185; 175;

198; 155; 158; 168; 190; 198; 180; 180; 175; 195; 176; 158.

Exemple 3

Dans une fabrique de chaussures un enquêteur s'intéresse à la langue maternelle des travailleurs.
Il obtient le relevé ci-dessous :

Bambara, Bambara, Bambara, Wolof, Sérère, Wolof, Wolof, Wolof, Diola,

Diola, Wolof, Diola, Wolof, Poular, Soninké, Diola, Sérère, Wolof, Soninké, Wolof, Sérère, Sérère, Diola, Poular, Balante, Poular, Sérère, Balante, Wolof, Wolof, Diola, Poular, Mancagne, Soninké, Bambara, Wolof, Sérère, Bambara, Diola, Wolof, Bambara, Wolof, Bambara, Sérère, Poular, Bambara, Soninké.

Exemple 4

A l'issue d'un devoir de mathématiques d'une classe de quatrième d'un collège, les notes suivantes ont été observées :

$$3,5; \quad 12; \quad 9; \quad 7; \quad 3,5; \quad 12; \quad 13; \quad 7; \quad 0; \quad 1;$$

$$3; \quad 5,5; \quad 17; \quad 18; \quad 20; \quad 10; \quad 6; \quad 12; \quad 9; \quad 18;$$

$$19; \quad 3; \quad 9; \quad 14; \quad 11; \quad 18; \quad 19; \quad 7; \quad 6,5; \quad 10.$$

2.2 Classement de données statistiques

2.2.1 Cas d'un caractère quantitatif discret

Dans le cas d'un caractère quantitatif discret, pour le traitement des données recueillies, on peut commencer par le classement dans l'ordre croissant ou décroissant des valeurs.

Le classement des valeurs dans l'ordre croissant de l'exemple 1 est le suivant :
5; 5; 5; 5; 5; 5; 5; 5; 7; 7; 7; 7; 8; 8; 8; 8; 8; 8; 12;
12; 12; 12; 12; 12; 12; 12; 12; 12; 12; 12; 12; 12; 12
13; 13; 17; 17; 17; 17; 17.
On a alors six modalités : 5; 7; 8; 12; 13; 17.

2.2.2 Cas d'un caractère quantitatif continu

Lorsque le nombre de modalités est assez élevé on les regroupe en intervalles appelés classes.
Une classe est un intervalle du type $[a, b[$ fermé à gauche et ouvert à droite.
La borne inférieure a appartient à la classe tandis que la borne supérieure b n'appartient pas à cette classe.
Dans ce cas, le classement dans l'ordre croissant des données facilite leur regroupement.
Pour l'exemple 2, le classement dans l'ordre croissant des valeurs recueillies est le suivant :

$$155; \quad 155; \quad 157; \quad 158; \quad 158; \quad 158; \quad 158; \quad 158; \quad 164; \quad 165; \quad 168; \quad 172$$

$$175; \quad 175; \quad 175; \quad 175; \quad 175; \quad 176; \quad 176; \quad 180; \quad 180; \quad 180; \quad 180; \quad 180$$

$$180; \quad 185; \quad 185; \quad 185; \quad 185; \quad 185; \quad 187; \quad 190; \quad 190; \quad 195; \quad 198; \quad 198.$$

On peut obtenir par exemple un regroupement selon les 5 classes suivantes :
$[1,50; 1,60[$, $[1,60; 1,70[$, $[1,70; 1,80[$, $[1,80; 1,90[$ et $[1,90; 2[$.

Amplitude d'une classe

Définition
On appelle amplitude de la classe $[a, b[$ la différence entre sa borne supérieure b et sa borne inférieure a : $b - a$.

Centre d'une classe

Définition
On appelle centre de la classe $[a, b[$ la demi-somme de ses deux bornes a et b : $\dfrac{a+b}{2}$.
Exemple 5 :
Soit la classe $[3, 7[$. Son amplitude est : $7 - 3 = 4$ et son centre est : $\dfrac{3+7}{2} = 5$.

Remarque
Il faut toujours proposer la borne inférieure de la première classe ainsi que l'amplitude commune des différentes classes pour harmoniser le regroupement en classes d'égale amplitude.

2.2.3 Cas d'un caractère qualitatif

Dans le cas d'un caractère qualitatif, un dépouillement permet de regrouper les valeurs identiques.
Pour l'exemple 3 on obtient :

Wolof, Wolof, Wolof, Wolof, Wolof, Wolof, Wolof, Wolof, Wolof, Wolof

Wolof, Wolof, Wolof, Mankagne, Soninké, Soninké, Soninké, Soninké, Balante,

Balante, Poular, Poular, Poular, Poular, Poular, Bambara, Bambara, Bambara

Bambara, Bambara, Bambara, Bambara, Bambara, Diola, Diola, Diola, Diola

Diola, Diola, Diola, Sérère, Sérère, Sérère, Sérère, Sérère, Sérère, Sérère.

On obtient les modalités suivantes :
Wolof ; Mankagne ; Poular ; Bambara ; Diola ; Sérère ; Soninké ; Balante.

2.3 Effectif partiel - Effectif total

Définition
On appelle effectif d'une modalité le nombre de fois que cette modalité est observée dans la population.
Définition
L'effectif d'une classe $[a, b[$ est le nombre de fois qu'un élément de cet intervalle est observé dans la population.

On note n_1 l'effectif de la modalité x_1 ou de la classe $[a_1, b_1[$, n_2 l'effectif de la modalité x_2 ou de la classe $[a_2, b_2[$, etc...

Définition

On appelle effectif d'une population le nombre d'individus de cette population.
On le note N.

Remarque

La somme des effectifs de toutes les modalités est égale à l'effectif N de la population :

$$N = n_1 + n_2 + n_3 + ...$$

Pour cette raison, N est appelé effectif total et par conséquent chaque effectif d'une modalité est appelé effectif partiel. Ainsi, n_1, n_2, n_3,... sont des effectifs partiels.

2.4 Fréquence d'une modalité - d'une classe

Définition

On appelle fréquence d'une modalité (ou d'une classe) le rapport de son effectif partiel n par l'effectif total N de la population. On note

$$f_1 = \frac{n_1}{N}, \quad f_2 = \frac{n_2}{N}, \quad f_3 = \frac{n_3}{N}, ...$$

n_1 est l'effectif partiel de la modalité x_1, n_2 celui de x_2 ... et N l'effectif total.
La fréquence d'une modalité permet d'avoir une idée de la représentativité de cette modalité.

Exemple 6

— Dans l'exemple 1, la fréquence de la modalité 5 est $\frac{8}{40}$ ou $0,2$ et celle de la modalité 8 est $\frac{6}{40}$ ou $0,15$.

— Dans l'exemple 2, la fréquence de la classe $[1,60; 1,70[$ est $\frac{3}{36}$ ou environ $0,083$ et celle de la classe $[1,80; 1,90[$ est $\frac{12}{36}$ ou environ $0,333$.

— Dans l'exemple 3, la fréquence de la modalité $Wolof$ est :$\frac{13}{47}$ ou environ $0,277$ et celle de la modalité *Bambara* est $\frac{5}{47}$ ou environ $0,106$.

Remarque :

Une fréquence est donnée par un nombre compris entre 0 et 1.

Pourcentage

On peut exprimer une fréquence f en pourcentage.
Par exemple pour la modalité x_1, d'effectif n_1, sa fréquence en pourcentages notée p_1 est donnée par :

$$p_1 = \frac{n_1}{N} \times 100 \text{ ou } p_1 = f_1 \times 100$$

Exemple 7

— Dans l'exemple 1, la fréquence de la modalité 5 est $\frac{8}{40} \times 100\% = 20\%$ et celle de la modalité 8 est $\frac{6}{40} \times 100\% = 15\%$.

— Dans l'exemple 2, la fréquence de la classe $[1,60;1,70[$ est $\frac{3}{36} \times 100\%$ ou environ $8,333\%$ et celle de la classe $[1,80;1,90[$ est $\frac{12}{36} \times 100\%$ ou environ $33,333\%$.

— Dans l'exemple 3, la fréquence de la modalité *Wolof* est $\frac{13}{47} \times 100\%$ ou environ $27,659\%$ et celle de la modalité *Bambara* est $\frac{5}{47} \times 100\%$ ou $10,638\%$.

2.5 Effectifs Cumulés d'une modalité - d'une classe

Lorsque le caractère étudié est quantitatif, il existe deux types d'effectifs cumulés :
les effectifs cumulés croissants (ECC) et les effectifs cumulés décroissants (ECD).
Lorsque le caractère est qualitatif le cumul des effectifs ou des fréquences n'a pas de sens.

2.5.1 Effectif cumulé croissant d'une modalité

Définition

On appelle effectif cumulé croissant d'une modalité, le nombre d'individus pour lesquels le caractère prend une valeur inférieure ou égale à cette modalité.

Exemple 8

Dans l'exemple 1, l'ECC de la modalité 5 est 8 et celui de la modalité 12 est $8+4+6+15 = 33$.

2.5.2 Effectif cumulé décroissant d'une modalité

Définition

On appelle effectif cumulé décroissant d'une modalité le nombre d'individus pour lesquels le caractère prend une valeur supérieure ou égale à cette modalité.

Exemple 9

Dans l'exemple 1, l'ECD de la modalité 5 est $5 + 2 + 15 + 6 + 4 + 8 = 40$ et celui de la modalité 8 est $5 + 2 + 15 + 6 = 28$.

2.5.3 Effectif Cumulé d'une classe

Définition

On appelle effectif cumulé croissant d'une classe $[a, b[$ le nombre d'individus pour lesquels le caractère prend une valeur strictement inférieure à b.

Exemple 10
Dans l'exemple 2, l'ECC de la classe $[1,60; 1,70[$ est $8+3 = 11$ et celui de la classe $[1,80; 1,90[$ est $8 + 3 + 8 + 12 = 31$.

Définition
On appelle effectif cumulé décroissant d'une classe $[a, b[$, le nombre d'individus pour lesquels le caractère prend une valeur supérieure ou égale à a.

Exemple11 Dans l'exemple 2, l'ECD de la classe $[1,60; 1,70[$ est : $5 + 12 + 8 + 3 = 28$ et celui de la classe $[1,80; 1,90[$ est $5 + 12 = 17$.

CHAPITRE 3

SÉRIE STATISTIQUE

3.1 Définition

On appelle série statistique l'ensemble des couples (x_1, n_1), (x_2, n_2), (x_3, n_3), etc...
Dans le cas d'un caractère quantitatif continu, on appelle série statistique l'ensemble des couples $([a_1, b_1[, n_1), ([a_2, b_2[, n_2)$, $([a_3, b_3[n_3)$, etc... Il est possible de définir la série statistique en remplaçant chaque classe par son centre.

Exemple 1
Voici les séries correspondant respectivement aux exemples 1, 2 et 3 donnés dans le chapitre 2.

$\Big\{(5, 8), (7, 4), (8, 6), (12, 14), (13, 2), (17, 5)\Big\}$.

$\Big\{([150, 160[, 8), ([160, 170[, 3), ([170, 180[, 8), ([180, 190[, 12), ([190, 200[, 5)\Big\}$.

$\Big\{$(Wolof, 13), (Poular, 5), (Soninké, 4), (Balante, 2), (Bambara, 8), (Sérère, 7), (Diola, 7), (Mancagne, 1)$\Big\}$.

3.2 Les tableaux statistiques

Une série statistique est souvent donnée à travers des tableaux de correspondance appelés tableaux statistiques.

3.2.1 Tableau des effectifs d'une série

Le tableau des effectifs est constitué des différentes modalités ou classes avec les effectifs partiels correspondants.

Exemple 2
Tableau des effectifs de l'exemple 1 du chapitre 2
On peut adopter l'une des dispositions suivantes :

Horizontalement

Modalités	5	7	8	12	13	17	Total
Effectifs	8	4	6	15	2	5	40

Nombre de sélections=modalités
Nombre de joueurs =Effectifs

Verticalement

Modalités	Effectifs
5	8
7	4
8	6
12	15
13	2
17	5
Total	40

Tableau des effectifs de l'exemple 2 du chapitre 2

Horizontalement

Classes	$[1,50;1,60[$	$[1,60;1,70[$	$[1,70;1,80[$	$[1,80;1,90[$	$[1,90;2,00[$	Total
Effectifs	8	3	8	12	5	36

Verticalement

Classes	Effectifs
$[1,50;1,60[$	8
$[1,60;1,70[$	3
$[1,70;1,80[$	8
$[1,80;1,90[$	12
$[1,90;2,00[$	5
Total	36

Tableau des effectifs de l'exemple 3 du chapitre 2
Horizontalement

Ethnies	Wolof	Poular	soninké	Balante	Bambara	Sérère	Diola	Mancagne	Total
Effectifs	13	5	4	2	8	7	7	1	47

Verticalement

Ethnies	Effectifs
Wolof	13
Poular	5
Soninké	4
Balante	2
Bambara	8
Sérère	7
Diola	7
Mancagne	1
Total	47

NB
Nous adopterons pour la suite une disposition verticale des tableaux.

3.2.2 Tableau des fréquences

On peut ajouter au tableau des effectifs les fréquences et les fréquences en pourcentages correspondant aux différentes modalités.
Exemple 3
Tableau des fréquences de l'exemple 1 du chapitre 2

Modalités	Effectifs	Fréquences	Fréquences(%)
5	8	$0,2$	20
7	4	$0,1$	10
8	6	$0,15$	15
12	15	$0,375$	$37,5$
13	2	$0,05$	5
17	5	$0,125$	$12,5$
Total	40	1	100

Tableau des fréquences de l'exemple 2 du chapitre 2

Classes	Effectifs	Fréquences	Fréquences(%)
$[1,50;1,60[$	8	$0,222$	$22,2$
$[1,60;1,70[$	3	$0,083$	$8,3$
$[1,70;1,80[$	8	$0,222$	$22,2$
$[1,80;1,90[$	12	$0,334$	$33,4$
$[1,90;2,00[$	5	$0,139$	$13,9$
Total	36	1	100

Tableau des fréquences de l'exemple 3 du chapitre 2

Ethnie	Nombre de joueurs	Fréquences	Fréquences(%)
Wolof	13	$0,277$	$27,7$
Poular	5	$0,106$	$10,6$
Soninké	4	$0,085$	$8,5$
Balante	2	$0,043$	$4,3$
Bambara	5	$0,170$	17
Sérère	8	$0,149$	$14,9$
Diola	7	$0,149$	$14,9$
Macagne	2	$0,021$	$2,1$
Total :	47	1	100

3.2.3 Tableaux des effectifs cumulés

On distingue le tableau des effectifs cumulés croissants(ECC) et celui des effectifs cumulés décroissants(ECD)

Exemple 4

Tableau des effectifs cumulés croissants de l'exemple 1 du chapitre 2

Modalités	Effectifs	Effectifs Cumulés Croissants
5	8	8
7	4	12
8	6	18
12	15	33
13	2	35
17	5	40
Total	40	

Tableaux des effectifs cumulés décroissants de l'exemple 1 du chapitre 2

Modalités	Effectifs	Effectifs Cumulés Décroissants
5	8	40
7	4	32
8	6	28
12	15	22
13	2	7
17	5	5
Total	40	

Tableau des effectifs cumulés croissants de l'exemple 2 du chapitre 2

Classes	Effectifs	Effectifs Cumulés Croissants
$[150, 160[$	8	8
$[160, 170[$	3	11
$[170, 180[$	8	19
$[180, 190[$	12	31
$[190, 200[$	5	36
Total	36	

Tableau des effectifs cumulés décroissants de l'exemple 2 du chapitre 2

Classes	Effectifs	Effectifs Cumulés Décroissants
$[150, 160[$	8	36
$[160, 170[$	3	28
$[170, 180[$	8	25
$[180, 190[$	12	17
$[190, 200[$	5	5
Total	36	

NB :
Toutefois le tableau des effectifs cumulés est celui qui contient à la fois les ECC et les ECD.

3.2.4 Tableau des fréquences cumulées

De la même manière que pour les effectifs on peut cumuler les fréquences d'une série à caractère quantitatif

Exemple 5
Tableau des fréquences cumulées croissantes (FCC) de l'exemple 1 du chapitre 2

Modalités	Effectifs	Fréquences	Fréquences Cumulées Croissantes
5	8	$0,2$	$0,2$
7	4	$0,1$	$0,3$
8	6	$0,15$	$0,45$
12	15	$0,375$	$0,825$
13	2	$0,05$	$0,875$
17	5	$0,125$	1
Total	40	1	

Tableau des fréquences cumulées décroissantes (FCD) de l'exemple 1 du chapitre 2

Modalités	Effectifs	Fréquences	Fréquences Cumulées Décroissantes
5	8	$0,2$	1
7	4	$0,1$	$0,8$
8	6	$0,15$	$0,7$
12	15	$0,375$	$0,55$
13	2	$0,05$	$0,175$
17	5	$0,125$	$0,125$
Total	40	1	

Tableau des fréquences cumulées croissantes de l'exemple 2 du chapitre 2

Classes	Effectifs	Fréquences	Fréquences Cumulées Croissantes
$[150, 160[$	8	$0,222$	$0,222$
$[160, 170[$	3	$0,083$	$0,305$
$[170, 180[$	8	$0,222$	$0,527$
$[180, 190[$	12	$0,334$	$0,861$
$[190, 200[$	5	$0,139$	1
Total	36	1	

Tableau des fréquences cumulées décroissantes de l'exemple 2 du chapitre 2

Classes	Effectifs	Fréquences	Fréquences Cumulées Décroissantes
$[150, 160[$	8	$0,222$	1
$[160, 170[$	3	$0,083$	$0,778$
$[170, 180[$	8	$0,222$	$0,695$
$[180, 190[$	12	$0,334$	$0,473$
$[190, 200[$	5	$0,139$	$0,139$
Total	36	1	

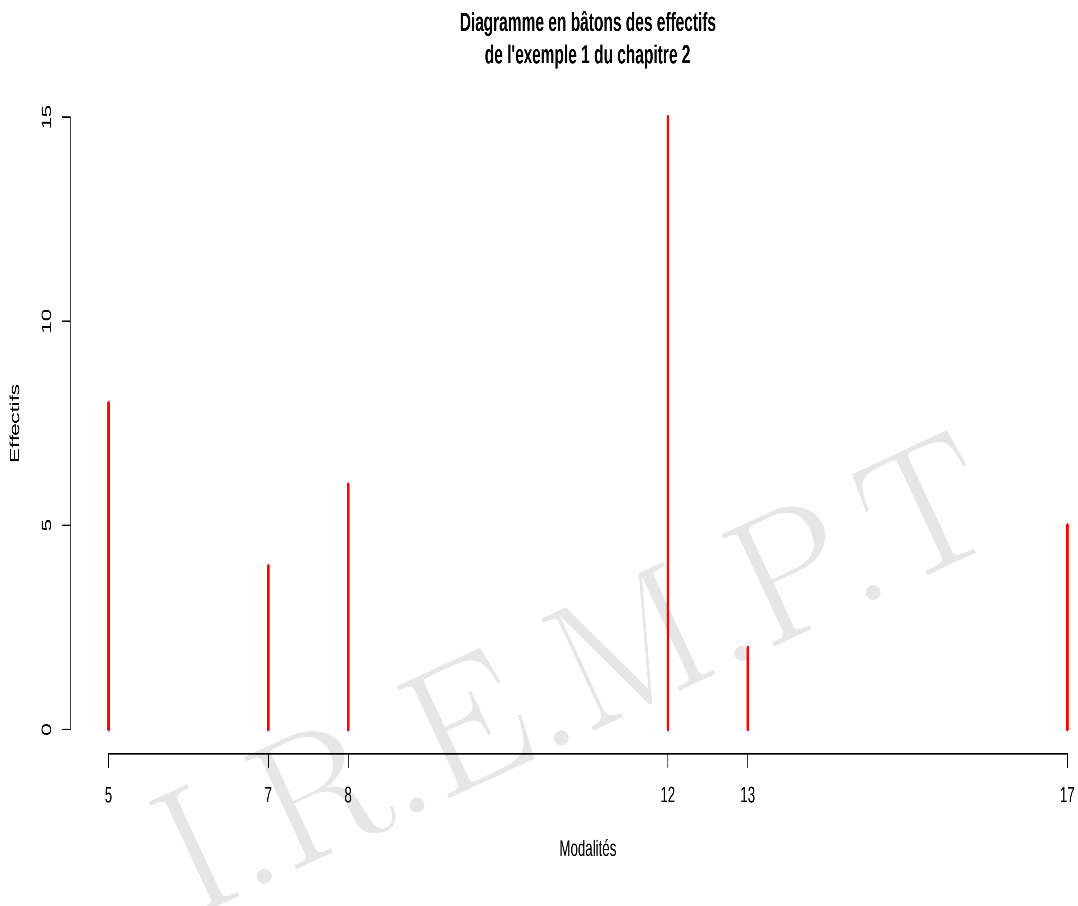
3.3 Représentation graphique d'une série statistique

3.3.1 Diagramme en bâtons des effectifs - des fréquences

Le diagramme en bâtons est une forme de représentation des séries à caractère quantitatif discret.
Dans un repère d'axes perpendiculaires, on porte en abscisses les modalités et en ordonnées les effectifs (ou les fréquences) correspondants.
A chaque modalité x_i, portée en abscisse, on fait correspondre un segment vertical (appelé

bâton) de longueur proportionnelle à l'effectif n_i ou à la fréquence f_i de cette modalité.

Exemple 6



3.3.2 Histogramme

L'histogramme est une forme de représentation des séries à caractère quantitatif continu. Mais pour un caractère quantitatif discret comportant plusieurs modalités, on peut faire un regroupement en classes et le représenter comme un caractère continu.

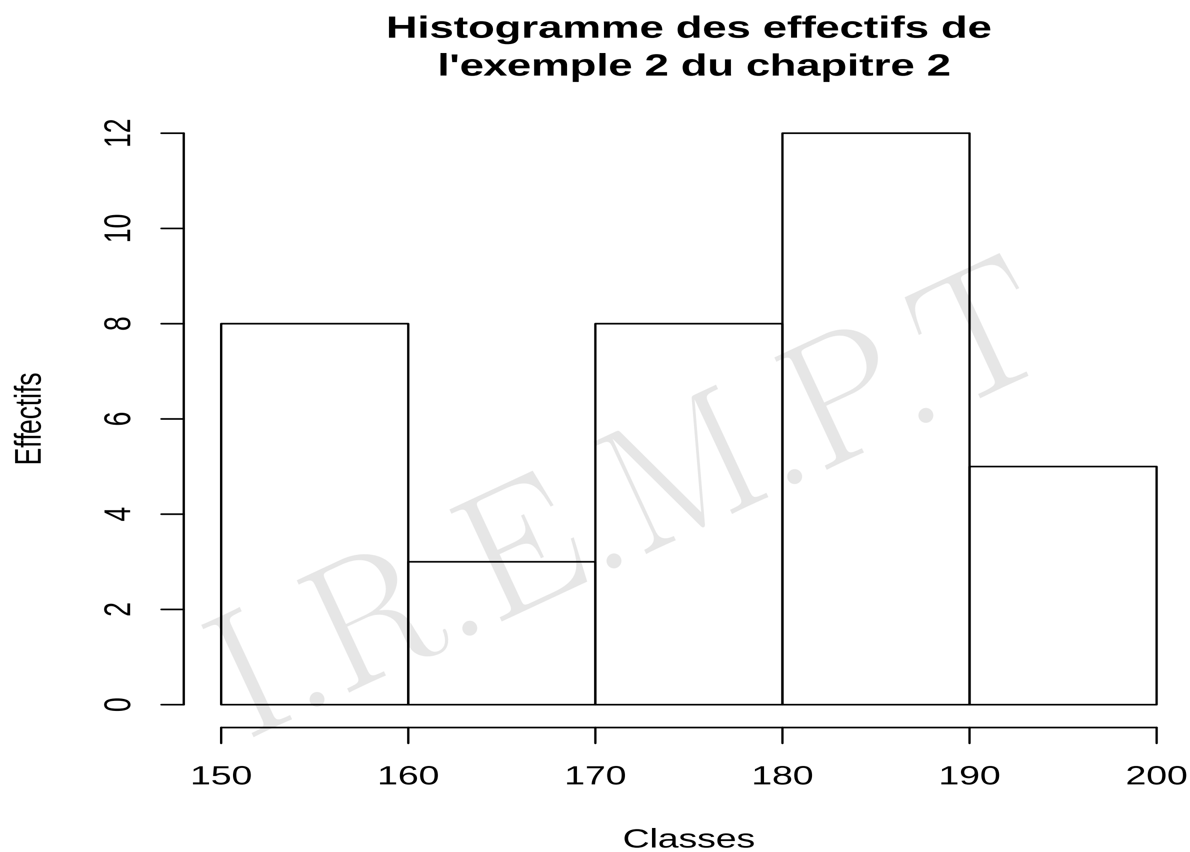
Histogramme des effectifs

Dans un repère d'axes perpendiculaires, on met les classes en abscisses et les effectifs partiels correspondants en ordonnées.
A chaque couple (classe ; effectif partiel) correspond un rectangle dont les dimensions sont l'amplitude et l'effectif de la classe.

Histogramme des fréquences

Dans un repère d'axes perpendiculaires, on met les classes en abscisses et les fréquences correspondantes en ordonnées.
A chaque couple (classe , fréquence) correspond un rectangle dont les dimensions sont l'amplitude et la fréquence de la classe.

Exemple 7



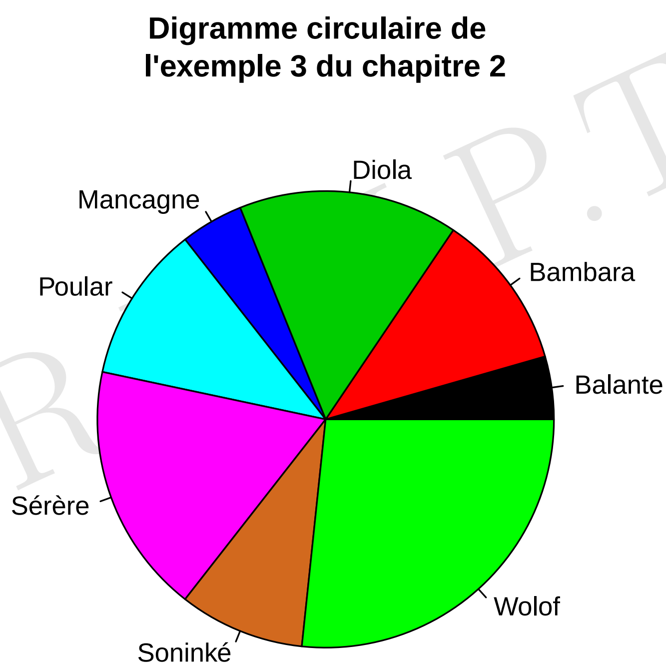
3.3.3 Diagramme circulaire

Le diagramme circulaire est une représentation d'un caractère qualitatif par un cercle. Chaque modalité est représentée par un secteur angulaire au centre dont la mesure est proportionnelle à l'effectif de cette modalité.
A l'effectif total N correspond l'angle au centre du cercle qui mesure 360 degrés.
A l'effectif partiel n_1 correspond un angle au centre du cercle de mesure $\alpha_1 = \dfrac{n_1 \times 360°}{N}$.
A l'ffectif partiel n_2 correspond un angle au centre de mesure $\alpha_2 = \dfrac{n_2 \times 360°}{N}$, etc...

Tableau des effectifs de l'exemple 3 du chapitre 2, avec les angles associés.

Ethnies	Effectifs	Angle (°)
Wolof	12	96
Poular	5	40
Soninké	4	32
Balante	2	16
Bambara	5	40
Sérère	8	64
Diola	7	56
Mancagne	2	16
Total	45	360

Le diagramme circulaire de cette série est le suivant:



3.3.4 Diagramme semi-circulaire

En utilisant un demi-cercle à la place d'un cercle, on obtient une représentation appelée diagramme semi-circulaire.

A l'effectif total N correspond un angle de mesure de 180 degrés.
A l'effectif partiel n_1 correspond un angle de mesure $\alpha_1 = \dfrac{n_1 \times 180°}{N}$.
A l'effectif partiel n_2 correspond un angle de mesure $\alpha_2 = \dfrac{n_2 \times 180°}{N}$, etc...
Tableau des effectifs de l'exemple 3 du chapitre 2, avec les angles associés .

Ethnies	Effectifs	Angle(°)
Wolof	12	48
Poular	5	20
Soninké	4	16
Balante	2	8
Bambara	5	20
Sérère	8	32
Diola	7	28
Mancagne	2	8
Total	45	180

Voici le diagramme semi-circulaire des effectifs de la série ci-dessus:

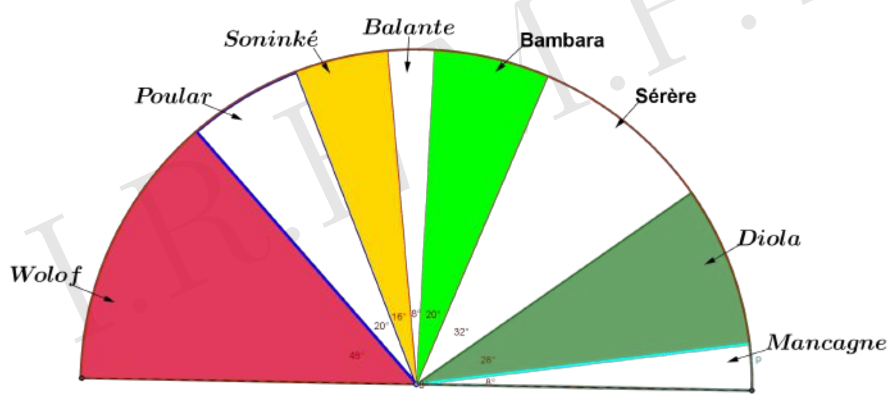


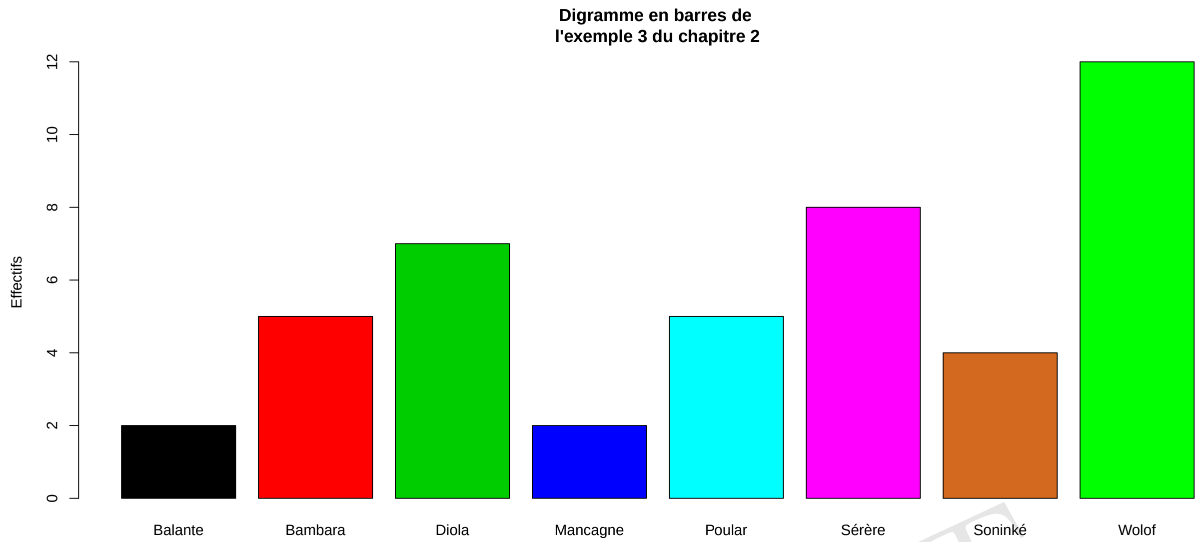
Diagramme semi-circulaire de l'exemple 3

Remarque
On obtient les diagrammes des fréquences de la même manière que pour les effectifs.

3.3.5 Diagrammes à bandes ou en barres

Le diagramme à bandes encore appelé diagramme en barres est une représentation d'un caractère qualitatif.
Les modalités sont représentées par des rectangles(bandes) de base constante et dont les hauteurs sont proportionnelles aux effectifs de ces modalités.

Voici le diagramme en barres des effectifs de l'exemple 3 du chapitre 2



3.3.6 Diagrammes cumulatifs des séries à caractère quantitatif

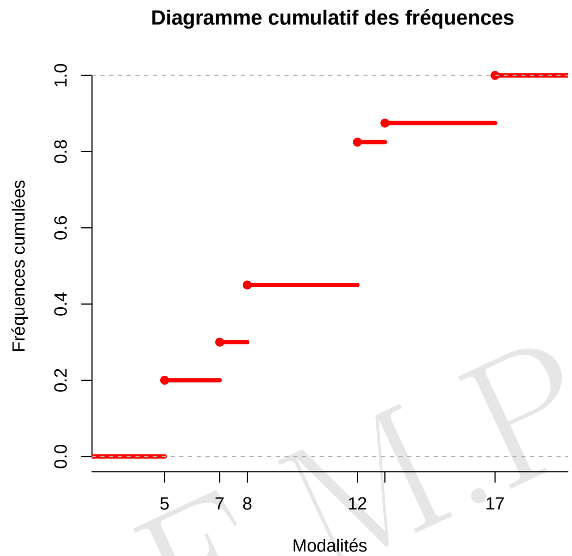
Cas discret

Considérons le tableau statistique des fréquences cumulées croissantes (FCC) de l'exemple 1 du chapitre 2.

Modalités	Effectifs	Fréquences	Fréquences Cumulées Croissantes
5	8	$0,2$	$0,2$
7	4	$0,1$	$0,3$
8	6	$0,15$	$0,45$
12	15	$0,375$	$0,825$
13	2	$0,05$	$0,875$
17	5	$0,125$	1
Total	40	1	

Pour représenter le diagramme cumulatif des fréquences, on détermine d'abord la fonction cumulative F définie par :

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 5 \\ 0,2 & \text{si} \quad 5 \leq x < 7 \\ 0,3 & \text{si} \quad 7 \leq x < 8 \\ 0,45 & \text{si} \quad 8 \leq x < 12 \\ 0,825 & \text{si} \quad 12 \leq x < 13 \\ 0,875 & \text{si} \quad 13 \leq x < 17 \\ 1 & \text{si} \quad x \geq 17 \end{cases}$$



Cas continu

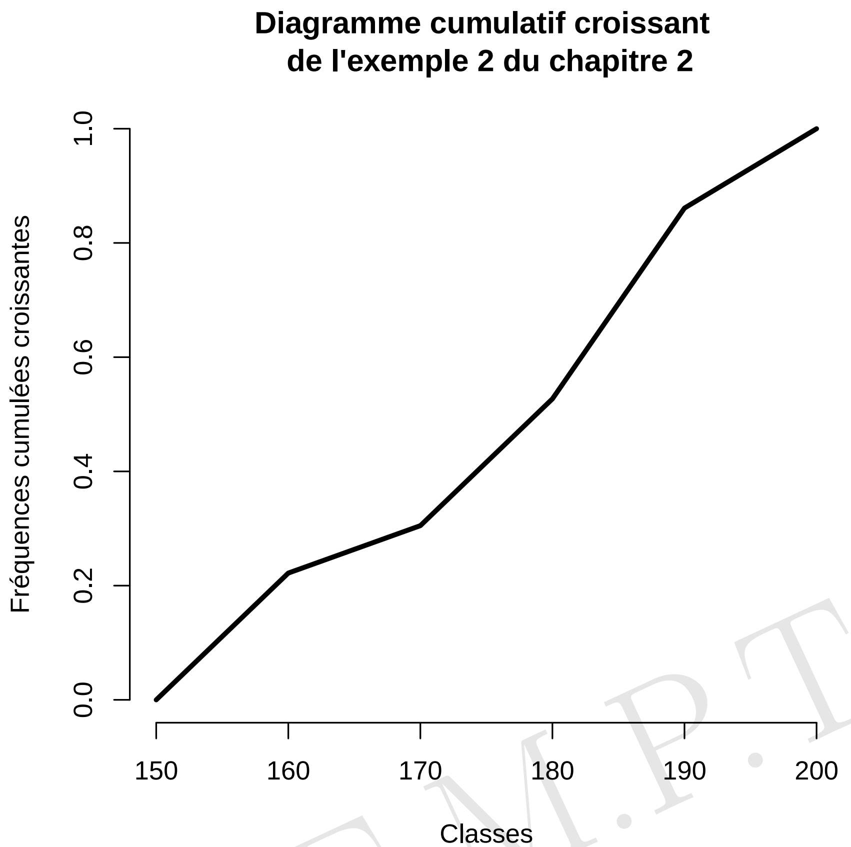
On distingue le digramme des fréquences cumulées croissantes et celui des fréquences cumulées décroissantes.

• Diagramme des fréquences cumulées croissantes.

Considérons le tableau des fréquences cumulées croissantes de l'exemple 2 du chapitre 2.

Classes	Effectifs	Fréquences	Fréquences Cumulées Croissantes
$[150, 160[$	8	$0,222$	$0,222$
$[160, 170[$	3	$0,083$	$0,305$
$[170, 180[$	8	$0,222$	$0,527$
$[180, 190[$	12	$0,334$	$0,861$
$[190, 200[$	5	$0,139$	1
Total	36	1	

Le diagramme cumulatif est la ligne polygonale qui joint les points de coordonnées : $(150; 0), (160; 0.222), (170; 0.305), (180; 0.527), (190; 0.861)$ et $(200; 1)$.



• Diagramme des fréquences cumulées décroissantes.
Considérons le tableau des fréquences cumulées décroissantes de l'exemple 2 du chapitre 2.

Classes	Effectifs	Fréquences	Fréquences Cumulées Décroissantes
$[150, 160[$	8	$0,222$	1
$[160, 170[$	3	$0,083$	$0,778$
$[170, 180[$	8	$0,222$	$0,695$
$[180, 190[$	12	$0,334$	$0,473$
$[190, 200[$	5	$0,139$	$0,139$
Total	36	1	

Le diagramme cumulatif est la ligne polygonale qui joint les points de coordonnées :
$(150; 1), (160; 0.778), (170; 0.695), (180; 0.473), (190; 0139)$ et $(200; 0)$.

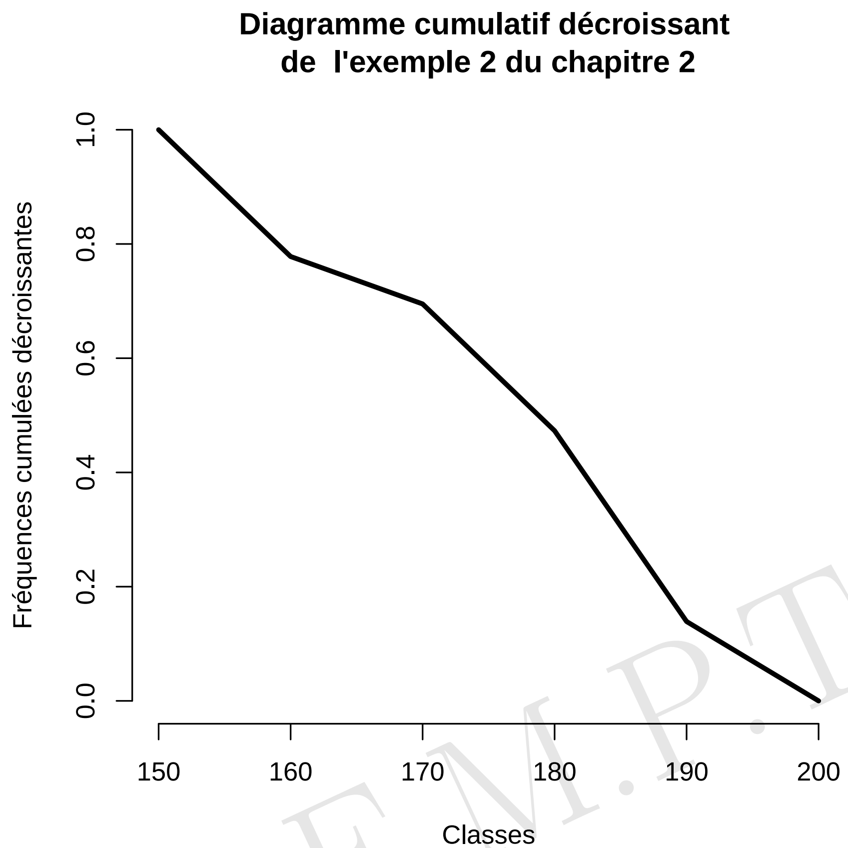
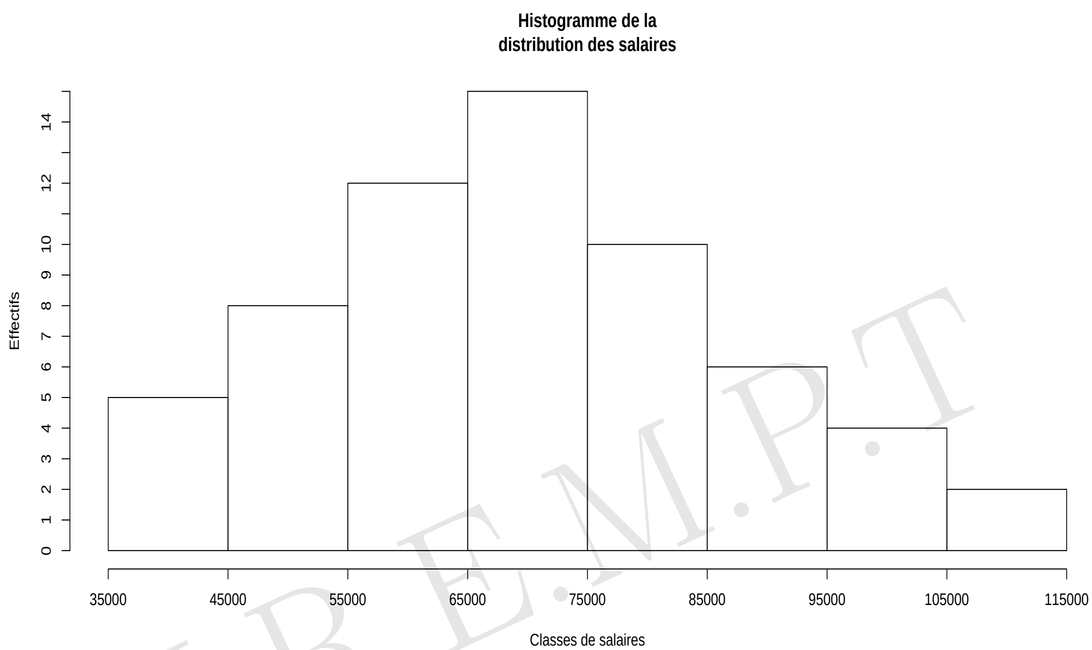


Diagramme cumulatif décroissant
de l'exemple 2 du chapitre 2
Fréquences cumulées décroissantes
0.0
0.2
0.4
0.6
0.8
1.0
150
160
170
180
190
200
Classes

3.4 Exercices sur les chapitres 2 et 3

Exercice 1

Le diagramme ci-dessous représente la répartition des salaires en francs CFA des employés d'une entreprise de confection de blouses scolaires.



1) Indique la nature du caractère étudié.

2) *a*) Dresser le tableau statistique de la série représentée en y mettant les fréquences.

b) Ajoute à ce tableau le centre de chaque classe, les effectifs cumulés et les fréquences cumulées en pourcentages.

3) Utilise les valeurs du tableau précédent pour répondre aux questions suivantes :

a) Combien d'ouvriers gagnent moins de 75.000 F ?

b) Combien d'ouvriers gagnent au moins 55.000 F ?

c) Quel est le pourcentage d'ouvriers qui gagnent au moins 65.000 F et moins de 95.000 F ?

4) Construire le polygone cumulatif des effectifs.

Exercice 2

Le tableau ci-dessous donne la répartition des élèves d'une classe de 3*e* selon le nombre d'absences au cours d'un semestre donné.

Nombre d'absences	Nombre d'élèves
0	24
1	9
2	4
3	3
5	8
6	8
8	6
9	4
10	6

1) Complète le tableau en y ajoutant les effectifs cumulés et les fréquences cumulées de chaque modalité.
2) Quel est le pourcentage des élèves qui ne se sont jamais absentés ?
3) Combien d'élèves se sont absentés au moins 3 fois ?
4) Combien d'élèves se sont absentés au moins 3 fois et au plus 8 fois ?
5) Dessine le diagramme en bâtons des effectifs et celui des effectifs cumulés croissants.

Exercice 3

Les élèves d'une classe ont obtenu les notes suivantes lors d'un contrôle de mathématiques :
$11 - 8 - 11 - 12 - 13 - 11 - 6 - 10 - 14 - 10 - 8 - 11 - 13 - 5 - 11 - 9 - 11 - 11 - 10 - 12 - 6 - 13 - 10 - 9 - 10 - 12 - 8 - 9 - 9 - 10$
1) Dresse le tableau statistique de cette série de notes (modalités et effectifs).
2) Dessine le diagramme en bâtons de la série.
3) Complète le tableau précédent en y mettant les effectifs cumulés croissants (ECC) et les effectifs cumulés décroissants (ECD).
4) Quel est le pourcentage d'élèves qui ont la moyenne ?
5) Quel est le pourcentage d'élèves qui ont moins de 9 ?
6) Quel est le pourcentage d'élèves qui ont plus de 11 ?

Exercice 4

On a relevé le poids (en "kilogramme") de 30 élèves d'une classe de troisième :

Poids(kg)	Effectif	ECC	ECD
$[a, 65[$	2		
$[65, 70[$	4		
$[70, 75[$	8		
$[75, 80[$	x		
$[80, b[$	10		
Total			

1) Détermine a et b sachant que les classes sont de même amplitude.
2) Détermine x puis complète le tableau ci-dessus.

3) Combien d'élèves pèsent moins de 75 kg ?
4) Combien d'élèves pèsent au moins 65 kg ?
5) Représente l'histogramme des effectifs.

Exercice 5

On étudie la situation matrimoniale des enseignants d'un établissement scolaire. Les résultats sont donnés dans le tableau ci-dessous.

Modalités	Marié	Veuf	Célibataire	Divorcé
Effectifs	35	5	40	10

1) Construis les diagrammes en barres, circulaire et semi circulaire de la situation matrimoniale.
2) Détermine le pourcentage d'enseignants non mariés.

Proposition de solutions

Exercice 1

1) Le caractère étudié est quantitatif continu.

2)$a) - b)$

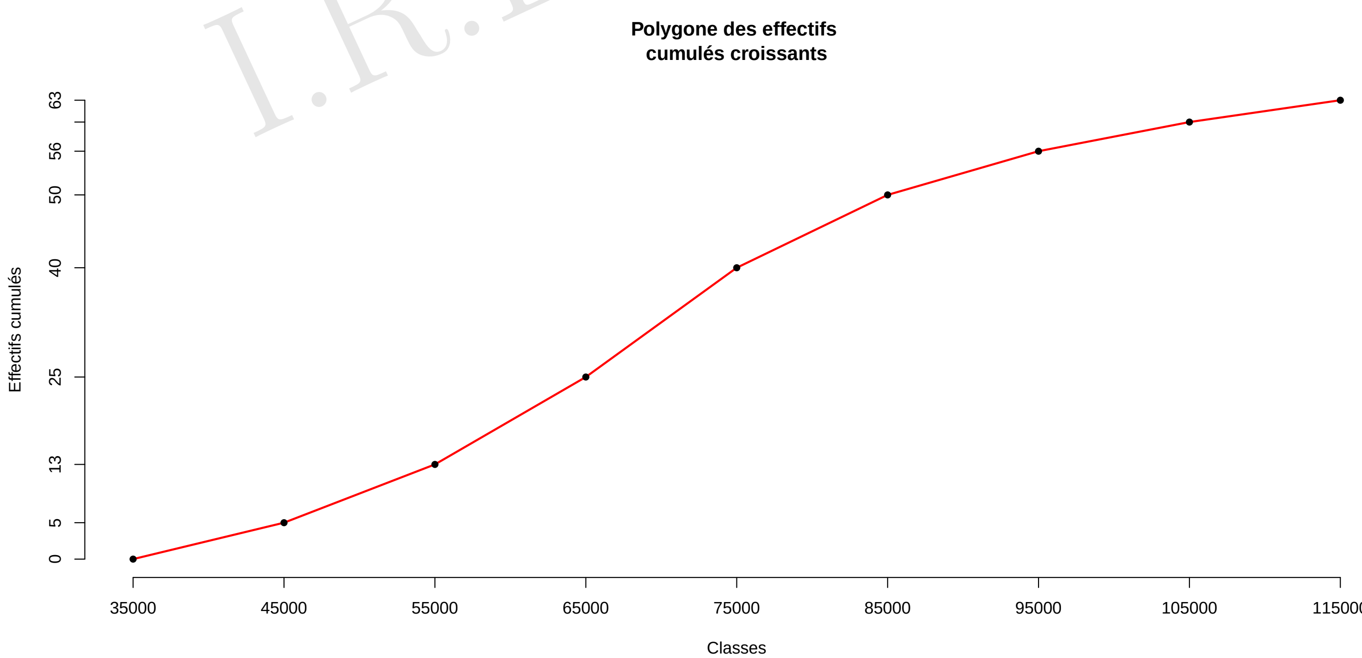
Classes	Effectifs	Fréquences (%)	ECC	ECD
[35000,45000[	5	7,94	5	63
[45000,55000[	8	12,07	13	58
[55000,65000[	12	19,05	25	50
[65000,75000[	15	23,81	40	38
[75000,85000[	10	16,87	50	23
[85000,95000[	6	9,52	56	13
[95000,105000[	4	6,35	60	7
[105000,115000[	3	4,76	63	3
Total	63	100	×	×

3)$a)$ Le nombre d'ouvriers qui gagnent moins de 75000 est donné par l'effectif cumulé croissant de la classe $[65000, 75000[$ qui est égal à 40.

$b)$ Le nombre d'ouvriers qui gagnent au moins 55000 est donné par l'effectif cumulé décroissant de la classe $[55000, 65000[$ qui est égal à 50.

$c)$ Le pourcentage d'ouvriers qui gagnent au moins 65000 et moins de 95000 est la somme des pourcentages des classes $[65000, 75000[$, $[75000, 85000[$ et $[85000, 95000[$ qui est égal à $50,2\%$.

4)



Exercice 2

1)

Nombre d'absences	Nombre d'élèves	ECC	ECD	FCC	FCD
0	24	24	72	0,33	1
1	9	33	48	0,46	0,67
2	4	37	39	0,51	0,54
3	3	40	35	0,56	0,49
5	8	48	32	0,67	0,44
6	8	56	24	0,78	0,33
8	6	62	16	0,86	0,22
9	4	66	10	0,92	0,14
10	6	72	6	1	0,08
Total	72	$\times$	$\times$	$\times$	$\times$

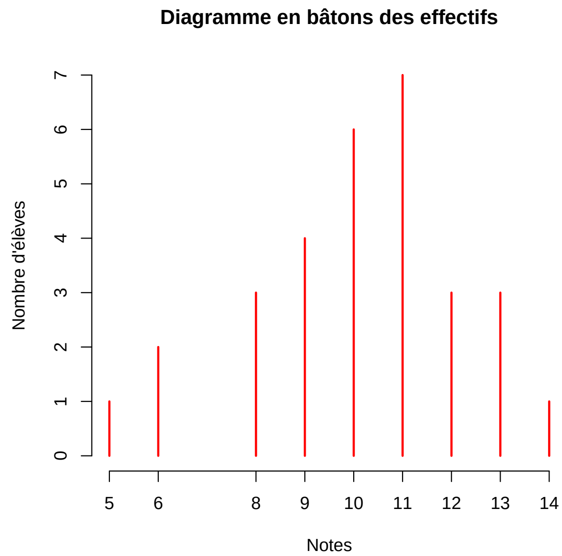
2) Le pourcentage des élèves qui ne se sont jamais absentés est égal à la fréquence en pourcentage de la modalité 0, soit 33%.
3) Le nombre d'élèves qui se sont absentés au moins 3 fois est égal à l'effectif cumulé décroissant de la modalité 3, soit 35.
4) Le nombre d'élèves qui se sont absentés au moins 3 fois et au plus 8 fois est égal à la somme des effectifs des modalités de 3 à 8, soit 25.

Exercice 3

1)

Notes	Effectifs	Effectifs cumulés croissants	Effectifs cumulés décroissants
5	1	1	30
6	2	3	29
8	3	6	27
9	4	10	23
10	6	16	17
11	7	23	10
12	3	26	7
13	3	29	4
14	1	30	1
Total	30	$\times$	$\times$

2)



3) Le nombre d'élèves qui ont la moyenne (c'est-à-dire au moins 10) correspond à l'effectif cumulé décroissant de la modalité 10 soit 17. Le pourcentage correspondant est alors $\frac{17 \times 100}{30} = 56,67\%$.
4) Le nombre d'élèves qui ont moins de 9 correspond à l'effectif cumulé croissant de la modalité 8 soit 6. Le pourcentage correspondant est alors $\frac{6 \times 100}{30} = 20\%$.
5) Le nombre d'élèves qui ont plus de 11 correspond à l'effectif cumulé décroissant de la modalité 12 soit 7. Le pourcentage correspondant est alors $\frac{7 \times 100}{30} = 23,33\%$.

Exercice 4

1) Les classes sont d'amplitude 5. Donc on aura $65 - a = 5$ et $b - 80 = 5$. Ainsi $a = 60$ et $b = 85$.
2) La somme des effectifs est $2 + 4 + 8 + x + 10$. Comme l'effectif total est égal à 30, on a $2 + 4 + 8 + x + 10 = 30$. Donc $x = 6$.

Poids(kg)	Effectif	ECC	ECD
$[60, 65[$	2	2	30
$[65, 70[$	4	6	28
$[70, 75[$	8	14	24
$[75, 80[$	6	20	16
$[80, 85[$	10	30	10
Total	30	$\times$	$\times$

3) Le nombre d'élèves pesant moins de 75 kg est égal à la somme des effectifs des classes $[60, 65[$, $[65, 70[$ et $[70, 75[$. Soit $2 + 4 + 8 = 14$.

4) Le nombre d'élèves pesant au moins 65 kg est égal à l'effectif cumulé décroissant de la classe $[65, 70[$; soit 28.

5)

CHAPITRE 4

LES CARACTÉRISTIQUES DE POSITION

Introduction

Lorsqu'on collecte des données, on utilise des tableaux statistiques pour les organiser et des représentations graphiques pour les visualiser. Mais cela ne permet de faire qu'une analyse sommaire. Pour mesurer davantage l'information contenue dans les données brutes, on définit des paramètres dits de position qui permettent de résumer les séries statistiques. Les caractéristiques (ou paramètres) usuelles sont : le mode, la moyenne arithmétique, la médiane et les quartiles.

4.1 Le mode

4.1.1 Définition

On appelle mode ou dominante d'une variable qualitative ou quantitative discrète la modalité ayant l'effectif le plus grand.
Pour une variable quantitative continue, on parle de classe modale c'est-à-dire la classe ayant l'effectif le plus grand. Dans ce cas on convient de prendre comme mode le centre de la classe modale.
Remarques :

• On peut définir le mode en remplaçant les effectifs par les fréquences.

• Une série peut avoir plusieurs modes (par exemple s'il y a deux modes, on dit que la série est bimodale).

• Une série peut ne pas avoir de mode (toutes les modalités ont le même effectif).

4.1.2 Exemples

Variable quantitative discrète

Une enquête portant sur le nombre d'appareils électroménagers détenus par famille dans un quartier de Dakar a donné les résultats suivants :

Nombre d'appareils électroménagers	Nombre de famille
0	2
1	5
2	10
3	53
4	30
Total	100

La modalité 3 a l'effectif le plus grand donc 3 est le mode de la série.
On peut dire que la plupart des familles disposent de 3 appareils électroménagers.

Variable quantitative continue

Une enquête portant sur la superficie de différents magasins d'une chaîne de produit de beauté a donné les résultats suivants :

Superficie de magasins ($en\, m^2$)	Nombre de magasins
$[20, 30[$	8
$[30, 40[$	7
$[40, 50[$	5
$[50, 60[$	13
$[60, 70[$	7
Total	40

La classe $[50, 60[$ a l'effectif le plus élevé, c'est donc la classe modale. Le mode est égal à 55.
On peut dire que la plupart des magasins ont une superficie comprise entre 50 m^2 et 60 m^2

Variable qualitative

Une étude portant sur la vente journalière d'un marchand de fruits a donné les résultats suivants

Fruits	Poids (en kg)
Pomme	12
Banane	30
Raisin	8
Orange	17
Mangue	30
Total	97

Les modalités Banane et Mangue sont les modes de la série.

4.2 La moyenne

La moyenne n'est définie que pour une variable quantitative.

4.2.1 Moyenne dans le cas d'une variable discrète

Soit X une variable discrète de modalités $x_1, x_2, ..., x_k$ dont les effectifs respectifs sont $n_1, n_2, ..., n_k$. On appelle moyenne de X le réel noté $\overline{x}$ et défini par

$$\overline{x} = \frac{x_1 \times n_1 + x_2 \times n_2 + \cdots + x_k \times n_k}{N},$$

où N est l'effectif total.

Remarque : on a aussi

$$\overline{x} = x_1 \times f_1 + x_2 \times f_2 + \cdots + x_k \times f_k,$$

où f_i est la fréquence de la modalité x_i, $i = 1, 2, ..., k$.

4.2.2 Moyenne dans le cas d'une variable continue

Soit X une variable continue groupée en classes : $[a_0, a_1[, [a_1, a_2[..., [a_{k-1}, a_k[$ d'effectifs respectifs $n_1, n_2, ..., n_k$. On appelle moyenne de X le réel noté $\overline{x}$ et défini par

$$\overline{x} = \frac{x_1 \times c_1 + x_2 \times c_2 + \cdots + x_k \times c_k}{N},$$

où c_i est le centre de la classe $[a_{i-1}, a_i[$, pour $i = 1, 2, ..., k$.

Remarque : on a aussi

$$\overline{x} = c_1 \times f_1 + c_2 \times f_2 + \cdots + c_k \times f_k$$

où f_i est la fréquence de la classe de centre c_i pour $i = 1, 2, ..., k$.

4.2.3 Exemples de calcul de la moyenne

Moyenne d'une série à variable discrète

- Les notes en SVT obtenues par les 50 élèves d'une classe de 3è*me* sont les suivantes :

Note (sur 20)	Effectif
8	15
10	10
12	12
15	8
18	5
Total	50

La moyenne de la classe se calcule ainsi :

$$\overline{x} = \frac{8 \times 15 + 10 \times 10 + 12 \times 12 + 15 \times 8 + 18 \times 5}{50} = 11,48.$$

On peut dire que les notes tournent autour de $11,48$.
• Les notes de mathématiques obtenues par un élève au cours d'un semestre sont les suivantes : 10; 8,5; 13,5; 11
Sa moyenne se calcule ainsi :

$$\overline{x} = \frac{10 + 8{,}5 + 13{,}5 + 11}{4} = 10,75.$$

Moyenne d'une série à variable continue

- Une course de vitesse dans une séance d'EPS a donné les résultats suivants :

Temps (en secondes)	Effectif	Centre de classe
$[10, 15[$	5	$12,5$
$[15, 20[$	10	$17,5$
$[20, 25[$	15	$22,5$
$[25, 30[$	10	$27,5$
Total	40	

Le temps moyen $\overline{x}$ de la course est :

$$\overline{x} = \frac{12,5 \times 5 + 17,5 \times 10 + 22,5 \times 15 + 27,5 \times 10}{40} = 21,25\, secondes.$$

On peut dire qu'un élève met environ $21,25$ secondes pour parcourir $100\, m$.

• Le tableau ci-dessous donne la répartition de 80 moutons d'un abattoir selon leurs poids.

Poids (kg)	Centre de classe	Effectif
$[15, 25[$	20	12
$[25, 35[$	30	48
$[35, 45[$	40	10
$[45, 55[$	50	7
$[55, 65[$	60	2
$[65, 75[$	70	1
Total		80

Le poids moyen est :

$$\overline{x} = \frac{20 \times 12 + 30 \times 48 + 40 \times 10 + 50 \times 7 + 60 \times 2 + 70 \times 1}{80} = 32,75\,kg.$$

Un mouton pèse environ $32,75\,kg$.

4.3 La médiane

4.3.1 Définition

La médiane notée Mé est une valeur qui partage la série (rangée en ordre croissant ou décroissant) en deux groupes de même effectif. Autrement dit c'est un nombre qui laisse autant de valeurs observées à sa gauche qu'à sa droite.

4.3.2 Détermination

Cas d'une variable discrète

Exemples

• Soit la série statistique suivante :

$$4, \quad 6, \quad 9, \quad 4, \quad 7, \quad 8, \quad 10, \quad 9, \quad 8, \quad 6, \quad 5.$$

Pour déterminer la médiane, on range les valeurs de la série dans l'ordre croissant :

$$4, \quad 4, \quad 5, \quad 6, \quad 6, \quad \mathbf{7}, \quad 8, \quad 8, \quad 9, \quad 9, \quad 10.$$

Il y a cinq observations avant 7 et cinq observations après. Donc la médiane est égale à 7 : Mé=7.

• Soit la série statistique suivante :

$$12, \quad 7, \quad 7, \quad 8, \quad 10, \quad 7, \quad 11, \quad 12, \quad 5, \quad 12, \quad 6.$$

Pour déterminer la médiane, on range les valeurs de la série en ordre croissant :

$$5, \quad 6, \quad 7, \quad 7, \quad 8, \quad 10, \quad 11, \quad 12, \quad 12, \quad 12, \quad 13, \quad 14.$$

On remarque qu'aucune valeur ne laisse autant d'observations à gauche que d'observations à droite. Cependant tout nombre appartenant à l'intervalle $]10, 11[$ peut être considéré comme une médiane de cette série.
Par convention on choisira comme médiane

$$\text{Mé} = \frac{10+11}{2} = 10,5.$$

Dans le cas général, on détermine la médiane selon les règles suivantes :
• Si l'effectif total N dans la série ordonnée est impair, alors la médiane correspond à l'observation de rang $\left(\frac{N+1}{2}\right)$.
• Si N est pair, alors il y a deux conventions :
Première convention
La médiane correspond à la demi-somme des observations de rang $\left(\frac{N}{2}\right)$ et $\left(\frac{N}{2}+1\right)$.
Deuxième convention
La médiane correspond à l'observation de rang $\left(\frac{N}{2}\right)$.
Remarques
• Si l'effectif total est impair, la médiane est une valeur observée. Si l'effectif total est pair, la médiane n'est pas toujours une valeur observée.
• Si l'effectif total est grand, pour déterminer la médiane, il est plus aisé d'utiliser le tableau des effectifs cumulés ou des fréquences cumulées.
• En pratique la première convention est la plus fréquemment utilisée. Cependant les valeurs de la médiane trouvées pour les deux conventions sont peu différentes.

Application
• Le tableau suivant indique le nombre d'heures supplémentaires effectuées par les travailleurs d'une entreprise au cours d'une semaine donnée.

Heures supplémentaires effectuées (h)	Nombre de travailleurs
1	3
2	5
3	7
4	2
5	10
6	8
Total	35

Pour déterminer la médiane, on complète le tableau avec les effectifs cumulés croissants comme suit :

Heures supplémentaires effectuées (h)	Effectifs	ECC
1	3	3
2	5	8
3	7	15
4	2	17
5	8	27
6	10	35
Total	35	

L'effectif total étant impair ($N = 35$), la médiane correspond à l'observation de rang $\dfrac{35+1}{2} = 18$. Ainsi Mé=5.

• On considère la série statistique donnant les masses en grammes des oeufs de poule d'un élevage. On donne le tableau des effectifs suivants :

Poids	50	51	52	53	54	55	56	57	58	59	60
Effectifs	2	3	7	14	24	15	17	6	3	2	7
ECC	2	5	12	26	50	65	82	88	91	93	100

L'effectif total ($N = 100$) est pair.
Selon la première convention, la médiane correspond à la demi-somme des observations de rang 50 et 51. Ainsi Mé=$\dfrac{54+55}{2} = 54,5$.
Selon la deuxième convention, la médiane correspond à l'observation de rang 50, c'est-à-dire Mé=54.

Cas d'une variable continue

Pour déterminer la médiane on peut utiliser la méthode graphique ou le calcul.

Détermination graphique

On trace le polygone des effectifs cumulés croissants ou celui des effectifs cumulés décroissants. La médiane est l'abscisse du point d'un des polygones d'ordonnée $\dfrac{N}{2}$.

Remarques

• Si on travaille avec les fréquences, la médiane correspond à l'abscisse du point d'ordonnée $0,5$.

• La médiane est l'abscisse du point d'intersection du polygone des ECC et celui des ECD.

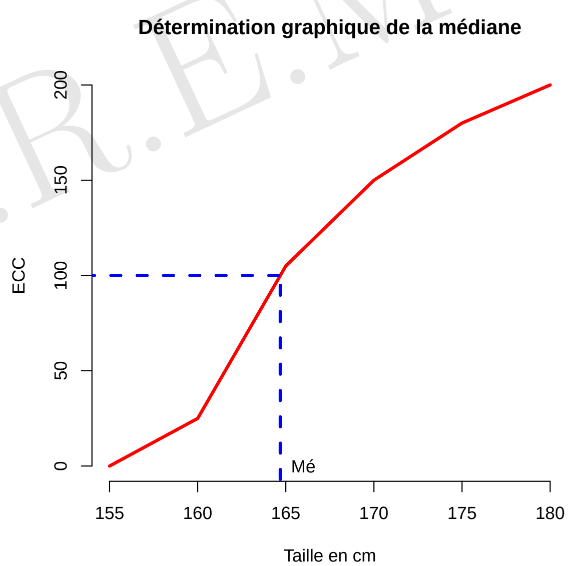
<u>Exemple</u>

On a mesuré la taille des candidats d'un jury de baccalauréat. Les résultats obtenus sont donnés dans le tableau ci-dessous :

Taille (cm)	Nombre de candidats
$[155, 160[$	25
$[160, 165[$	80
$[165, 170[$	45
$[170, 175[$	30
$[175, 180[$	20
Total	200

Déterminons la taille médiane des candidats. Pour cela, on calcule les effectifs cumulés croissants. Ce qui conduit au tableau suivant :

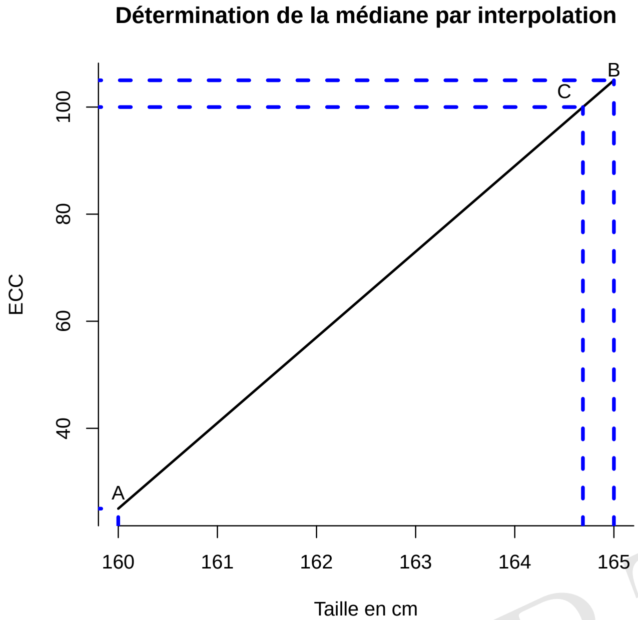
Taille (cm)	Effectifs	ECC
$[155, 160[$	25	25
$[160, 165[$	80	105
$[165, 170[$	45	150
$[170, 175[$	30	180
$[175, 180[$	20	200
Total	200	$\times$



Calcul de la médiane d'une série continue par interpolation affine

En utilisant la colonne des ECC : la classe médiane est la première dont l'ECC est supérieur ou égal à $\frac{N}{2}$.

Dans notre exemple, la classe médiane est $[160, 165[$.



Considérons les points $A(160, 25)$, $B(165, 105)$ et $C(\text{Mé}, 100)$. Puisqu'ils sont alignés, on peut déterminer le coefficient directeur de la droite qui les contient en utilisant deux de ces points (A et B ou A et C ou B et C). Puisque le coefficient directeur est unique, on peut alors établir la formule suivante :

$$\frac{y_B - y_A}{x_B - x_A} = \frac{y_C - y_A}{x_C - x_A},$$

appelée formule d'interpolation affine.
En remplaçant les coordonnées de A, B et C par leurs valeurs dans notre exemple, on obtient

$$\frac{105 - 25}{165 - 160} = \frac{100 - 25}{\text{Mé} - 160},$$

c'est-à-dire

$$\frac{80}{5} = \frac{75}{\text{Mé} - 160}.$$

Ce qui donne

$$\frac{\text{Mé} - 160}{75} = \frac{5}{80},$$

soit

$$\text{Mé} - 160 = \frac{5}{80} \times 75.$$

On trouve $\text{Mé} = 164,6875$.

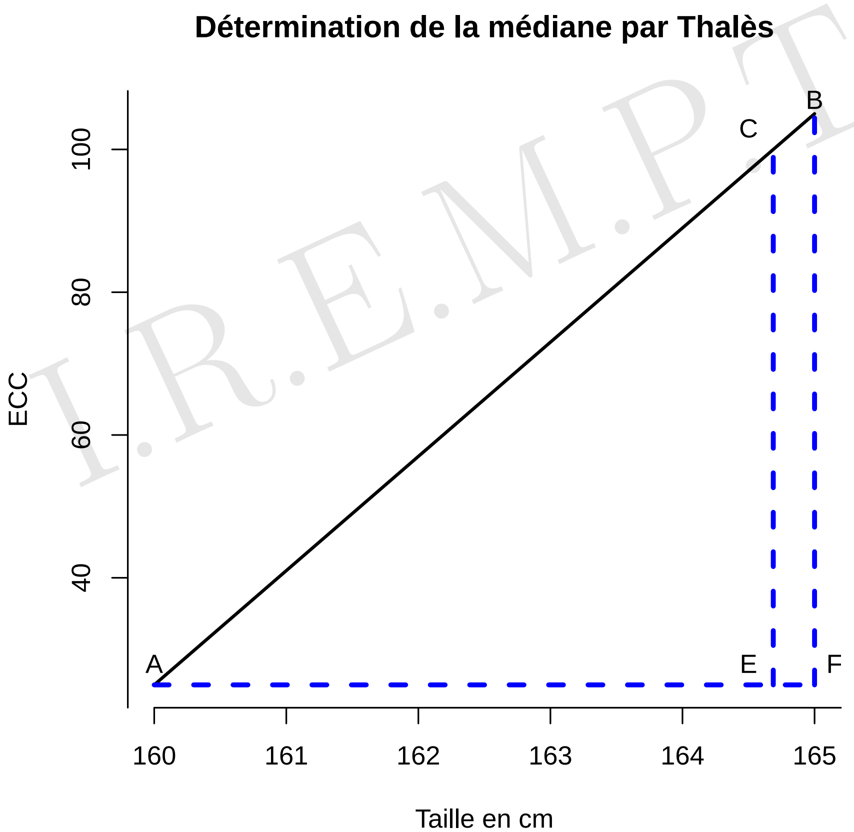
<u>Généralisation</u>
Soient $[a;b[$ la classe médiane, n_1 son ECC, n_2 celui de la classe qui la précède, et N l'effectif total de la population.
La médiane Mé s'obtient par :

$$\frac{\text{Mé} - a}{\frac{N}{2} - n_2} = \frac{b-a}{n_1 - n_2}.$$

<u>Remarque</u>
On peut déterminer de même la médiane en utilisant les fréquences.

Calcul de la médiane d'une série continue avec le théorème de Thalès

Aux points $A(160, 25), B(165, 105)$ et $C(\text{Mé}, 100)$ définis précédemment associons les points $E(\text{Mé}, 25)$ et $F(165, 25)$. Les droites (CE) et (BF) sont parallèles, donc les triangles ACE et ABF sont "en configuration de Thalès" (voir figure ci-dessous).



En utilisant le théorème de Thalès, on obtient :

$$\frac{AE}{AF} = \frac{EC}{BF}.$$

Or, $AE = \text{Mé} - 160$; $AF = 165 - 160$; $EC = 100 - 25$ et $BF = 105 - 25$.
Donc
$$\frac{\text{Mé} - 160}{165 - 160} = \frac{100 - 25}{105 - 25}.$$
D'après ce qui précède : $\text{Mé} = 164,6875$

4.4 Les quartiles d'une série statistique

Les quartiles sont les trois valeurs qui partagent la série ordonnée en quatre séries de même effectif.
Le premier quartile
C'est la valeur qui laisse 25% des observations à sa gauche et 75% des observations à sa droite. On le note q_1.
Le deuxième quartile
Il correspond à la médiane Mé.
Le troisième quartile
C'est la valeur qui laisse 75% des observations à sa gauche et 25% des observations à sa droite. On le note q_3.

4.4.1 Détermination arithmétique des quartiles d'une série discrète

Considérons une série d'effectif total N dont les observations sont rangées dans l'ordre croissant.
Détermination du premier quartile
• Si $\frac{N}{4}$ n'est pas un nombre entier, alors q_1 est l'observation de rang égal au plus petit entier supérieur à $\frac{N}{4}$.
• Si $\frac{N}{4}$ est un nombre entier, alors il y a deux conventions :

Première convention :
q_1 est l'observation de rang $\frac{N}{4}$.

Deuxième convention :
q_1 est la demi somme des observations de rang $\frac{N}{4}$ et $\frac{N}{4} + 1$.

Détermination du troisième quartile

• Si $\frac{3N}{4}$ n'est pas un nombre entier, alors q_3 est l'observation de rang égal au plus petit entier supérieur à $\frac{3N}{4}$.

• Si $\frac{3N}{4}$ est un nombre entier, alors il y a deux conventions :

Première convention : q_3 est l'observation de rang $\frac{3N}{4}$.

Deuxième convention : q_3 est la demi somme des observations de rang $\frac{3N}{4}$ et $\frac{3N}{4}+1$.

Exemples

• Le tableau suivant indique le nombre d'heures supplémentaires effectuées par les travailleurs d'une entreprise au cours d'une semaine donnée.

Heures supplémentaires effectuées (h)	Effectifs	ECC
1	3	3
2	5	8
3	7	15
4	2	17
5	8	27
6	10	35
Total	35	

Ici l'effectif total $N = 35$.

$\frac{35}{4} = 8,75$ n'est pas entier. Par conséquent q_1 est l'observation de rang 9, soit $q_1 = 3$.

$3 \times \frac{35}{4} = 26,25$ n'est pas entier. Par conséquent q_3 est l'observation de rang 27, soit $q_3 = 5$.

• On considère la série statistique donnant les masses en grammes des oeufs de poule d'un élevage. On donne le tableau des effectifs suivants :

Poids	50	51	52	53	54	55	56	57	58	59	60
Effectifs	2	3	7	14	24	15	17	6	3	2	7
ECC	2	5	12	26	50	65	82	88	91	93	100

L'effectif total $N = 100$.

$\frac{N}{4} = 25$ et $3 \times \frac{N}{4} = 75$ sont des entiers.

Première convention : q_1 est l'observation de rang 25 et q_3 est l'observation de rang 75, soient $q_1 = 53$ et $q_3 = 56$.

Deuxième convention : q_1 est la demi somme des observations de rang 25 et 26, tandis que q_3 est la demi somme des observations de rang 75 et 76, soient $q_1 = 53$ etb $q_3 = 56$.

Ici, on remarque que les valeurs trouvées sont identiques quelle que soit la convention utilisée.

4.4.2 Détermination arithmétique des quartiles d'une série continue

Exemple

La répartition de 100 personnes selon leur temps de sieste (en min) a donné le tableau suivant :

Temps (en min)	Effectif
$[10, 15[$	10
$[15, 20[$	35
$[20, 25[$	30
$[25, 30[$	15
$[30, 35[$	10
Total	100

Pour déterminer les quartiles, on complète le tableau avec les effectifs cumulés et les fréquences cumulées croissants.

Temps (en min)	Effectif	ECC
$[10; 15[$	10	10
$[15; 20[$	20	30
$[20; 25[$	25	55
$[25; 30[$	40	95
$[30; 35[$	5	100

On procède de la même manière que pour le calcul de la médiane par interpolation linéaire.

• q_1 appartient à la première classe dont l'ECC est supérieur ou égal à $\frac{N}{4}$.

$N = 100$ et $\dfrac{N}{4} = 25$. Donc $q_1 \in [15, 20[$. Ainsi,

$$\frac{q_1 - 15}{20 - 15} = \frac{25 - 10}{30 - 10}.$$

Donc $q_1 = 18,75$.

• q_3 appartient à la première classe dont l'ECC est supérieur ou égal à $\dfrac{3N}{4}$.

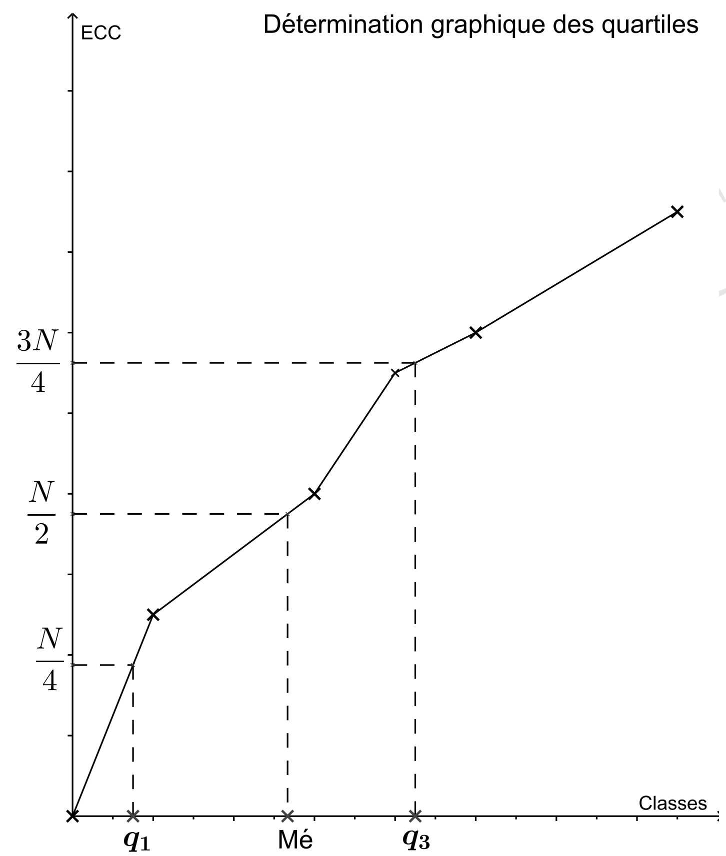
$N = 100$ et $\frac{3N}{4} = 75$. Donc $q_3 \in [25, 30[$. Ainsi,

$$\frac{q_3 - 25}{30 - 25} = \frac{75 - 55}{95 - 55}.$$

Donc $q_3 = 27,50$.

Détermination graphique des quartiles d'une série continue

On construit le polygone des ECC :

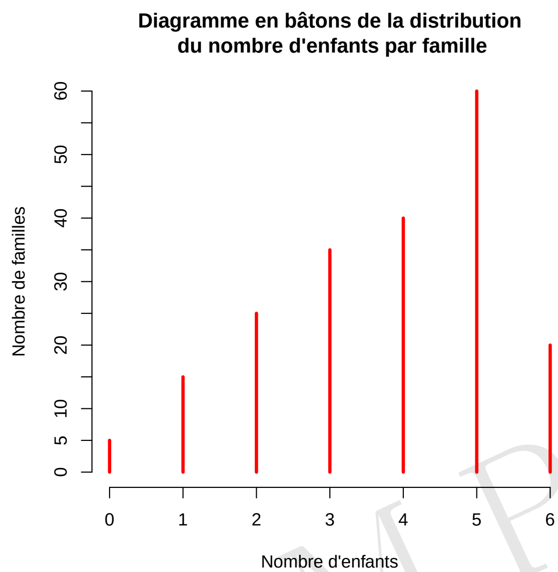


Les antécédents de $\frac{N}{4}, \frac{N}{2}$ et $\frac{3N}{4}$ sont respectivement q_1, Mé et q_3.

4.5 Exercices sur le chapitre 4

Exercice 1

Le relevé du nombre d'enfants dans chacune des familles d'un quartier a donné le diagramme en bâtons suivant :



1) Quel est le mode de la série ?
2) *a*) Dans combien de familles le relevé a-t-il été fait ?
   b) Calcule la moyenne de la série.
3) Détermine la médiane de la série et les autres quartiles.

Exercice 2

Au cours d'une semaine d'un mois de juin dans la région de Dakar, les températures (en degrés) observées jour après jour en un moment donné sont respectivement :

$$26; \quad 32; \quad 24; \quad 29; \quad x; \quad 31; \quad 24.$$

1) Calcule la température observée au cinquième jour pour que la température moyenne de la semaine soit égale à 28°.
2) Quelle est la température modale ?
3) Quelle est la température médiane ?

Exercice 3

A la question :
"Combien de messages téléphoniques avez-vous reçus au cours de la journée ?"

30 personnes ont donné les réponses suivantes :

5; 2; 0; 4; 6; 7; 5; 8; 6; 10; 0; 12; 4; 15; 2; 3;

0; 15; 14; 16; 9; 8; 6; 17; 18; 10; 7; 8; 11; 12.

1) Range ces nombres dans l'ordre croissant.
2) Les données ci-dessus sont regroupées en classes d'amplitude 4.
Complète alors le tableau suivant :

Classes	Effectifs	Centres de classe	ECC
$[0;4[$			
$[4;8[$			
$[8;12[$			
$[12;16[$			
$[16;20[$			

3) Quelle est la classe modale de cette série ?
4) Calcule le nombre moyen de messages reçus.
5) Détermine le nombre médian de messages reçus.

Exercice 4

Dans le but d'équiper le laboratoire de son établissement, un principal achète des ordinateurs.
1) Complète le tableau suivant :

Prix (en FCFA)	Nombre d'ordinateurs	ECC	ECD
80.000		1	
90.000		6	
100.000		23	
110.000		38	
120.000		49	
130.000		57	

2) Détermine le prix modal et le prix moyen des ordinateurs.
3) Détermine le prix médian des ordinateurs ainsi que les autres quartiles.

Exercice 5

Le tableau ci-dessous donne la répartition des salaires des travailleurs d'une fabrique de chaussures.

Salaires en milliers de francs	Nombre de travailleurs
$[40; 55[$	12
$[55; 70[$	15
$[70; 85[$	8
$[85; 100[$	10
$[100; 115[$	5

1) Recopie puis complète le tableau avec les fréquences cumulées (en pourcentages).
2) Déduis-en le pourcentage des travailleurs qui gagnent au moins $85000\,F$.
3) Interprète la fréquence cumulée croissante de la classe $[55; 70[$.
4) Calcule la moyenne des salaires dans cette fabrique.
5) Construis l'histogramme des fréquences cumulées décroissantes et son polygone puis calcule le salaire médian de cette série en appliquant le théorème de Thalès.

Exercice 6

Une enquête portant sur la distance "maison-école" menée au sein d'un groupe d'élèves a donné la répartition ci-dessous :

Distance (en km)	Nombre d'élèves
$[3; 4,8[$	20
$[4,8; 6,6[$	16
$[6,6; 8,4[$	14
$[8,4; 10,2[$	12
$[10,2; 12[$	18

1) Recopie puis complète le tableau avec les effectifs cumulés et les fréquences.
2) Donne la classe médiane.Justifie.
3) Quelle est la distance moyenne que parcourent ces élèves pour aller à l'école ?
4) Détermine le nombre d'élèves qui parcourent au moins $6,6$ km et celui des élèves qui parcourent moins de $8,4$ km.
5) Construis l'histogramme des effectifs cumulés décroissants puis calcule la distance médiane en appliquant le théorème de Thalès.

Exercice 7

Le tableau ci-dessous donne la répartition des absences relevées au sein d'une classe.

Nombre d'absences	ECC
0	3
1	10
2	15
3	27
4	42
5	46
6	49
7	50

Détermine :
1) le mode,
2) la moyenne,
3) les quartiles de la série.

Propositions de solutions

Exercice 1
1) 5 est le mode de la série.
2) a) Le nombre N de familles dans lesquelles le relevé a été fait correspond à l'effectif total. Il s'obtient en faisant la somme des effectifs des différentes modalités.

$$N = 5 + 15 + 25 + 35 + 40 + 60 + 20 = 200.$$

b) La moyenne $\overline{x}$ de la série est donnée par

$$\overline{x} = \frac{0 \times 1 + 1 \times 15 + 2 \times 25 + 3 \times 35 + 4 \times 40 + 5 \times 60 + 6 \times 20}{200} = \frac{750}{200} = 3,75.$$

NB : Puisque la variable représente le nombre d'enfants, on peut arrondir la moyenne à 4.
3) N est pair.
Selon la première convention, la médiane est la demi-somme des observations de rang 100 et de rang 101, soit $Mé = \frac{4+4}{2} = 4$.
Selon la deuxième convention, la médiane est l'observation de rang $\frac{N}{2} = 100$, soit $Mé = 4$.
$\frac{N}{4} = 50$.
Selon la première convention, le premier quartile est la demi-somme des observations de rang 50 et de rang 51, soit $q_1 = \frac{3+3}{2} = 3$.
Selon la deuxième convention, le troisième quartile est l'observation de rang $\frac{N}{4} = 50$, soit $q_1 = 3$.
$\frac{3N}{4} = 150$.
Selon la première convention, le troisième quartile est la demi-somme des observations de rang 150 et de rang 151, soit $q_3 = \frac{5+5}{2} = 5$.
Selon la deuxième convention, le troisième quartile est l'observation de rang $\frac{3N}{4} = 150$, soit $q_3 = 5$.
Exercice 2
La température moyenne est égale à

$$\frac{26 + 32 + 24 + 29 + x + 31 + 24}{7}.$$

Pour déterminer x, il suffit de résoudre l'équation

$$19 = \frac{26 + 32 + 24 + 29 + x + 31 + 24}{7},$$

on trouve $x = 30°$.
2) La température modale est 24° car il a l'effectif le plus grand.

3) Pour déterminer la médiane, ordonnons la série comme suit :

$$24; \quad 24; \quad 26; \quad 29; \quad 30; \quad 31; \quad 32.$$

Comme $N = 7$ est impair, la médiane est l'observation de rang $\dfrac{7+1}{2} = 4$, soit Mé $= 29°$.

Exercice 3

1) Rangeons les valeurs dans l'ordre croissant.

$$0; \quad 0; \quad 0; \quad 2; \quad 2; \quad 3; \quad 4; \quad 4; \quad 5; \quad 5; \quad 6; \quad 6; \quad 6; \quad 7; \quad 7; \quad 8$$

$$8; \quad 8; \quad 9; \quad 10; \quad 10; \quad 11; \quad 12; \quad 12; \quad 14; \quad 15; \quad 15; \quad 16; \quad 17; \quad 18.$$

2) Complétons le tableau

Classes	Effectifs	Centres de classe	ECC
$[0, 4[$	6	2	6
$[4, 8[$	9	6	15
$[8, 12[$	7	10	22
$[12, 16[$	5	14	27
$[16, 20[$	3	18	30

3) $[4, 8[$ est la classe qui a le plus grand effectif , donc c'est la classe modale.
4) Soit m le nombre moyen de messages reçus. On a alors

$$m = \frac{2 \times 6 + 6 \times 9 + 10 \times 7 + 14 \times 5 + 18 \times 3}{30} = 8,66.$$

On trouve $m = 8,66$.
5) $\dfrac{N}{2} = 15$ et l'abscisse du point d'ordonnée 15 sur le polygone cumulatif est 8. Donc Mé $= 8$.

Exercice 4

1)

Prix (en FCFA)	Nombre d'ordinateurs	ECC	ECD
80.000	1	1	57
90.000	5	6	56
100.000	17	23	51
110.000	15	38	34
120.000	11	49	19
130.000	8	57	8

2) La modalité qui a l'effectif le plus grand est $100.000\,F$, donc le mode est égal à $100.000\,F$.

$$m = \frac{80.000 \times 1 + 90.000 \times 5 + 100.000 \times 17 + 110.000 \times 15 + 120.000 \times 11 + 130.000 \times 8}{57} = 109.474$$

Le prix moyen des ordinateurs est de $109.474\ FCFA$.

3) L'effectif total N étant impair, la médiane est l'observation de rang $\dfrac{N+1}{2} = 29$. on trouve Mé $= 110.000$. Le prix médian est donc de $110.000\,F$.

Les autres quartiles :

$\dfrac{N}{4} = \dfrac{57}{4} = 14,25$ n'est pas un nombre entier. Donc q_1 est l'observation de rang 15. Donc $q_1 = 100.000\,F$.

$\dfrac{3N}{4} = \dfrac{3 \times 57}{4} = 42,75$ n'est pas un nombre entier. Donc q_3 est l'observation de rang 43. Donc $q_3 = 180.000\,F$.

4.6 Exercices de Synthèse

Exercice 1 (B.F.E.M 2019)

1. Dis comment obtenir la valeur de la médiane d'une série statistique ordonnée à caractère quantitatif discret et d'effectif total N.
2. Le tableau ci-dessous donne la répartition des salaires mensuels et leurs proportions dans le personnel d'une entreprise.

fonctions	Fréquences en pourcentages	salaires
Cadres supérieurs	5	450 000
Agents de production	45	350 000
Personnels administratifs	15	200 000
Chauffeurs	5	150 000
Agents de sécurité	10	100 000
Agents commerciaux	20	175 000

   a) Indique le caractère étudié et sa nature.

   b) Calcule le salaire moyen mensuel dans cette entreprise

3. Calcule le salaire médian de cette entreprise sachant qu'il y a exactement 2 cadres qui y travaillent.
4. Construis le diagramme des fréquences cumulées croissantes de cette série.

Exercice 2 (B.F.E.M 2018)

On considère la liste des notes obtenues par les élèves d'une classe de troisième, lors d'un devoir de mathématiques.

5; 8; 7; 8; 9; 6; 10; 11; 15; 13; 10; 18; 16; 15; 12; 9; 14;

16; 17; 15; 10; 16; 17; 8; 9; 10; 16; 9; 10; 7; 10; 6; 12; 13;

11; 13; 18; 10; 11; 6; 10; 13; 17; 12; 11; 12; 9; 16; 17; 14.

1) Regroupe ces notes en classes d'amplitude 3, en commençant par la classe $[5;8[$[1].
2) Calcule l'effectif cumulé croissant de chaque classe.
3) Calcule la note moyenne.
4) Trace le diagramme des effectifs cumulés croissants.
5) Détermine graphiquement la médiane de cette série.

1. Nous avons jugé nécessaire d'apporter cette précision

Exercice 3 (B.F.E.M 2017)

Les notes des 160 candidats à un concours sont consignées dans le tableau suivant :

Notes	[10 ;12[	[12 ;14[	[14 ;16[	[16 ;18[	[18 ;20[
Fréquences	0,3	x	0,2	0,15	y

1) Donne une interprétation de la valeur 0,3 ; fréquence de la classe [10 ; 12[.
2) Calcule x et y sachant que 25% des élèves ont une note supérieure ou égale à 16.
3) On donne $x = 0,25$ et $y = 0,1$.
   a) Calcule la moyenne des notes.
   b) Construis le diagramme des fréquences cumulées décroissantes.

Exercice 4 (B.F.E.M 2016)

Une série statistique à caractère quantitatif continu, groupée en classes d'amplitude 10 compte 5 classes de centres respectifs : C_1, C_2, C_3, C_4 et C_5 d'effectifs respectifs : n_1, n_2, n_3, n_4 et n_5.

1. Donne l'expression de sa moyenne .
2. Lors d'un recrutement au service militaire, les tailles de 100 candidats ont été répertoriées dans le tableau ci-dessous.

Tailles en cm	[135,145[	[145,155[	[155,165[	[165,175[	[175,185[
Fréquences	0,12	a	0,28	0,32	b
E.C.C					

   Sachant que la moyenne de cette série est de 161 cm , calcule a et b.
3. Pour la suite, tu prendras $a = 0,18$ et $b = 0,10$.
   a) Recopie et complète le tableau.
   b) Combien de candidats ont une taille au moins égale à 165 cm ?
   c) Détermine graphiquement la classe médiane de la série.

Exercice 6 (B.F.E.M 2015)

Le tableau statistique ci-dessous donne la répartition des usagers transportés en une journée par une entreprise de transport selon le prix du ticket de section acheté.

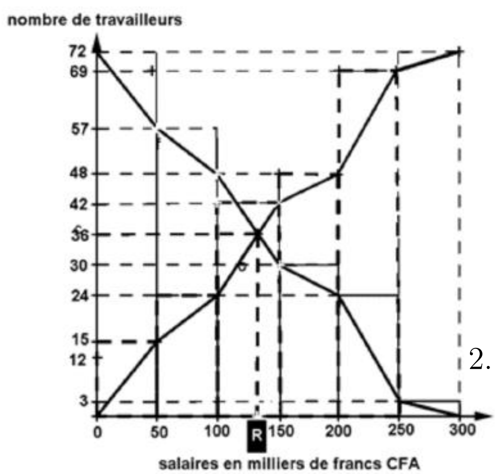
Prix du ticket de section en FCFA	100	150	200	250	300	350
Nombre de tickets	2480	1060	820	960	780	1100
Effectifs cumulés croissants	2480	3540	4360	5320	6100	7200
Effectifs cumulés décroissants	7200	4720	3660	2840	1880	1100

1. Quel est le caractère statistique étudié ?

2. Combien cette entreprise a-t-elle transporté d'usagers ce jour ?
3. Donne les modalités du caractère étudié.
4. Quel est le nombre d'usagers ayant acheté un ticket valant moins de $250F$?
5. Quel est le nombre d'usagers ayant acheté un ticket valant au moins $250F$?
6. Quel est le prix médian du ticket de section de ce jour (médiane de cette série) ?
7. Calcule le prix moyen du ticket de section de ce jour (la moyenne de cette série).
8. Construis le diagramme circulaire de la série.

Exercice 7 (B.F.E.M 2014)

Dans une petite et moyenne entreprise ou PME on étudie la répartition des salaires des travailleurs.
Le schéma ci-dessous en représente l'histogramme des ECC et celui des ECD tracés dans un même repère.



1. Déduis du schéma :
   a le caractère étudié puis précise sa nature,
   b) le nombre de travailleurs, dans cette PME,
   c) le nombre de travailleurs qui gagnent au moins $100000F$
   d) le nombre de travailleurs qui gagnent moins de $150000FCFA$,
   e) le nombre de travailleurs qui gagnent entre $150\,000F$ et $200\,000F$
2. Reproduis avec soin l'histogramme des effectifs cumulés croissants en prenant en abscisses$1cm$ pour $50\,000F$ et en ordonnées $0,5cm$ pour 3 travailleurs.
3. Donne la signification statistique du salaire R sur le schéma

Exercice 8 (B.F.E.M 2013)

Une enquête portant sur le nombre de filles fréquentant une classe de terminales scientifiques, menée dans les 50 établissements scolaires d'une localité, a donné le relevé ci-dessous :

12; 11; 10; 14; 8; 0; 5; 10; 7; 10; 14; 10; 13; 14; 4; 18; 10;

10; 7; 10; 4; 10; 13; 11; 13; 18; 4; 13; 12; 18; 17; 0; 6; 5;

5; 6; 10; 10; 9; 7; 11; 4; 15; 17; 16; 16; 15; 10; 0; 3.

1) a) Quelle est la population étudiée ?
   b) Quel est le caractère étudié ? Quelle est sa nature ?
   c) Donne une modalité de ce caractère et son effectif partiel.
2) a) Calcule l'effectif moyen de filles en terminales scientifiques dans ces établissements.

b) Détermine la médiane de cette série statistique.

c) Dans combien d'établissements scolaires a-t-on au moins 10 filles en classes de terminales scientifiques ?

3) *a*) Regroupe les données recueillies en classes d'amplitude 5.

b) Dresse le tableau statistique de la série comprenant l'effectif et l'effectif cumulé décroissant de chacune des classes.

4) Construis l'histogramme des effectifs cumulés décroissants et le diagramme des effectifs cumulés décroissants de cette série.

Exercice 9 (B.F.E.M 2011)

Les lutteurs d'une écurie sont répartis en cinq classes de poids d'amplitude 15 kg.
On a les classes suivantes : $[80; 95[$, $[95; 110[$, $[110; 125[$, $[125; 140[$ et $[140; 155[$ qui seront appelées dans la suite catégorie 1, 2, 3, 4 et 5.

1) Les lutteurs de la catégorie 2 sont au nombre de 6 et représentent 12% de l'effectif de l'écurie.
Montre qu'il y a 50 lutteurs dans cette écurie.

2) La mesure de l'angle correspondant à la catégorie 3 dans le diagramme circulaire de la série est de 36°.
Montre que le nombre de lutteurs de cette catégorie est égal à 5.

3) La fréquence de la catégorie 4 est 0,3. Vérifie que cette catégorie compte 15 lutteurs.

4) L'effectif de la catégorie 5 est le tiers de l'effectif de la catégorie 1.
Montre qu'il y a 6 lutteurs dans la catégorie 5.

5) Etablis le tableau des effectifs cumulés croissants des classes de poids, puis déduis-en la classe médiane.

4.7 Exercices complémentaires

Exercice 1

Le département Marketing de la société Rola-Cola souhaite étudier le comportement des consommateurs vis-à-vis de Rola-Cola et de la marque concurrente Cora-Cola. Une enquête a été réalisée auprès des clients choisis au hasard dans un centre commercial et à qui on a posé les questions suivantes :

Question 1 : Quelle boisson préférez-vous ?
a) Rola-Cola (RC)
b) Cora-Cola (CC).

Question 2 : Avez-vous déjà acheté Rola-Cola ? Oui (O) ou Non (N).

Question 3 : Quelle est votre réaction à la phrase "J'aime les boissons au Cola sucrées" ?
a)D'accord (DA)
b) Pas sûr (PS)
c) Pas d'accord (PA).

Question 4 : Combien de litres de boisson au Cola votre famille a-t-elle bu le mois dernier ?
Les réponses des clients ont été les suivantes :

$n°$ client	Q_1	Q_2	Q_3	Q_4	$n°$ client	Q1	Q2	Q3	Q4	$n°$ client	Q_1	Q_2	Q_3	Q_4
1	RC	O	DA	1	11	RC	O	DA	6	21	RC	O	PS	6
2	CC	N	PA	6	12	CC	O	PA	2	22	RC	O	DA	8
3	RC	O	DA	4	13	RC	O	PS	4	23	CC	N	PA	2
4	CC	O	PS	7	14	RC	O	DA	7	24	CC	N	PS	17
5	RC	O	PS	2	15	RC	O	PS	11	25	RC	O	DA	5
6	CC	O	DA	12	16	RC	O	PA	4	26	CC	N	PS	14
7	CC	N	PA	6	17	CC	O	DA	3	27	RC	O	PA	15
8	RC	O	PA	2	18	CC	N	PS	10	28	RC	O	PA	5
9	RC	O	PS	9	19	RC	O	DA	5	29	CC	O	PS	9
10	RC	O	PS	5	20	CC	O	PA	7	30	RC	O	PS	4

1) Précise la population et les caractères étudiés. Indique la nature de chaque caractère. Quel est l'effectif total de la population ?

2) Dresse le tableau statistique du caractère représentant la réponse à la question Q3. Représente graphiquement ce caractère de deux façons différentes.

3) Pour la présentation des différents calculs effectués dans les questions suivantes, on construira dès le début un unique tableau présentant l'ensemble des résultats demandés ou utiles.

a) Présente les données obtenues avec la question Q_4 dans un tableau classes/effectifs, en prenant pour classes $[0,5;3,5[$, $[3,5;6,5[$, $[6,5;9,5[$, $[9,5;12,5[$, $[12,5;15,5[$ et $[15,5;18,5[$. Ce type de présentation des données correspond-il à la nature de la variable étudiée ?

b) Représente graphiquement les résultats présentés dans le tableau précédent.

c) Calcule les fréquences cumulées (croissantes) de la distribution, trace le polygone des fréquences cumulées. Détermine graphiquement la médiane puis retrouve ce résultat par le calcul.

d) Calcule la moyenne et le mode de la distribution.

Exercice 2

On considère la série statistique suivante : quantitée (en grammes) de lait entier en poudre absorbés par des bébés de 2 mois en une journée (résultats d'observations portant sur la collectivité de 315 bébés de la pouponnière X...)

Consommation	[40,45[	[45,50[	[50,55[	[55,60[	[60,65[	[65,70[	[70,75[	[75,80[	[80,85[	[85,90[
Nombre de bébés	8	11	31	61	54	58	43	25	17	7

1) Représente l'histogramme et le polygone des effectifs ainsi que le polygone cumulatif de cette série.
2) Détermine le mode de cette série. Interpréte le résultat.
3) Calcule la valeur de la médiane de cette série. Contrôle le résultat au moyen du graphique. Interpréte le résultat.
4) Calcule, à partir de cette série, la quantité moyenne de lait entier en poudre absorbée par un bébé de 2 mois. Interpréte le résultat.
5) Détermine les quartiles de cette série

CHAPITRE 5

R UN LOGICIEL DE TRAITEMENT DE DONNÉES STATISTIQUES

5.1 Introduction

L'évolution de la techcnologie engendre une évolution des programmes. Pour lutter contre la fracture numérique, le Sénégal intègre les TICE dans l'enseignement des mathématiques. C'est pourquoi, nous présentons ci-après le logiciel R permettant à l'élève Sénégalais d'utiliser cet outil pour traiter des données.

5.2 Présentation de R

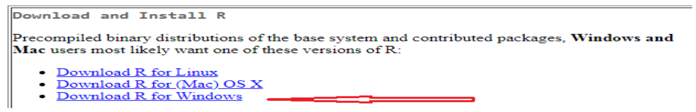
R est un langage de programmation complet, simple et évolutif. C'est un logiciel libre permettant d'effectuer des calculs statistiques et des graphiques. R possède un site web officiel "http ://www.r-project.org/". Il existe des versions Windows et Linux de R, qu'on peut télécharger sur le site web. La documentation complète est également disponible sur ce site. Le logiciel a été nitié en 1990 par Robert Gentleman et Ross Ihaka, chercheurs au département de Statistique de l'Université d'Auckland en Nouvelle-Zélande, d'où le nom R .

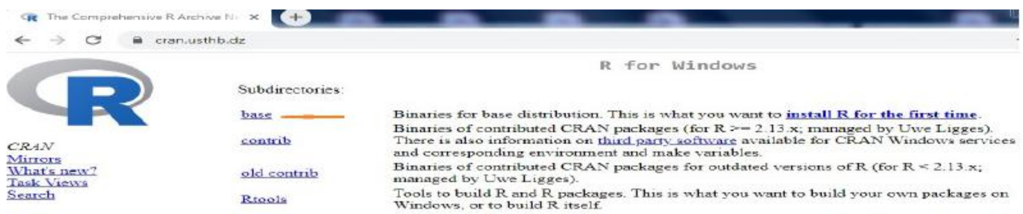
5.3 Procédure de téléchargement et installation de R

• Aller sur la page web http ://www.r-project.org .
• Dans la colonne de gauche, à l'onglet Download, cliquer sur le lien CRAN.
• Dans la liste qui apparaît, choisir le lien de Algeria, en cliquant sur http ://cran.usthb.dz/

• Dans la fenêtre Download and Install R, choisir le lien qui correspond à votre système d'exploitation. Nous choisirons par exemple Download R for Windows.

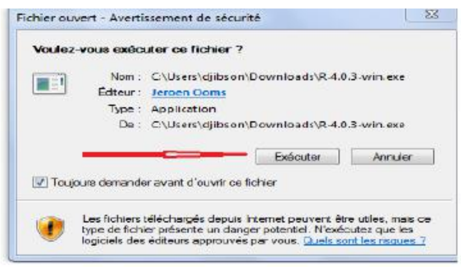
- Cliquer sur base, pour Binaries for base distributions.
- Dans la liste qui apparaît, cliquer sur le fichier R-4.0.3 for Windows (32/64 bit)
- Exécuter ce fichier.
- Suivre les instructions jusqu'à l'installation complète du logiciel.





5.4 Premiers contacts avec R

• Vous lancez le logiciel R en cliquant sur l'icône R. Une seule fenêtre s'ouvre. Elle se nomme R console. Il s'agit d'une fenêtre de commandes dans laquelle on soumet des commandes et des résultats sont affichés.
• Une fois le logiciel lancé, le symbole > apparaît automatiquement en début de chaque ligne de commandes. Ce symbole > signifie que R est prêt à travailler. Il ne faut pas taper ce symbole au clavier car il est déjà présent en début de ligne sur la R Console. C'est à la suite de ce symbole > que vous pourrez taper les commandes R.
• Une commande doit se terminer par un retour de chariot pour afficher le résultat.
Exemple
$> 2/5+3$
[1] 3.4
Une commande peut s'étaler sur plusieurs lignes. Il est possible de mettre plusieurs commandes sur une même ligne, à condition de les séparer par des points-virgules. Autrement dit, le symbole ; permet d'enchaîner plusieurs commandes dont les résultats seront donnés les unes sous les autres.
Exemple
$> 2+5; 5^2; 3/10$



[1] 7
[1] 25
[1] 0.3
Les différents éléments d'une commande peuvent ou non être séparés par un ou plusieurs espaces. Cela n'a aucune importance.
<u>Exemple</u>
$> 2+5\,;\ 2\ +\ 5$
[1] 7
[1] 7
• Si on exécute une commande et que le signe + apparaît en début de la ligne suivante, cela signifie que la commande est incomplète. Il faut alors la compléter ou reprendre la saisie si nécessaire.
• Pour les nombres décimaaux R utilise un point.
<u>Exemple</u>
sous R $3,141593$ s'écrit 3.141593.
• Chaque nombre entre crochets indique la rang du premier élément de la ligne.

Exemple
En tapant la commande runif(15) sous R, on obtient :
[1] 0.8667563 0.6079356 0.9320504 0.6609379 0.8389820 0.3689098 0.6056121
[8] 0.8284623 0.4754653 0.6556009 0.6327941 0.9229922 0.2339661 0.8749937
[15] 0.9862678.
La commande runif(15) produit 15 nombres au hasard compris entre 0 et 1 dont 0.8667563 est le premier, 0.8284623 est le huitième et 0.9862678 est le quinzième.
• Les lignes de commandes déjà tapées peuvent être rappelées avec les flèches ↑ et ↓ du clavier.
• Il est possible de se déplacer sur une ligne de commandes non encore exécutée avec les flèches → et ← du clavier.
• Le symbole # sert à insérer des commentaires dans un programme R. Autrement dit, sur une ligne, tout ce qui se trouve après un # est ignoré par R lors de l'exécution d'un programme.
Exemple
$> 2+5*7; 2+5\# *7$
[1] 37
[1] 7
• Pour nettoyer la fenêtre ≪ R console ≫, vous tapez Ctrl L.
• Pour quitter R, vous utilisez les commandes suivantes : q() oubien quit().
Si vous travaillez avec la version française de R, la question *Sauver une image de la session* ? est alors posée : R propose de sauvegarder le travail effectué. Trois réponses sont possibles : Oui, Non ou Annuler. En tapant Annuler, la procédure de fin de session sous R est annulée. Si vous tapez Oui, cela permet que les commandes tapées pendant la session soient conservées en mémoire et soient donc ≪ rappelables ≫.
Par contre si vous travaillez avec la version anglaise, La question *Save workspace image* ? est posée. Les trois réponses possibles sont alors : y (pour yes), n (pour no) ou c (pour cancel, annuler).
• Si vous quittez R en choisissant la sauvegarde de l'espace de travail, deux fichiers sont créés :
(i) le fichier *.Rdata* qui contient des informations sur les variables utilisées,
(ii) le fichier *.Rhistory* qui contient l'ensemble des commandes utilisées.
Lors de la prochaine session de travail, les objets contenus dans un fichier *.Rdata* positionné dans le répertoire courant seront automatiquement chargés dans l'environnement de travail.

4) R, une machine à calculer

Le logiciel R peut être utilisé comme une simple calculatrice. Pour cela il suffit d'écrire directement l'opération dans la console et d'appuyer sur *Entrée* :
Exemple :
$(2+5)*3+1$
[1] 22
On peut ainsi faire toutes les opérations que l'on souhaite grâce aux opérateurs $+$, $-$, $*$, $/$, $\wedge$ (puissance), etc.

R possède aussi un très grand nombre de fonctions mathématiques prédéfinies dont :

cos()/sin()/tan()	Cosinus/Sinus/Tangente
abs()	Valeur absolue
sqrt()	Racine carrée (*square root*)

5) Les fichiers scripts
Il est plus pratique de saisir les commandes R dans une fenêtre script. Un script est un fichier texte contenant des commandes R que l'on peut exécuter ensemble. Cela permet de travailler plus efficacement lorsque l'on a beaucoup de commandes à saisir sans vouloir une réponse à chaque étape de la saisie.
Pour créer ou ouvrir un nouveau script, il suffit de se placer sur la fenêtre R Console et de cliquer sur Fichier.
Pour exécuter des commandes à partir de la fenêtre de script il suffit de les sélectionner, de faire un clic droit et suivre les instructions.
Pour sauvegarder un script, il suffit, lorsque la fenêtre de script est active, de sélectionner l'entrée "Sauver" du menu "Fichier".

5.5 Taitement de données statistiques avec R

5.5.1 Cas d'une variable qualitative

Reprenons les données de l'exemple 3 du chapitre 2.

Bambara, Bambara, Bambara, Wolof, Sérère, Wolof, Wolof, Wolof, Diola, Diola, Wolof, Diola, Wolof, Poular, Soninké, Diola, Sérère, Wolof, Soninké, Wolof, Sérère, Sérère, Diola, Poular, Balante, Poular, Sérère, Balante, Wolof, Wolof, Diola, Poular, Mancagne, Soninké, Bambara, Wolof, Sérère, Bambara, Diola, Wolof, Bambara, Wolof, Bambara, Sérère, Poular, Bambara, Soninké.

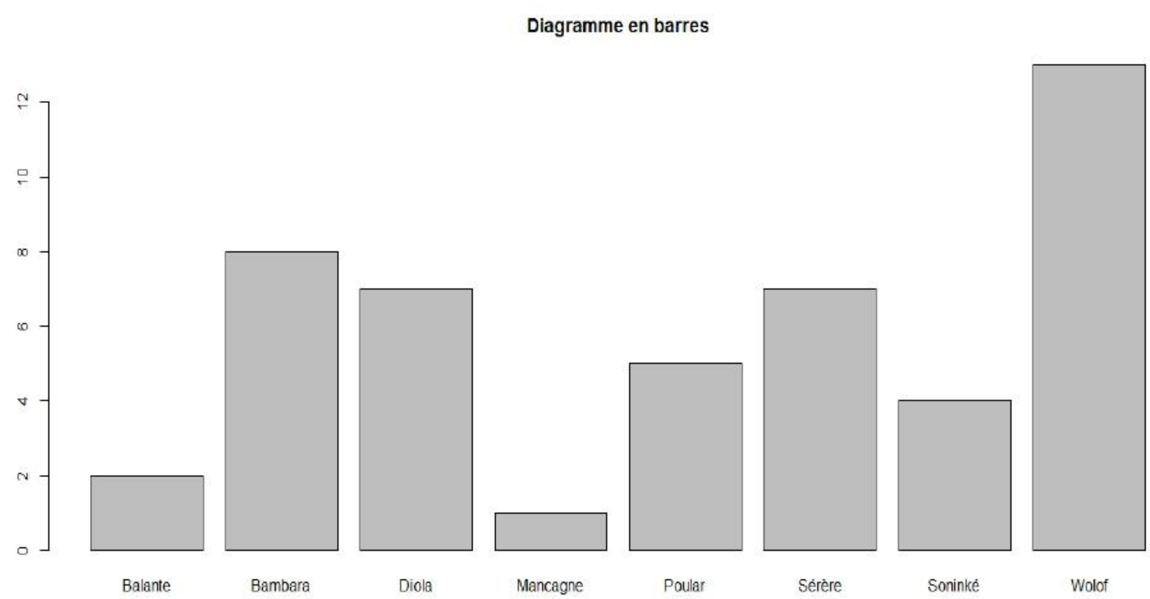
Pour dresser le tableau statistique on procède comme suit :
Ethnie=c("Bambara", "Bambara", "Bambara", "Wolof", "Sérère", "Wolof", "Wolof", "Wolof", "Diola", "Diola", "Wolof", "Diola", "Wolof", "Poular", "Soninké", "Diola", "Sérère", "Wolof", "Soninké", "Wolof", "Sérère", "Sérère", "Diola", "Poular", "Balante", "Poular", "Sérère", "Balante", "Wolof", "Wolof", "Diola", "Poular", "Mancagne", "Soninké", "Bambara", "Wolof", "Sérère", "Bambara", "Diola", "Wolof", "Bambara", "Wolof", "Bambara", "Sérère", "Poular", "Bambara", "Soninké")
table(Ethnie)
Voici ce que l'on obtient
Ethnie

Balante	Bambara	Diola	Mancagne	Poular	Sérère	Soninké	Wolof
2	8	7	1	5	7	4	13

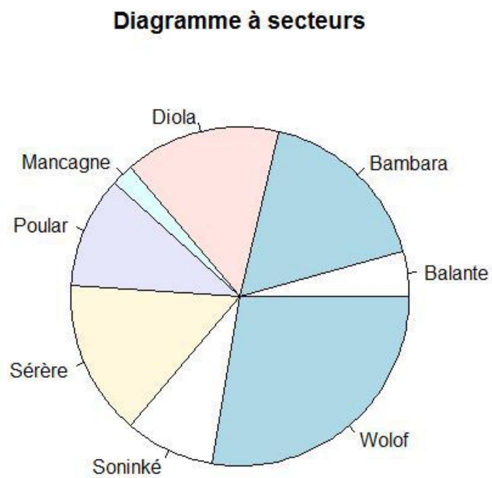
Pour obtenir le tableau statistique avec les fréquences, on exécute les commandes suivantes :
Eth=c(table(Ethnie))
data.frame(Effectifs=Eth,Fréquences=Eth/sum(Eth))
Cette dernière ligne de commandes permet d'obtenir les résultats à 10^{-3} près.

	Effectifs	Fréquences
Balante	2	0,043
Bambara	8	0.170
Diola	7	0.149
Mancagne	1	0.021
Poular	5	0.106
Sérère	7	0.149
Soninké	4	0.085
Wolof	13	0.277

Pour obtenir le diagramme en barres, on utilise les commandes suivantes :
barplot(table(Ethnie),main="Diagramme en barres")



Pour obtenir le diagramme à secteurs, on utilise les commandes suivantes :
pie(table(Ethnie),main="Diagramme à secteurs")



5.5.2 Cas d'une variable quantative discrète

Reprenons les données de l'exemple 1 du chapitre 2.

5; 7; 17; 5; 12; 12; 12; 5; 7; 12; 8; 12; 12; 12; 13; 5; 12; 8; 8; 8;

12; 13; 12; 5; 6; 7; 12; 5; 17; 17; 12; 12; 12; 12; 7; 17; 8; 5; 8; 17.

Pour dresser le tableau statistique on procède comme suit :
nombre_de_selctions=c(5,7, 17, 5, 12, 12, 12, 5, 7, 12, 8, 12, 12, 12, 13, 5, 12, 8, 8, 8, 12, 13, 12, 5, 6, 7, 12, 5, 17, 17, 12, 12, 12, 12, 7, 17, 8, 5, 8, 17)
Tableau=table(nombre_de_selctions) # Donne le tableau statistique
Pour obtenir le tableau statistique(Modalités, Effectifs, Fréquences), il suffit d'excécuter les commandes suivantes :
se1=c(Tableau)
data.frame(Effectifs=se1,Fréquences=se1/sum(se1))
Voici ce que l'on obtient

	Effectifs	Fréquences
5	7	0.175
6	1	0.025
7	4	0.100
8	6	0.150
12	15	0.375
13	2	0.050
13	5	0.125

Pour obtenir le tableau statistique (Modalités, Effectifs, Effectifs cumulés croissants (ECC)), il suffit d'excécuter les commandes suivantes :
se1=c(Tableau)
data.frame(Effectifs=se1,ECC=cumsum(se1))
Voici ce que l'on obtient

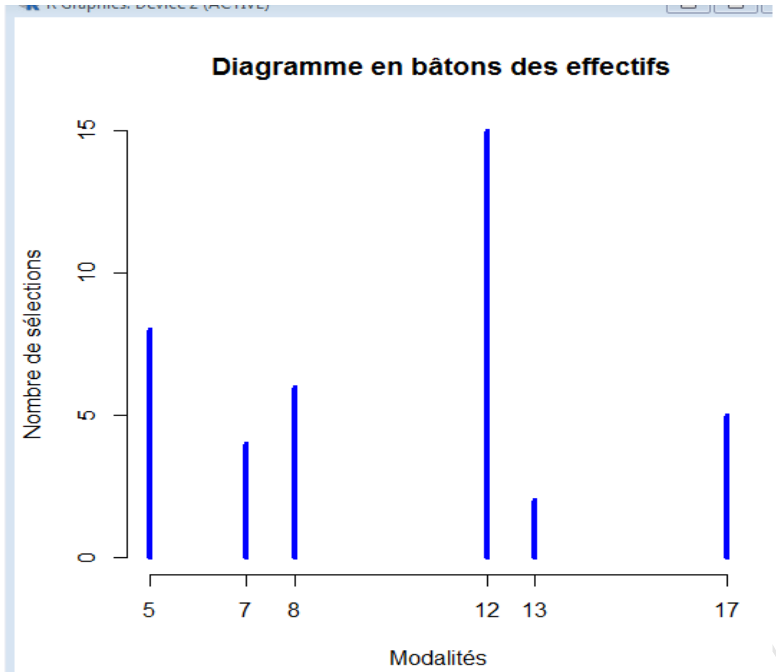
	Effectifs	ECC
5	7	7
6	1	8
7	4	12
8	6	18
12	15	33
13	2	35
13	5	40

Pour obtenir les paramètres de position (valeurs minimale et maximale, médiane, quartiles, moyenne), il suffit de taper la commande suivante :
summary(nombre_de_selctions)
Voici ce que l'on obtient

```
> summary(nombre_de_selctions)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    5.0     7.0    12.0    10.2    12.0    17.0
>
```

Pour obtenir le diagramme en bâtons, on utilise les commandes suivantes :
w=c(rep(5,8),rep(7,4),rep(8,6),rep(12,15),rep(13,2),rep(17,5))
plot(table(w), type="h",col="blue",lwd=4,,xlab="Modalités", ylab="Nombre de sélections", main="Diagramme en bâtons des effectifs",frame=0)

Pour obtenir le diagramme cumulatif des fréquences, on utilise les commandes suivantes :
w=c(rep(5,8),rep(7,4),rep(8,6),rep(12,15),rep(13,2),rep(17,5))
plot(ecdf(w),col="red",lwd=4,xaxt="n",xlab="Modalités",ylab="Fréquences cumulées", main=" Diagramme cumulatif des fréquences",frame=0)
axis(1,c(0,5,7,8,12,13,17,25))



5.5.3 Cas d'une variable quantitative continue

Reprenons les données de l'exemple 2 du chapitre 2 relatives au relevé des tailles (en centimètre) des 36 professeurs d'un établissement d'enseignement moyen :

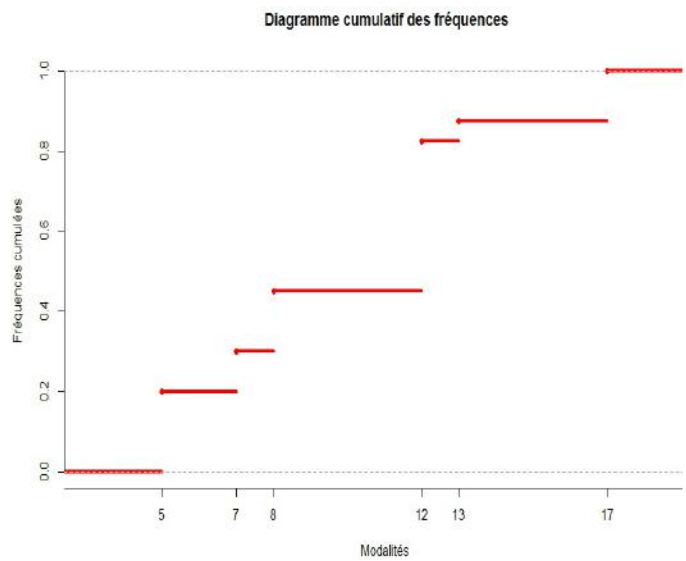
155; 164; 187; 185; 176; 158; 157; 158; 185; 175; 175; 180;

185; 190; 180; 165; 175; 180; 158; 185; 172; 180; 185; 175;

198; 155; 158; 168; 190; 198; 180; 180; 175; 195; 176; 158.

Pour obtenir le tableau statistique(Classes, Effectifs, Fréquences), il suffit d'excécuter les commandes suivantes :

```
v=c(155, 164, 187, 185, 176, 158, 157, 158, 185, 175, 175, 180,
185, 190, 180, 165, 175, 180, 158, 185, 172, 180, 185, 175,
198, 155, 158, 168, 190, 198, 180, 180, 175, 195, 176, 158)
tableau2=table(cut(v,breaks=c(150,160,170,180,190,200),right=FALSE))
se2=c(tableau2)
round(data.frame(Effectifs=se2,Fréquences=se2/sum(se2)),3)
```



Voici ce que l'on obtient

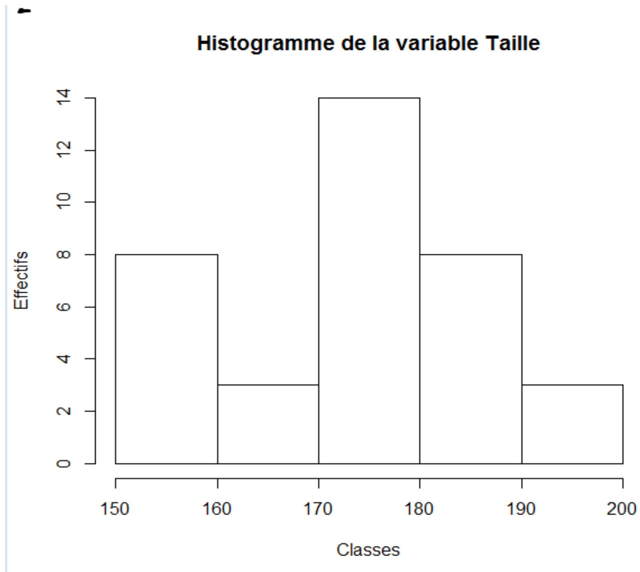
	Effectifs	Fréquences
$[150, 160[$	8	0.222
$[160, 170[$	3	0.083
$[170, 180[$	8	0.222
$[180, 190[$	12	0.333
$[190, 200[$	5	0.139

Pour obtenir le tableau statistique (Classes, Effectifs, Effectifs cumulés croissants (ECC)), il suffit d'excécuter les commandes suivantes :
se2=c(tableau2)
data.frame(Classes=se2,ECC=cumsum(se2))
Voici ce que l'on obtient

	Effectifs	Fréquences
$[150, 160[$	8	8
$[160, 170[$	3	11
$[170, 180[$	8	19
$[180, 190[$	12	31
$[190, 200[$	5	36

Pour obtenir les paramètres de position (valeurs minimale et maximale, médiane, quartiles, moyenne), il suffit de taper la commande suivante :
summary(v)
Pour obtenir l'histogramme des effectifs, il suffit de taper la ligne de commandes suivantes :
hist(v,breaks=c(150,160,170,180,190,200),xlab="Classes",ylab="Effectifs", main="Histogramme

de la variable Taille")



Le diagramme cumulatif croissant est la ligne polygonale qui joint les points de coordonnées : (150, 0) ; (160, 0.222) ; (170, 0.305) ; (180,0.527) ; (190, 0.861) et (200, 1).
Pour l'obtenir, il suffit de taper la ligne de commandes suivantes :
plot(c(150,160,170,180,190,200),c(0,0.222,0.305,0.527,0.861,1),type="line",xlab="Classes", ylab="Fréquences cumulées croissantes",main="Diagramme cumulatif croissant",lwd=3,frame=0)
Voici ce que l'on obtient

5.6 Conclusion

Pour le calcul des paramètres numériques, le logiciel ne fait pas de distinction entre une variable discrète et une variable continue. Car quelque soit le nombre d'observations, le logiciel peut calculer en un temps réduit tous ces paramètres.
Si on a des données regroupées en classes, il sera nécessaire d'effectuer un travail papier crayon pour déterminer certains de ces paramètres.

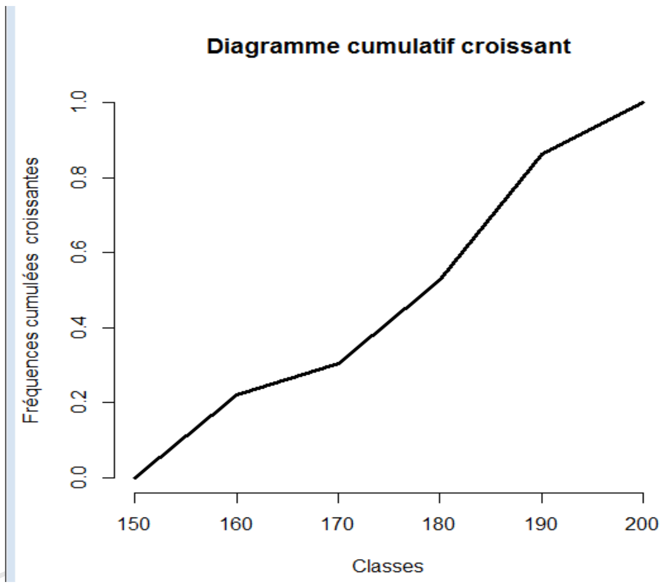


Diagramme cumulatif croissant
Fréquences cumulées croissantes
0.0
0.2
0.4
0.6
0.8
1.0
150
160
170
180
190
200
Classes