

STATISTIQUE DESCRIPTIVE



GHIZLAN LOUMRHARI
FÉVRIER 2024

La séance d'aujourd'hui



- Prise de contact
- Plan du cours
- Bibliographie
- Déroulement du cours
- Evaluation
- Introduction générale

Tables des matières



Plan du cours

1. **Introduction**
2. **Chapitre 1** : Tableaux et graphiques des distributions unidimensionnelles
3. **Chapitre 2** : Les caractéristiques de tendance centrale
4. **Chapitre 3** : Les caractéristiques de dispersion
5. **Chapitre 4** : Les caractéristiques de forme
6. **Chapitre 5** : Les caractéristiques de concentration

Références (sélectives)



- ❑ Gérard Calot, « Cours de statistique descriptive », Dunod.
- ❑ Bernard Py , « Statistique descriptive », Economica.
- ❑ Murray R. Spiegel et Larry J. Stephens, « Statistiques », Mc Graw Hill.
- ❑ Bernard Delmas « Statistique descriptive », Nathan Université.
- ❑ Fouzi Mourji et Najib Mikou « Eléments de mathématiques et statistiques pour économistes »

Déroulement des cours



- ❑ cours est présenté sous forme de slides assez détaillés présentant les principaux concepts, théories, définitions et des applications.
- ❑ Les étudiants doivent participer activement pendant les séances à travers des implications dans la résolutions des différentes applications.
- ❑ Consacrer une séance pour corriger une planche de TD à la fin de chaque chapitre.

Définition de la statistique



Le terme statistique admet trois acceptations possibles :

- ❑ Les statistiques désignent les chiffres et les données se rapportant à un phénomène particulier. Ainsi, on parle des statistiques du chômage pour désigner, par exemple, l'évolution du taux de chômage au Maroc sur une période de temps donnée ;
- ❑ La statistique peut désigner une caractéristique d'un échantillon ; la moyenne dans un échantillon est par exemple une statistique ;
- ❑ La statistique est « une science » qui regroupe l'ensemble des méthodes numériques et mathématiques qui permettent la description, la synthèse et l'explication des données.

Définition de la statistique



La statistique en tant que science est généralement divisée en deux grandes parties :

- ❑ la statistique descriptive dont le but est la description et la présentation simplifiée des données et,
- ❑ la statistique inférentielle, ou mathématique, dont l'objet est d'appréhender les caractéristiques d'une population en se basant sur une information incomplète *i.e.* un échantillon.

Définition de la statistique



- ❑ La statistique descriptive est une « science » qui s'attache à décrire le plus simplement et le plus clairement possible les données et les informations relatives à des grandeurs observées. Pour ce faire, elle utilise des méthodes et des outils, essentiellement mathématiques, qui constituent le corps de cette « science » et qu'on peut regrouper sous l'expression de réduction de données.
- ❑ La statistique descriptive peut s'attacher à décrire les données et les informations statistiques d'un point de vue d'une seule variable, de deux variables ou de plusieurs variables. On parle respectivement de statistique univariée, bivariée ou multivariée (analyse des données).

Définition de la statistique



- ❑ Comme toutes les sciences, la statistique utilise un vocabulaire propre qu'il est nécessaire de connaître et de pouvoir utiliser avec précision.
- ❑ C'est Roland Fisher qui, le premier à utilisé l'expression *réduction des données* pour définir la statistique descriptive. Cette expression désigne le calcul de caractéristiques, qui permettent de résumer une distribution statistique.
- ❑ En réalité, la statistique fut appliquée pour la première fois à la démographie dont elle a hérité la terminologie.

Définition de la statistique



- ❑ Comme toutes les sciences, la statistique utilise un vocabulaire propre qu'il est nécessaire de connaître et de pouvoir utiliser avec précision.
- ❑ C'est Roland Fisher qui, le premier à utilisé l'expression *réduction des données* pour définir la statistique descriptive. Cette expression désigne le calcul de caractéristiques, qui permettent de résumer une distribution statistique.
- ❑ En réalité, la statistique fut appliquée pour la première fois à la démographie dont elle a hérité la terminologie.

Vocabulaire et mots clés



Population : On appelle population un grand ensemble au sens mathématique (fini ou infini) auquel s'intéresse la statistique. À titre d'exemples, peuvent être considérés comme des populations, les ensembles suivants :

- ❑ Les étudiants inscrits en première année à l'ENCG au titre de l'année universitaire 2023 – 2024 ;
- ❑ Le parc automobile de la région Tanger-Tetouan pour l'année 2023.
- ❑ Le nombre d'actions échangées à la bourse de Casablanca pendant une journée donnée.

Vocabulaire et mots clés



Comme on peut le constater, les ensembles considérés comme des populations sont très variés. Le plus important est que ces ensembles doivent être définis sans ambiguïté en comportant les précisions nécessaires telles que la date et le lieu comme dans nos exemples.

Vocabulaire et mots clés



Individus : On appelle individus ou unités statistiques les éléments qui composent une population. Si nous gardons nos précédents exemples, les individus ou les unités statistiques sont : les étudiants, les voitures et les actions.

Ainsi, on appelle individus les éléments qui composent une population qu'ils soient des êtres humains des objets ou toute autre chose.

Vocabulaire et mots clés



Caractère : Les individus d'une population peuvent être identifiés en fonction de plusieurs caractéristiques.

Chaque caractéristique est appelée caractère.

Un caractère est tout ce qui peut différencier un individu d'un autre : L'âge des étudiants, le niveau de qualification des employés, la couleur des voitures, par exemples, sont des caractères.

Vocabulaire et mots clés



Modalité : La modalité est la situation ou l'état dans lequel peut se trouver un caractère. Si le caractère considéré est la couleur des voitures, par exemple, les modalités sont : noire, blanche, rouge, etc. Si le caractère étudié est l'âge, les modalités seront alors 18ans, 19 ans, etc. Enfin, le caractère sexe admet deux modalités : masculin et féminin.

Vocabulaire et mots clés



Echantillon : Ainsi nous l'avons déjà précisé, la statistique s'intéresse à l'étude de grands ensembles (parfois infinis), il est souvent difficile voire impossible, ou simplement très coûteux d'étudier toute la population. Le seul moyen dont nous disposons pour avoir des informations sur la population à laquelle on s'intéresse est d'en observer une partie. Le sous-ensemble observé est appelé échantillon.

C'est pour cette raison que les recensements des populations sont très espacés dans le temps.

Vocabulaire et mots clés



Enquête : On appelle enquête l'ensemble des moyens mis en œuvre pour recueillir des informations relatives à une population.

- Lorsque l'enquête porte sur tous les individus de la population on parle d'enquête **exhaustive ou de recensement**. C'est le cas par exemple du recensement général de la population.
- Lorsque l'enquête porte sur une partie seulement de la population, on parle **d'enquête partielle ou de sondage**. C'est le cas par exemple des sondages d'opinions.
- Parmi tous ces termes qu'on vient de présenter brièvement, le concept fondamental en statistique est celui de caractère c'est pourquoi nous allons détailler .

Typologie des caractères



Un caractère est toute caractéristique qui permet de différencier et de distinguer les individus d'une même population. Si par exemple on s'intéresse à la population des entreprises installées dans la région de Tanger, il est possible de les classer et de les différencier, entre autres, par rapport aux caractéristiques suivantes :

- Effectif (nombre de salariés employés) ;
- Catégorie socioprofessionnelle ;
- Chiffre d'affaires ;
- Niveau d'étude des employés ;
- Masse salariale ;
- Secteur d'activité ;
- Nombre de filiales ;
- Sexe du PDG.

Typologie des caractères



Ces caractéristiques sont autant de caractères. Dans ces cas les modalités sont

- 20, 50, 200, 1000, etc. ;
- Ouvrier, employé, cadre, etc. ;
- 50000, 200000, 1000000 ;
- Primaire, secondaire, supérieur ;
- 30000, 469400, 750890 ;
- Agroalimentaire, textile, Industrie chimie, services, etc. ;
- 2, 3, 4 ;
- Féminin, masculin ;

Typologie des caractères



Les caractères qualitatifs

- ❑ Un caractère est dit qualitatif s'il n'est pas mesurable *i.e.* ses modalités ne sont pas des valeurs numériques mais des états, des catégories, des opinions, des couleurs, etc. Dans notre exemple, les caractères qualitatifs sont le niveau d'études et le secteur d'activité.
- ❑ Les caractères qualitatifs peuvent être de deux types : les variables **qualitatives nominales** et les variables qualitatives ordinales. Lorsque la moyenne est impossible à calculer et elles ne peuvent pas être ordonnées (leur classement n'a pas de sens). C'est le cas du sexe, de la nationalité ou du secteur d'activité dans notre exemple. Par contre, lorsque la moyenne peut être calculée mais ne présente aucun sens mais que le classement des modalités est possible et admet un sens et un intérêt, le caractère est dit **qualitatif ordinal**. C'est le cas du niveau d'études et de la catégorie socioprofessionnelle par exemple.

Typologie des caractères



Les caractères quantitatifs

- ❑ Un caractère est dit quantitatif lorsque ses modalités sont mesurables : des valeurs numériques et que la moyenne de ces valeurs a un sens. Dans ce cas, on parle de variables statistiques. Ces dernières se scindent en deux catégories : **les variables statistiques discrètes et les variables statistiques continues.**
- ❑ Une variable statistique est dite **discrète** lorsque ses modalités ne peuvent pas prendre n'importe quelle valeur sur un intervalle donné mais sont un ensemble fini ou dénombrable de valeurs entières et réelles. C'est le cas de la variable "nombre d'enfants par ménage" qui admet 0, 1, 2, 3, ... comme modalités. Elle ne peut pas prendre de valeur décimale ; la modalité 2,3 enfants n'a pas de sens. Dans notre exemple l'effectif est une variable statistique discrète.

Typologie des caractères



Les caractères quantitatifs

- ❑ Une variable statistique est dite **continue** lorsqu'elle peut prendre n'importe quelle valeur sur un intervalle donné c.a.d. lorsque le nombre de modalités est infini. La variable "poids d'un individu" peut prendre les valeurs 70 kg ou 80 kg par exemple, mais elle peut prendre n'importe quelle valeur intermédiaire. Ainsi, les modalités 70 kg 100 ou 73 kg 99 peuvent très bien exister. Dans notre exemple, le chiffre d'affaire et la masse salariale sont des variables continues.
- ❑ Les variables statistiques continues lorsqu'elles prennent un nombre très élevé de modalités peuvent être assimilées à des variables continues.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles



1.LES TABLEAUX STATISTIQUES

2.LES REPRÉSENTATIONS GRAPHIQUES

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles



La première tâche du statisticien, avant de procéder à la réduction des données, est de présenter les informations et les données collectées lors d'enquêtes sous forme de tableaux et de graphiques afin d'en faciliter la lecture.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

1. Définition

Un tableau statistique ou distribution statistique est un tableau qui met en relation d'un côté les modalités notés x_i et de l'autre côté le nombre d'individus n_i présentant ce caractère.

Tab-1. Distribution des fréquences absolues

Modalités x_i	Fréquence absolues n_i
x_1	n_1
x_2	n_2
\vdots	\vdots
x_i	n_i
\vdots	\vdots
x_k	n_k
Total	$n = \sum n_i$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

2. Exemple

Supposons qu'on souhaite avoir une idée sur le niveau des étudiants en statistique descriptive durant l'année universitaire 2023-2024. On réalise un sondage auprès de 10 étudiants et on obtient le résultat suivant :

15 ; 12 ; 08 ; 12 ; 10 ; 15 ; 18 ; 12 ; 14 ; 12

Cette liste de données est appelée **données brutes**. Elles doivent être traitées. Elles peuvent être présentées dans un premier temps sous forme d'une série statistique c.a.d. sous forme d'une suite de valeurs ordonnées

08 ; 10 ; 10 ; 12 ; 12 ; 12 ; 12 ; 15 ; 15 ; 18

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques



2. Exemple

Ensuite, lorsque la taille de l'échantillon est importante, elles peuvent être présentées sous forme d'une distribution des fréquences

Tab-2. Distribution des fréquences absolues

Note x_i	Effectif n_i
8	1
10	2
12	4
15	2
18	1
Total	$n = 10$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

2. Exemple Tab-2. Distribution des fréquences absolues

Note x_i	Effectif n_i
8	1
10	2
12	4
15	2
18	1
Total	$n = 10$

□ Ce tableau donne le nombre d'individus qui présentent chacune des modalités.

- Il est préférable de donner plutôt la part d'individus qui présentent chacune des modalités ce qui s'obtient en utilisant le concept de fréquence relative.
- La fréquence relative, notée f_i , est la part d'individus représentant la modalité. C'est le rapport entre la fréquence absolue et l'effectif total

$$f_i = \frac{n_i}{n} \equiv \frac{n_i}{n} \times 100$$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques



2. Exemple

Si nous reprenons notre exemple, nous obtenons :

Tab-3. Distribution des fréquences relatives

Note x_i	Effectif n_i	Fréquence f_i	Fréquence f_i en %
8	1	0,1	10
10	2	0,2	20
12	4	0,4	40
15	2	0,2	20
18	1	0,1	10
Total	$n = 10$	$\sum f_i = 1$	$\sum f_i = 100$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

2. Exemple

Tab-3. Distribution des fréquences relatives

Note x_i	Effectif n_i	Fréquence f_i	Fréquence f_i en %
8	1	0,1	10
10	2	0,2	20
12	4	0,4	40
15	2	0,2	20
18	1	0,1	10
Total	$n = 10$	$\sum f_i$ = 1	$\sum f_i$ = 100

La lecture du tableau se fait de la manière suivante :

- ❑ 10% des étudiants ont obtenu une note égale à 8/20 ;
- ❑ 20% des étudiants ont obtenu une note égale à 10/20.

La distribution des fréquences telle qu'elle vient d'être présentée est le tableau de base en statistique. Il peut contenir d'autres informations selon la nature du caractère.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

3. les caractères qualitatifs

Exemple. Nous avons réalisé une enquête auprès d'une entreprise de 100 salariés qui a relevé qu'elle emploie 60 ouvriers, 30 ouvriers qualifiés et 10 cadres.

Il est possible alors de construire la distribution suivante :

Tab-4. Distribution des fréquences relatives

Catégorie x_i	Effectif n_i	Fréquence f_i
Ouvriers	60	0,6
Ouvriers qualifiés	30	0,3
Cadres	10	0,1
Total	$n = 100$	$\sum f_i = 1$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables discrètes

- ❑ Les opérations appliquées aux caractères qualitatifs demeurent valables dans le cas des variables statistiques. Mais comme le caractère est ordonné on pourra cumuler les fréquences.
- ❑ Il sera alors possible de répondre directement (en lisant simplement le tableau) à des questions de type :

« combien d'individus présentent une modalité est au moins égale à une certaine valeur » ou bien « combien d'individus présentent une modalité est au plus égale à une certaine valeur ».

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables discrètes

Répondre à ce type de questions revient à utiliser le concept « fréquences cumulées ». Ces dernières se calculent en additionnant les fréquences absolues ou relatives et on les notes N_i et F_i :

- N_i^+ : Fréquences absolues cumulées croissantes
- N_i^- : Fréquences absolues cumulées décroissantes
- F_i^+ : Fréquences relative cumulées croissantes
- F_i^- : Fréquences relative cumulées décroissantes

Les calculs peuvent se faire dans un même tableau

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables discrètes

Tab-5. Distribution des fréquences relatives et fréquences cumulées

x_i	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
x_1	n_1	f_1	n_1	$n_k + \dots + n_1$ $= n$	f_1	$f_k + \dots + f_1$
x_2	n_2	f_2	$n_1 + n_2$	$n_k + \dots + n_2$	$f_1 + f_2$	$f_k + \dots + f_2$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_i	n_i	f_i	$n_1 + \dots + n_i$	$n_k + \dots + n_i$	$f_1 + \dots + f_i$	$f_k + \dots + f_i$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_k	n_k	f_k	$n_1 + \dots + n_k = n$	n_k	$f_1 + \dots + f_k = 1$	f_k

Nombre d'individus

Part (%) d'individus

Total	$\sum = n$	$\sum = 1$	Au plus (moins de)	Au moins (plus de)	Au plus (moins de)	Au moins (plus de)
--------------	------------	------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------



Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables discrètes

Reprenons notre exemple

Tab-6. Distribution des fréquences relatives et fréquences cumulées

x_i	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
8	1	0,1	1	10	0,1	1
10	2	0,2	3	9	0,3	0,9
12	4	0,4	7	7	0,7	0,7
15	2	0,2	9	3	0,9	0,3
18	1	0,1	10	1	1	0,1
Total	$n = 10$	$\sum = 1$	Au plus (moins de)	Au moins (plus de)	Au plus (moins de)	Au moins (plus de)

Ainsi on peut voir que :

- 3 étudiants ont obtenu au plus 10/20
- 90% des étudiants ont obtenu au moins 10/20

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables discrètes

Application : une enquête réalisée au Maroc, auprès de 40 ménages, donne la distribution suivante sur le nombre d'enfants par ménage.

Tab-7. Distribution du nombre d'enfants par ménage

x_i	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
0	2	0,05	2	40	0,05	1
1	4	0,1	6	38	0,15	0,95
2	6	0,15	12	34	0,3	0,85
3	14	0,35	26	28	0,65	0,7
4	8	0,2	34	14	0,85	0,35
5	4	0,1	38	6	0,95	0,15
5 et +	2	0,05	40	2	1	0,05
Total	$n = 40$	$\sum = 1$	Au plus (moins de)	Au moins (plus de)	Au plus (moins de)	Au moins (plus de)

- (i) Le nombre de ménages ayant au plus 3 enfants
- (ii) Le pourcentage de ménages qui ont au moins 2 enfants
- (iii) Le nombre de ménages qui ont un nombre d'enfants inférieur à 4 enfants
- (iv) Le pourcentage de ménages ayant plus de 4 enfants

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables discrètes

Réponse : Après avoir calculé : N_i^+ , N_i^- , F_i^+ et F_i^-

- (i) Le nombre de ménages ayant au plus 3 enfants est 26 soit : $N_3^+ = n_0 + n_1 + n_2 + n_3 = 2 + 4 + 6 + 14 = 26$
- (ii) Le pourcentage de ménages ayant au moins 2 enfants est 85%, soit : $F_2^- = f_2 + f_3 + f_4 + f_5 + f_{5et+} = 0,15 + 0,35 + 0,2 + 0,1 + 0,05 = 0,85 = 85\%$
- (iii) Un nombre d'enfants inférieur à 4 : 0,1,2,3 soit au plus 3 enfants, donc on cherche $N_3^+ = 26$
- (iv) Le pourcentage de ménages ayant plus de 4 enfants : 5 et 5 et + soit au moins 5, donc on cherche $F_5^- = 15\%$
Ainsi, 15% de ménages ont plus de 4 enfants.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables Continues

Définitions et notations : lorsqu'on a affaire à une variable continue, le nombre de modalités est très élevé. Il devient alors très utile de les ranger en plusieurs classes, auxquelles on fait correspondre le nombre d'individus (ou leur %).

Chaque classe, notée i , est symbolisée par un intervalle avec deux limites ; inférieure et supérieure notées b_{i-1} et b_i . Par convention, on retient un intervalle fermé à gauche et ouvert à droite. Si certains salariés d'une entreprise gagnent entre 2000 et 3000DH, la classe s'écrira $[2000,3000[$.

La taille ou la largeur de l'intervalle est appelé l'amplitude et notée a_i ; c'est la différence entre les deux limites. Soit : $a_i = b_i - b_{i-1}$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables Continues

Selon les cas, les amplitudes peuvent soit varier soit être constantes d'une classe à une autre. Parfois, il est nécessaire d'opérer une correction des fréquences. Enfin pour terminer avec nos définitions, notons qu'on appelle valeur de la classe ou centre de la classe, notée c_i , une valeur qui sert à représenter une classe. Elle s'obtient en additionnant les deux limites et en divisant par deux. Soit : $c_i = \frac{(b_i + b_{i-1})}{2}$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables Continues

Exemple : la distribution des salaires dans une entreprise de 200 salariés est donnée par le tableau suivant :

Tab-8. Distribution des salaires

Catégories	x_i	Effectifs n_i	Amplitudes a_i	Centres des classes c_i
Ouvriers	[1500,2500[80	1000	2000
Ouvriers qualifiés	[2500,3500[60	1000	3000
Employés	[3500,5500[40	2000	4500
Cadres	[5500,8500[20	3000	7000
Total		$n = 200$		

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables Continues

Exemple : Correction des fréquences

Tab-9. Distribution des salaires et fréquences corrigées

Catégories	x_i	n_i	f_i	a_i	a_i/\bar{a}	$f_i / (a_i/\bar{a})$
Ouvriers	[1500,2500[80	0,4	1000	1	0,4
Ouvriers qualifiés	[2500,3500[60	0,3	1000	1	0,3
Employés	[3500,5500[40	0,2	2000	2	0,1
Cadres	[5500,8500[20	0,1	3000	3	0,033
Total		$n = 200$	$\sum = 1$	$\bar{a} = 1000$		

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

1. Les tableaux statistiques

4. les variables quantitatives / Les variables Continues

Exemple : Pour le reste, il est toujours possible de calculer les fréquences cumulées :

Tab-10. Distribution des salaires et fréquences relatives

x_i	n_i	f_i	N_I^+	N_i^-	F_i^+	F_i^-
[1500,2500[80	0,4	80	200	0,4	1
[2500,3500[60	0,3	140	120	0,7	0,6
[3500,5500[40	0,2	180	60	0,9	0,3
[5500,8500[20	0,1	200	20	1	0,1
	$n = 200$	$\sum = 1$	Moins de	Plus de	Moins de	Plus de

Le tableau se lit de la manière suivante :

- 20% des salariés (soit 40) gagnent entre 3 500 et 5 500DH
- 90% des salariés (soit 180) gagnent moins de 5 500DH
- 60% des salariés (soit 120) gagnent plus de 2 500DH

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques



Malgré le fait que les tableaux statistiques contiennent la totalité de l'information statistique collectée, on a souvent recours à des représentations graphiques pour mieux faciliter et visualiser leurs lectures. L'inconvénient des graphiques c'est qu'ils sont moins précis que les tableaux.

Il existe plusieurs types de graphiques, le choix de celui à utiliser dépend de la nature du caractère étudié.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

1. les caractères qualitatives / Le graphique à secteurs

Deux types de graphiques sont essentiellement utilisés dans le cas des caractères qualitatifs : **le graphique à secteurs** et **le graphique en tuyaux d'orgue**.

Ce graphique appelé également **diagramme circulaire** ou **camembert** (en référence au célèbre fromage français), s'obtient en divisant un cercle en autant de parts (secteurs) que le caractère présente de modalités. L'angle de chaque part noté w_i , est proportionnel à la fréquence de la modalité. Soit

$$w_i = 360^\circ \times f_i$$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

1. les caractères qualitatifs / Le graphique à secteurs

Exemple : reprenons notre exemple sur la distribution des qualifications

Tab-11. Distribution des qualifications

x_i	n_i	f_i	$w_i = f_i \times 360^\circ$
Ouvriers	60	0,6	216°
Ouvriers qualifiés	30	0,3	108°
Cadres	10	0,1	36°
Total	$n = 100$	$\sum = 1$	$\sum = 360^\circ$

Le graphique à secteurs se présente de la manière suivante :

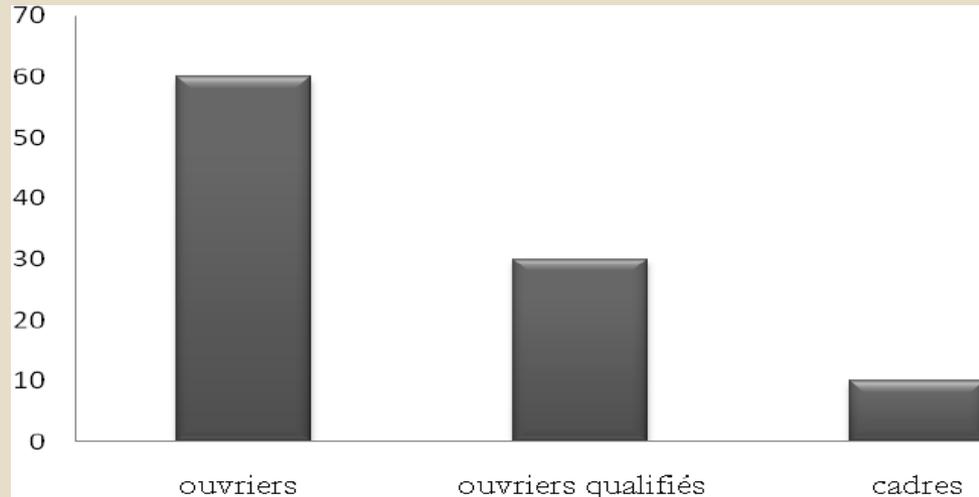


Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

1. les caractères qualitatifs / Le diagramme en tuyaux d'orgue

Dans ce graphique, on représente chaque modalité par un rectangle. Tous les rectangles ont la même base mais leur hauteur est proportionnelle aux fréquences des modalités.



Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les caractères discrètes / Le diagramme en bâtons

Le diagramme en bâtons sert à représenter les fréquences d'une variable discrète. Il s'agit d'un graphique tracé sur un plan à deux dimensions. On porte les modalités, x_i sur l'axe des abscisses et les fréquences (n_i ou f_i) sur l'axe des ordonnées. A chaque modalité x_i on fait correspondre un segment vertical (un bâton) de longueur proportionnelle à n_i ou f_i

Exemple. Reprenons notre exemple sur la distribution du nombre d'enfants par ménage.

2.2.1 Tab-12. Distribution du nombre d'enfants

x_i	n_i	f_i
0	2	0,05
1	4	0,1
2	6	0,15
3	14	0,35
4	8	0,2
5	4	0,1
5 et +	2	0,05
Total	$n = 40$	$\sum = 1$

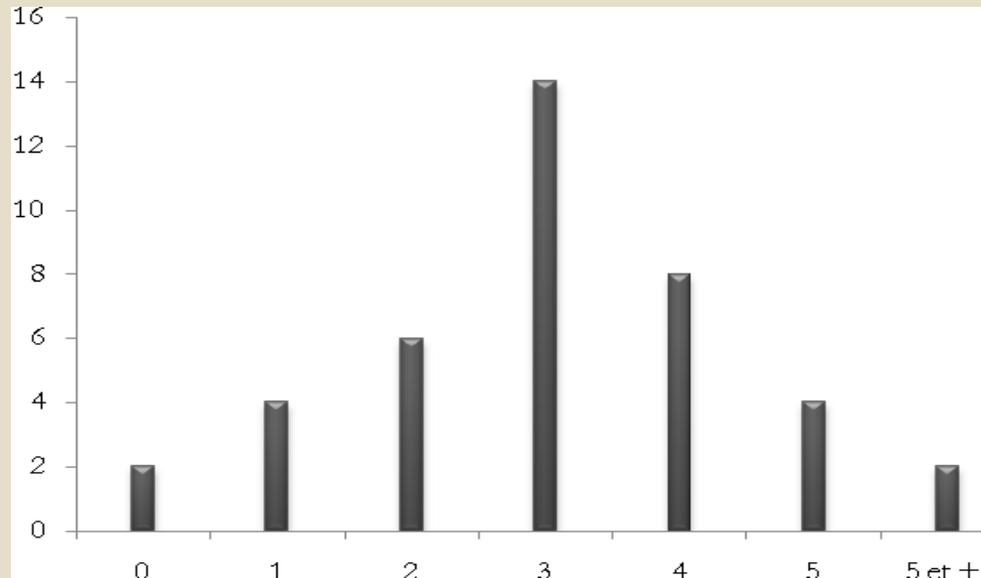
$$\sum = 1$$

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les caractères discrètes / Le diagramme en bâtons

x_i	n_i	f_i
0	2	0,05
1	4	0,1
2	6	0,15
3	14	0,35
4	8	0,2
5	4	0,1
5 et +	2	0,05
Total	$n = 40$	$\sum f_i = 1$



Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les caractères discrètes / La courbe cumulative

Cette courbe sert à représenter les fréquences cumulées. On donne l'exemple de la courbe relative à la fréquence relative cumulée. C'est-à-dire la courbe de la fonction $F(x_i)$. Sur l'axe des abscisses on porte les modalités x_i et sur l'axe des ordonnées les valeurs F_i^+ .

Reprenons notre exemple et calculons les F_i^+

Tab-13. Distribution du nombre d'enfants et fréquence cumulée

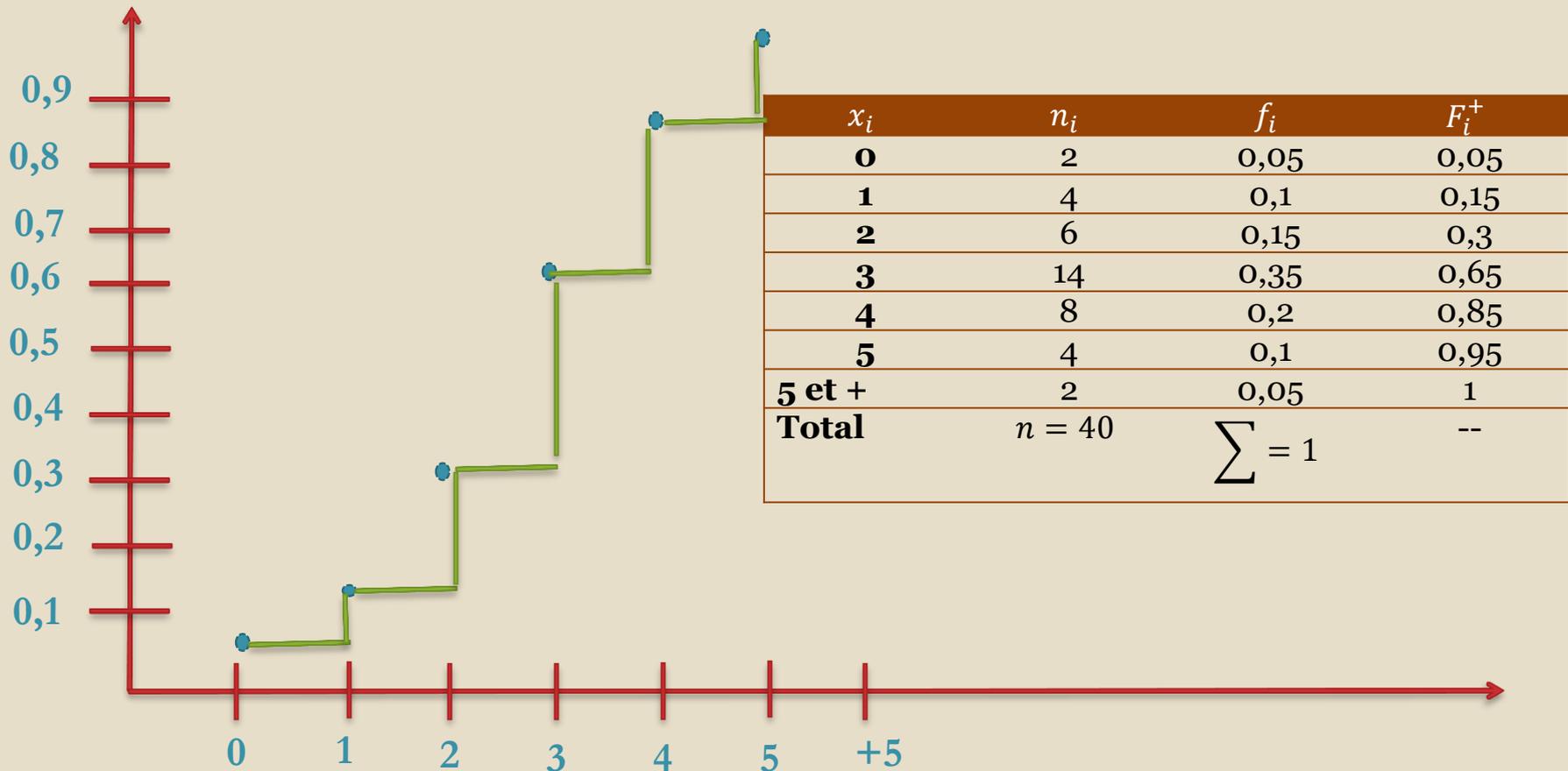
x_i	n_i	f_i	F_i^+
0	2	0,05	0,05
1	4	0,1	0,15
2	6	0,15	0,3
3	14	0,35	0,65
4	8	0,2	0,85
5	4	0,1	0,95
5 et +	2	0,05	1
Total	$n = 40$	$\sum = 1$	--

Pour la courbe cumulative elle se présente en escaliers avec des paliers horizontaux.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les caractères discrètes / La courbe cumulative



Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les variables continues

Pour les variables continues, les deux graphiques appropriés sont **l'histogramme** et **la courbe cumulative**.

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les variables continues / L'histogramme

L'histogramme est un graphique qui ressemble au diagramme en tuyaux d'orgue. C'est-à-dire des rectangles. Sauf que maintenant sont des bases proportionnelles aux amplitudes et que leurs surfaces sont proportionnelles aux fréquences.

Lorsque les amplitudes ne sont pas constantes, il faut corriger les effectifs.

Exemple. Reprenons notre exemple avec la distribution des salaires

Tab-14. Distribution des salaires

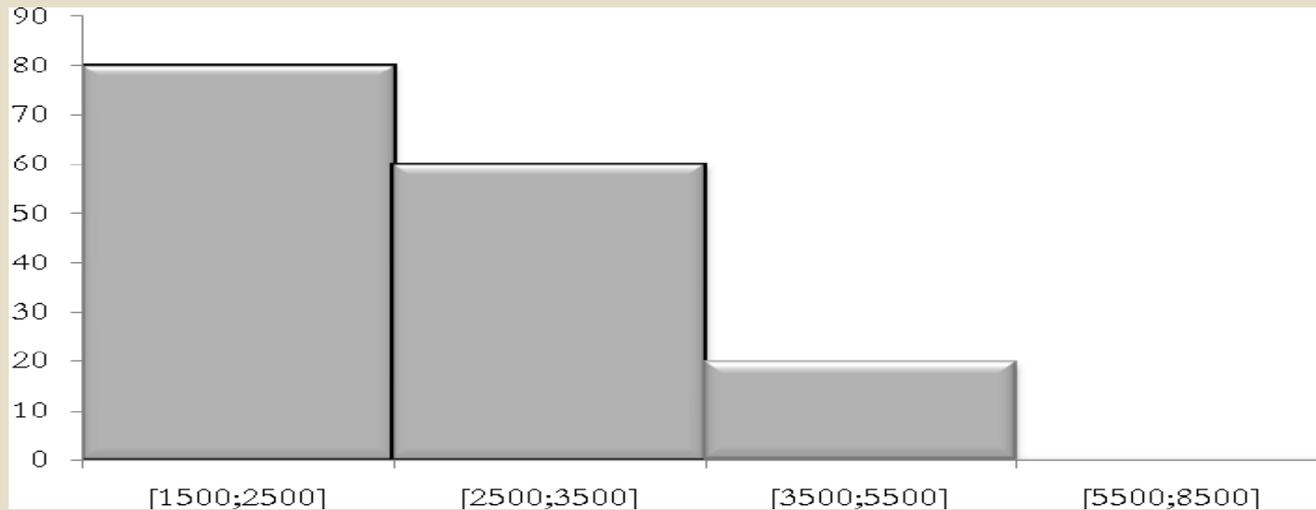
x_i	n_i	f_i	a_i	a_i/\bar{a}	n_i^c
[1500,2500[80	0,4	1000	1	80
[2500,3500[60	0,3	1000	1	60
[3500,5500[40	0,2	2000	2	20
[5500,8500[20	0,1	3000	3	6,6
Total	$n = 200$	$\sum = 1$	\bar{a} $= 1000$		

Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les variables continues / L'histogramme

Histogramme des fréquences



Chapitre 1. Tableaux et graphiques des distributions unidimensionnelles

2. Représentations graphiques

2. les variables continues / La courbe cumulative

De la même manière que pour les variables discrètes, il est possible de tracer la courbe des fréquences cumulées. Pour la courbe des fréquences relatives cumulées croissantes par exemple, on obtient une courbe monotone croissante.

Exemple. Calculons F^+ dans l'exemple précédent.

Tab-15. Distribution des salaires et fréquence cumulée

x_i	n_i	f_i	F_i^+
[1500,2500[80	0,4	0,4
[2500,3500[60	0,3	0,7
[3500,5500[40	0,2	0,9
[5500,8500[20	0,1	1
Total	$n = 200$	$\sum = 1$	