



STATISTIQUE DESCRIPTIVE

Enseignant : Houssam Hamdouch

Email: hamdouchhoussam@gmail.com

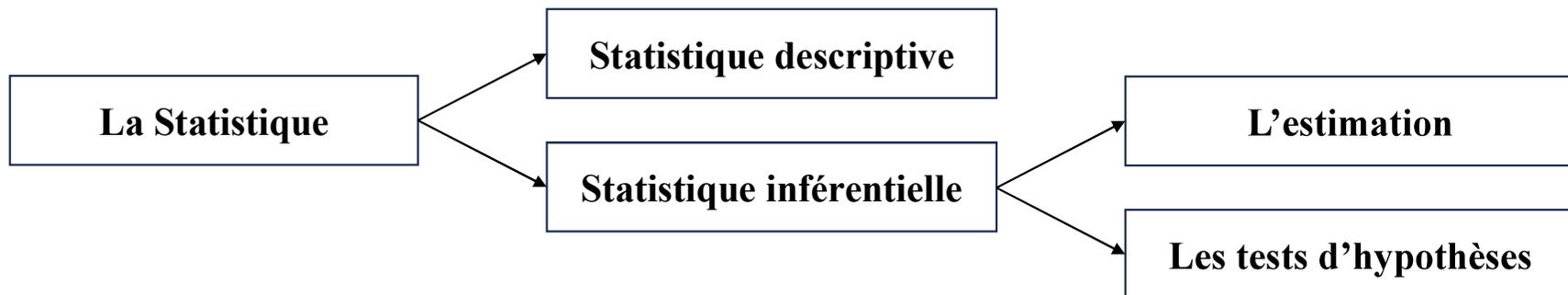
2024/2025

Introduction Générale

- **La statistique et Les statistiques**

- **Les statistiques** : est un ensemble de données ou d'informations relative à un phénomène, Ce sont des données chiffrés (moyenne, pourcentage, etc)
- Exemple : - Les importation du Maroc entre 2014 et 2024,
- **La statistique** : est un ensemble de *méthodes* qui permettent de décrire et analyser des données, Elle nous permet de tirer des conclusions, et de prendre des décisions raisonnables.

La statistique peut se distinguer en deux approches



Plan du Cours

01

Statistique descriptive a une dimension

1. Vocabulaire,
 2. Caractères statistiques,
 3. Tableaux et représentations graphiques,
 4. Paramètres de position,
 5. Paramètre de dispersion,
 6. Paramètre de forme,
 7. Paramètre de concentration
-

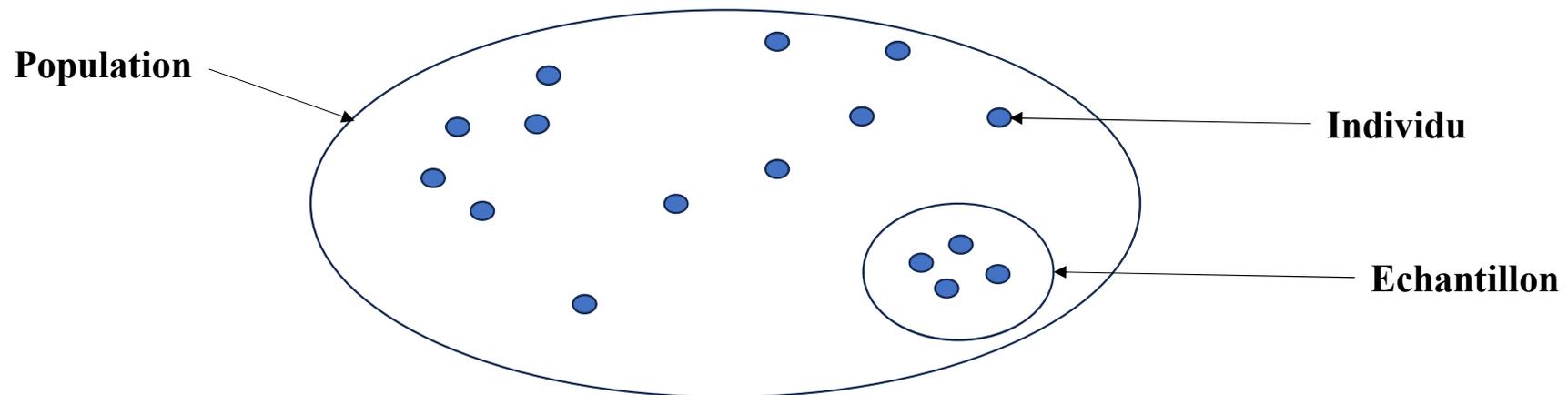
02

Statistique descriptive à deux dimensions

1. Les tableaux statistiques à deux caractères
 2. L'ajustement linéaire
-

1. Vocabulaire

- **Population** : Est un ensemble d'objets homogènes sur lesquels porte une étude (aux quel on s'intéresse),
Exemple : Etudiants d'une classe, Habitants d'un quartier, La population des PME. Du Maroc.
- **Individus (unités statistiques)** : Sont les éléments qui composent une population.
- **Echantillon** : Un sous ensemble d'une population.
- **Taille de la population** : Le nombre des individus qui la composent,
- **Série statistique** : correspond à la représentation simultanée, généralement sous forme de tableau, des valeurs du caractère étudié.

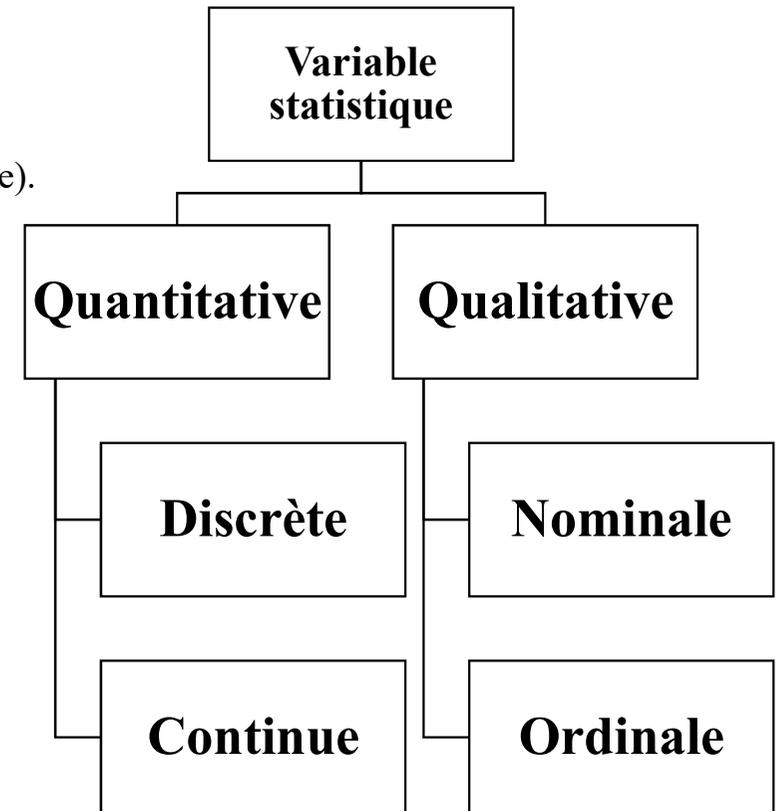


2. Caractère statistique

- On appelle **caractère statistique** (ou **variable statistique**) un critère relatif permettant d'effectuer des observations sur les individus d'une population.
- **Modalité** : toute valeur qui peut prendre une variable statistique (un caractère).

Les caractères attribuables à chaque individu de la population, peuvent être :

- **Quantitatives** mesurés par des nombres, elles peuvent prendre deux formes:
 - **Variables discrètes** : se sont des variables dont l'ensemble des modalités est discret : elles ne prennent que des valeurs isolées les unes des autres.
Exemple : nombre d'enfants par ménage, le nombre de véhicules dans une ville.
 - **Variables continues** : prennent des valeurs d'un intervalle (nombre infini de valeurs). Exemple : Salaire, âge, poids, diamètre d'une pièce.
- **Qualitatifs** : désignés par des noms, par exemple : Sexe : « Masculin », « Féminin »; couleur des yeux : « Bleu », « Marron », « Noir », et « Vert »; Mention au Bac, etc.
 - **Variables ordinales** : ses modalités peuvent être *naturellement ordonnées*, autrement dit, cette variable correspond, à *des noms qui peuvent être classés dans un ordre précis*. Exemple : Mention au Bac.
 - **Variables nominales** ou qualitative pure, Exemple couleur des yeux.



Schémas récupératif de type de caractère

3. Tableau et représentation graphique

a). Cas d'un caractère qualitatif et quantitatif discret

Soit X: un caractère qualitatif ou quantitatif discret, étudié sur une population de n individus :

$\{(M_1, n_1), \dots, (M_n, n_k)\}$

Modalités	Effectifs	Fréquences	Fréquences cumulées	...
M_i	n_i	$f_i = \frac{n_i}{\sum_i n_i}$	F_i	
M_1	n_1	f_1	$F_1 = f_1$	
M_2	n_2	f_2	$F_2 = F_1 + f_2$	
\vdots	\vdots	\vdots	\vdots	
M_i	n_i	f_i	$F_i = \sum_{j=1}^i f_j$ $F_i = F_{i-1} + f_i$	
\vdots	\vdots	\vdots	\vdots	
M_k	n_k	f_k	$F_k = 1$	
Total	N	1		

3. Tableau et représentation graphique

a). Cas d'un caractère qualitatif et quantitatif discret

- **Exemple :** Une enquête réalisée dans un village porte sur le nombre d'enfants à charge pour 229 familles.

M_i	n_i	f_i (en %)	F_i (en %)
0	48	20.96	20.96
1	65	28.38	49.34
2	44	19.21	68.56
3	27	11.79	80.35
4	19	8.30	88.65
5	15	6.55	95.20
6	8	3.49	98.69
7	2	0.87	99.56
8	1	0.44	100
Total	229	100	

3. Tableau et représentation graphique

a). Cas d'un caractère qualitatif et quantitatif discret

- **Exemple 2:** On dispose d'un échantillon de 20 familles, interrogées sur le nombre d'enfants. La série obtenue est la suivante : 4 0 1 1 2 2 2 3 3 4 2 3 3 5 2 1 3 3 4 5

x_i	n_i	f_i	ECC_i	ECD_i	FCC_i	FCD_i
0						
1						
2						
3						
4						
5						
Total						

3. Tableau et représentation graphique

a). Cas d'un caractère qualitatif et quantitatif discret

- **Solution exemple 2:** On dispose d'un échantillon de 20 familles, interrogées sur le nombre d'enfants. La série obtenue est la suivante : 4 0 1 1 2 2 2 3 3 4 2 3 3 5 2 1 3 3 4 5

x_i	n_i	f_i (en %)	ECC_i (en %)	ECD_i (en %)	FCC_i (en %)	FCD_i (en %)
0	1	05	1	20	5	100
1	3	15	4	19	20	95
2	5	25	9	16	45	80
3	6	30	15	11	75	55
4	3	15	18	5	90	25
5	2	10	20	2	100	10
Total	20	100	Au plus	Au moins	Au plus	Au moins

Il y a 19 familles (95 %) qui ont **au moins** 1 enfant,

Il y a 9 familles (45 %) qui ont **au plus** 2 enfants

3. Tableau et représentation graphique

Application sur caractère qualitatif nominal

Une entreprise de fabrication emploie 50 salariés dans différents postes :

- Ouvrier (1),
- Ingénieur (2),
- Technicien (3),
- Manager (4),
- Directeur (5).

Les données recueillies sur les 50 salariés sont les suivantes :

1, 2, 3, 4, 2, 4, 1, 1, 2, 4, 1, 2, 1, 1, 4, 2, 4, 1, 3, 1, 3, 2, 2, 4, 5, 3, 1, 3, 4, 2, 1, 2, 1, 3, 4, 2, 1, 4, 3, 1, 2, 4, 1, 3, 2, 4, 1, 3, 2, 4

- 1) Créez un tableau des effectifs pour chaque poste.
- 2) Calculez la fréquence de chaque modalité et exprimez-la en pourcentage dans le même tableau,

3. Tableau et représentation graphique

Application sur caractère qualitatif nominal

Solution :

x_i	n_i	f_i (en %)
1	15	30
2	13	26
3	9	18
4	12	24
5	1	2
Total	50	100

3. Tableau et représentation graphique

b). Cas d'un caractère quantitatif continu

Les modalités sont regroupées en classes. Soit k le nombre de ces classes

Modalités $[x_i, x_{i+1}[$	Centre $c_i = \frac{x_i + x_{i+1}}{2}$	n_i	f_i	F_i	Amplitude $a_i = x_{i+1} - x_i$
$[x_1, x_2[$	c_1	n_1	f_1	$F_1 = f_1$	a_1
$[x_2, x_3[$	c_2	n_2	f_2	$F_2 = F_1 + f_2$	a_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$[x_i, x_{i+1}[$	c_i	n_i	f_i	$F_i = \sum_{j=1}^i f_j$ $F_i = F_{i-1} + f_i$	a_i
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$[x_{k-1}, x_k[$	c_k	n_k	f_k	$F_k = 1$	a_k
Total		N	1		

3. Tableau et représentation graphique

b). Cas d'un caractère quantitatif continu

Exemple : Les 13 étudiants d'une Licence ont passé un examen. La répartition des notes obtenues par ces étudiants est donnée par : 6 – 8 – 11 – 10 – 9 – 11,5 - 10,25 - 9,5 - 8,5 – 7 – 5 – 14 – 15,25

Note $[x_i, x_{i+1}[$	Centre c_i	Effectif n_i	Fréquence f_i	Fréquence cumulé F_i	Amplitude a_i
[0,4[
[4,8[
[8,12[
[12,16[
[16,20[
Total					

3. Tableau et représentation graphique

b). Cas d'un caractère quantitatif continu

Solution :

Note $[x_i, x_{i+1}[$	Centre c_i	Effectif n_i	Fréquence f_i (%)	Fréquence cumulé F_i (%)	Amplitude a_i
[0,4[2	0	0	0	4
[4,8[6	3	23,07	23,07	4
[8,12[10	8	61,53	84,6	4
[12,16[14	2	15,38	99,98	4
[16,20[18	0	0	99,98	4
Total		13	99,98	Au plus	

$n_3 = 8$ le nombre d'étudiants ayant des notes entre 8 et 12

$f_3 = 61,53\%$ pourcentage des étudiants ayant des notes entre 8 et 12

$F_3 = 84,6\%$ des étudiants ayant au plus une note de 12

3. Tableau et représentation graphique

b). Cas d'un caractère quantitatif continu

Application : La série suivante représente le poids réel d'un échantillon de 42 boîtes de céréales de marques différentes:

300-250-242-320-410-420-500-512-600-630-710-350-200-210-250-310-320-400-460-380-
510-560-740-240-270-280-340-360-230-290-310-430-460-530-640-680-720-450-470-640-
780-440

Il est possible de représenter des tableaux avec des classes d'amplitude égale ou inégale.

3. Tableau et représentation graphique

b). Cas d'un caractère quantitatif continu

Solution : Tableau statistique avec des classes d'amplitudes égales:

Classes de poids	Centre C_i	Effectifs (ni)	Fréquence (fi)	Amplitude a_i
[200 – 300[250	10	0,24	100
[300 – 400[350	9	0,21	100
[400 – 500[450	9	0,21	100
[500 – 600[550	5	0,12	100
[600 – 700[650	5	0,12	100
[700 – 800[750	4	0,10	100
Total		N=42	1	

3. Tableau et représentation graphique

b). Cas d'un caractère quantitatif continu

Solution : Tableau statistique avec des classes d'amplitudes inégales:

Classes de poids	Centre C_i	Effectifs (n_i)	Fréquence (f_i)	Amplitude a_i
[200 – 250[225	5	0,12	50
[250 – 350[300	11	0,26	100
[350 – 500[425	12	0,29	150
[500 – 600[550	5	0,12	100
[600 – 800[700	9	0,21	200
Total		N=42	1	

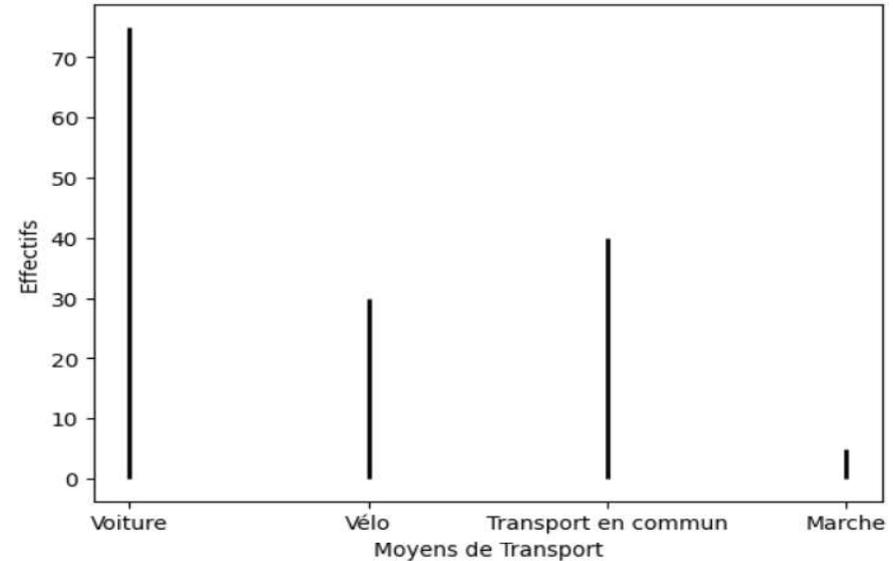
4. Représentation graphique

a). Diagramme en bâtons. Cas d'un caractère qualitatif et quantitatif discret

Application : On s'intéresse à la "répartition des moyens de transport" utilisés par une population de 150 individus. Voici le tableau des résultats obtenus :

Représenter graphiquement l'effectif et la fréquence de chaque moyen de transport dans cette population.

Moyen de transport	Effectif n_i
Voiture	75
Vélo	30
Transport en commun	40
Marche	5



4. Représentation graphique

b). Diagramme en secteur : (Caractère qualitatif et quantitatif discret)

Les diagrammes circulaires, ou semi-circulaires, sont des représentations où un disque ou un demi-disque est divisé en tranches, ou secteurs, correspondant aux modalités observées. La surface de chaque tranche est proportionnelle à l'effectif ou à la fréquence de chaque modalité. La circonférence totale (360°) représente l'effectif global.

Exemple : Tableau de distribution de la variable "Mention au Bac" pour un groupe de 15 étudiants.

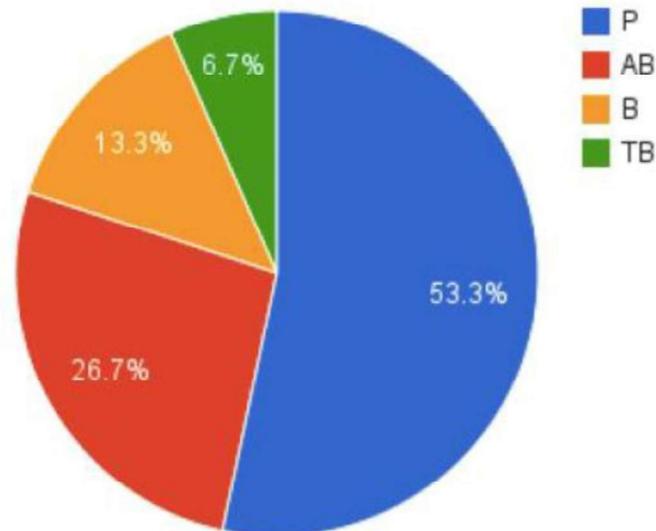
Mention au Bac	Effectifs	Fréquences
P	$n_P = 8$	$f_P = 8/15 = 0.533$
AB	$n_{AB} = 4$	$f_{AB} = 4/15 = 0.267$
B	$n_B = 2$	$f_B = 2/15 = 0.133$
TB	$n_{TB} = 1$	$f_{TB} = 1/15 = 0.067$
	effectif total $N = 15$	$f_P + f_{AB} + f_B + f_{TB} = 1$

4. Représentation graphique

b). Diagramme en secteur : Cas d'un caractère qualitatif et quantitatif discret

Solution de l'exemple :

Diagramme en secteurs des mentions



Les secteurs angulaires seront obtenus en multipliant les fréquences par 360 (360 représentant l'effectif total).
 $angle(M_i) = f_i \times 360$, où f_i est compris entre 0 et 1.

4. Représentation graphique

c). Courbe cumulative

Courbe cumulative est une courbe « en escalier » dont les paliers sont horizontaux,

Propriétés de la courbe cumulative :

La fonction F est définie sur \mathbb{R} et à valeurs dans $[0,1]$:

(i) $F(x) = 0$ si $x < x_1$

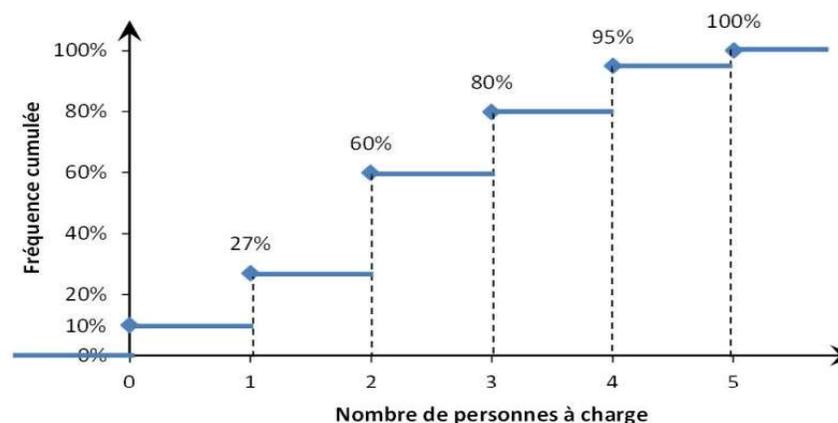
(ii) $F(x) = 1$ si $x \geq x_k$

(iii) $F(-\infty) = 0$ et $F(+\infty) = 1$

(iv) F est constante dans chaque intervalle séparant deux modalités $[x_i, x_{i+1}[$.

Exemple : Un assistant social d'un Centre Public d'Aide Sociale a étudié le nombre de personnes à charge parmi les bénéficiaires demandant une assistance. Pour cela, il a sélectionné aléatoirement 100 dossiers et a relevé, dans chacun d'eux, le nombre de personnes à charge déclaré,

Nbre pers. à charge x_{oj}	Eff. n_j	Eff. cumulé N_j	Fréq. f_j	Fréq. cumulée F_j
0	10	10	0.10	0.10
1	17	27	0.17	0.27
2	33	60	0.33	0.60
3	20	80	0.20	0.80
4	15	95	0.15	0.95
5	5	100	0.05	1
Total	$n = 100$		1	



4. Représentation graphique

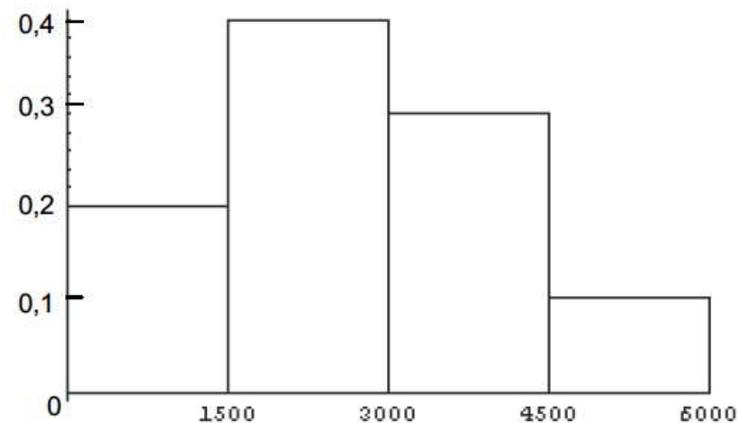
d). Histogramme: (Caractère quantitatif continue)

L'histogramme est particulièrement adapté pour représenter les variables quantitatives regroupées en classes. Dans le cas où les classes ont des amplitudes égales, sa construction est simple et directe. Toutefois, lorsque les amplitudes des classes diffèrent, une correction doit être appliquée en suivant la méthode décrite ci-dessous.

- **Cas des amplitudes égales**

Exemple : Considérons 100 ménages classés en fonction de leur revenu mensuel en euros, répartis dans des classes d'amplitudes égales à 1 500 euros.

Classe de revenu	n_i	f_i
$[0;1500[$	20	0,2
$[1500;3000[$	40	0,4
$[3000;4500[$	30	0,3
$[4500;6000[$	10	0,1



Les surfaces de rectangles S_i est proportionnelle à la fréquence f_i de la classe i .

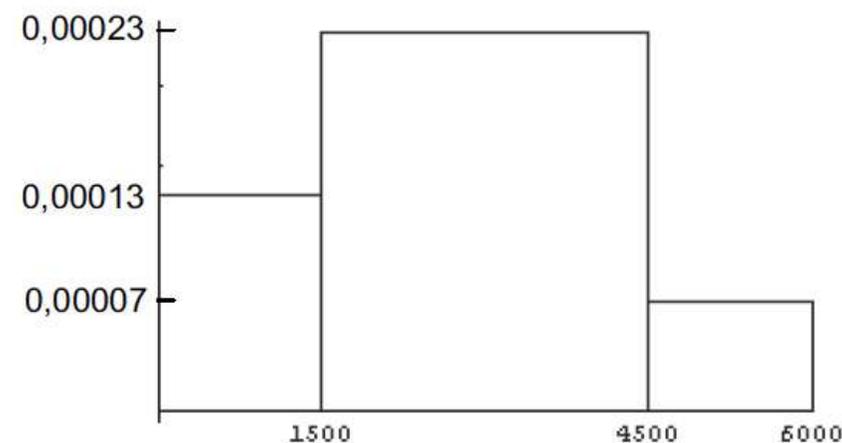
4. Représentation graphique

d). Histogramme: (Caractère quantitatif continue)

- **Cas des amplitudes inégales**

Supposons que les données de l'exemple précédent soient regroupées en classes d'amplitudes inégales ($[0,1500]$, $[1500,4500]$, $[4500,6000]$). Dans ce cas, une correction est nécessaire pour prendre en compte les différences d'amplitudes. Cette correction consiste à diviser la fréquence de chaque classe par l'amplitude correspondante. On obtient ainsi une fréquence corrigée, notée h_i .

Classe de revenu	Amplitude de classe (a_i)	n_i	f_i	$h_i = f_i/a_i$
$[0;1500[$	1500	20	0,2	0,00013
$[1500;4500[$	3000	70	0,7	0,00023
$[4500;6000[$	1500	10	0,1	0,00007

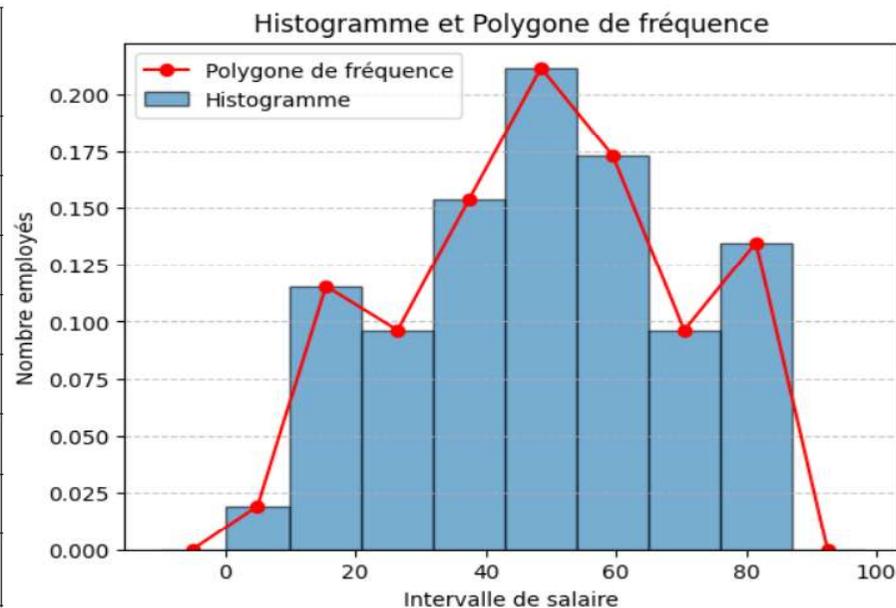


4. Représentation graphique

e). Polygone de fréquences

Le polygone des fréquences est tracé en reliant les milieux des bords supérieurs des rectangles de l'histogramme. Il s'agit d'une ligne brisée dont les points de départ et d'arrivée se trouvent sur l'axe des abscisses.

Salaire (milliers)	Nombre d'employés
[0 – 10[50
[11 – 21[300
[22 – 32[250
[33 – 43[400
[44 – 54[550
[55 – 65[450
[66 – 76[250
[66 – 76[350



4. Représentation graphique

e). Polygone de fréquences

Application : Le tableau suivant présente la répartition du personnel d'une entreprise en fonction de leur salaire mensuel en dirhams (DH).

Représenter graphiquement l'histogramme et le polygone de fréquence.

Salaire mensuel en dhs	Effectif n_i
[1500 – 2000[26
[2000 – 2500[12
[2500 – 3000[11
[3000 – 3500[6
[3500 – 4000[5
Total	60