

Statistique descriptive:

1 - Généralités:

La statistique a pour objet de recueillir des observations portant sur des sujets présentant une certaine propriété et de traduire ces observations par des nombres qui permettent d'avoir des renseignements sur cette propriété.

Le but: de la statistique descriptive est de structurer et de représenter l'information contenue dans les données.

Vocabulaire Statistique:

- 1 - **population:** est l'ensemble des sujets observés.
- 2 - **le caractère:** est la propriété étudiée sur ces sujets.
- 3 - **Modalités:** sont les diverses cas prise par le caractère.
- 4 - **Unité Statistique:** chaque élément de la population appelé unité statistique.

Expt: l'étude du nombre d'enfant à charge dans les 26 familles de l'Avenue Aouati Mostefa.

①

Exp 2: l'étude de l'état matrimonial des Travailleurs de la société S.N.C.

Exp 3: l'étude de salaire annuel des enseignants de l'U.N.C

Solution 1:

- la population : 26 familles.
- Unité Statistique: chaque famille.
- caractère: nombre d'enfant.
- Modalité: 0, 1, 2, 3, ...

Solution 2:

- population: les Travailleurs.
- caractère: l'état matrimonial.
- Modalité: Marié, divorcé, célibataire, veuf.

Solution 3:

- population: les enseignants.
- caractère: Les salaires annuel.
- Modalité: [2500, 30000[, ...

Les Types de caractères

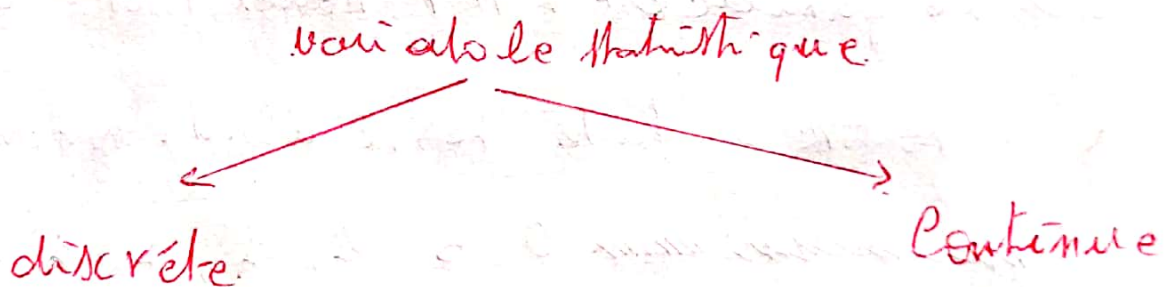
1. caractère quantitatif

un caractère est dit quantitatif quand ces différentes modalités sont mesurables par des nombres comme Ex 1 et 3

2. caractère qualitatif

un caractère est dit qualitatif quand ces différentes modalités ne peuvent être désignées que par leur qualité, comme Ex 2.

- Variable Statistique = le caractère quantitatif est appelé variable statistique.



la variable statistique est dite discrète lorsque ces modalités sont des nombres fixes comme Ex 1.

la variable statistique est dite continue quand elle peut prendre n'importe quelle valeur dans un intervalle. Comme Ex 3.

Chapitre 02: Variable Statistique discrète.

Définition: On appelle variable statistique discrète ou caractère quantitatif discrète toute Application:

$$X: P \longrightarrow \mathbb{R}$$

$$\omega \longrightarrow X(\omega) = x_i.$$

$$X(P) = \{x_1, x_2, \dots, x_R\} \quad x_1 < x_2 < \dots < x_R$$

Exp: En 2004 on a enregistré le nombre d'enfant des 26 famille de L'A.V. Aouati Mostefa, on note

X : le nombre d'enfants. Les résultats sont données:

~~procentage~~ 0, 5, 1, 1, 1, 1, 1, 0, 2

2, 3, 1, 2, 3, 4, 1, 1, 3, 2, 2, 1, 1, 0, 0, 2, 3

1: Fréquence

Fréquences \rightarrow absolues n_i : donnent le nombre d'unités en valeur absolues.

Relatives f_i : $f_i = \frac{n_i}{N}$
 $N \rightarrow$ effectif total.

(4)

- effectifs cumulés en n_R : $N_{iR} = \sum_{i=1}^k n_i$

- Fréquences relatives cumulées en n_R : $F_R = \sum_{i=1}^k f_i$

le Tableau Statistique de notre Exp:

n_i	n_i	f_i	N_{iR}	F_R
0	4	0,153	4	0,153
1	10	0,384	14	0,537
2	6	0,230	20	0,767
3	4	0,153	24	0,919
4	1	0,0384	25	0,9574
5	1	0,0384	26	1
Σ	$N=26$	$\hat{=} 1$		

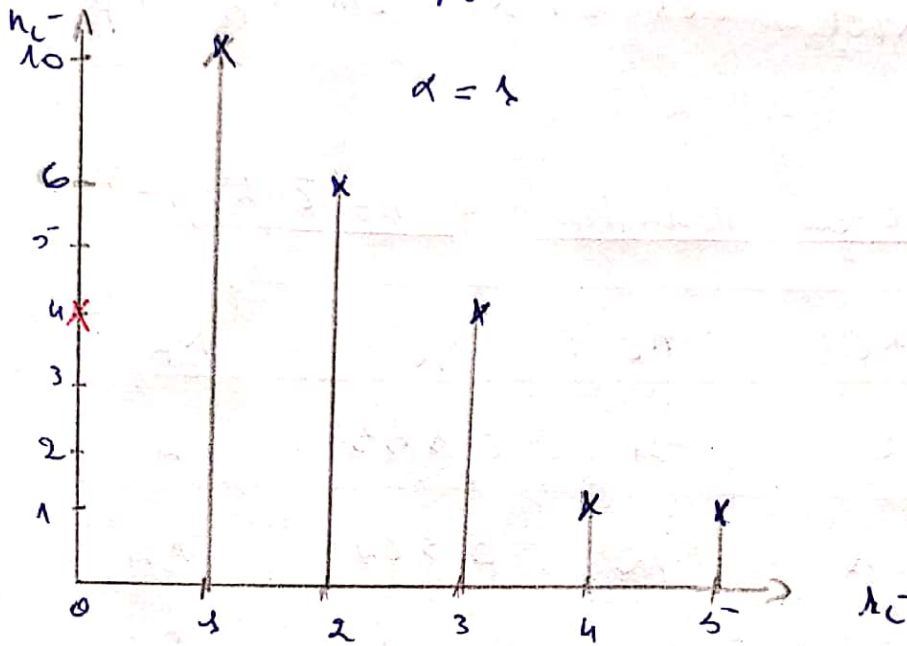
Représentation graphique

pour un caractère quantitatif discret on a deux types de représentation graphiques sont cyclébre : le diagramme différentiel et le diagramme cumulatif.

1- diagramme différentiel est appelé aussi diagramme en bâtons dans un repère orthogonal en association à

(5)

chaque valeur de la variable n_i un segment de droite, de longueur αn_i ou αf_i ($\alpha \in \mathbb{R}_+^*$).



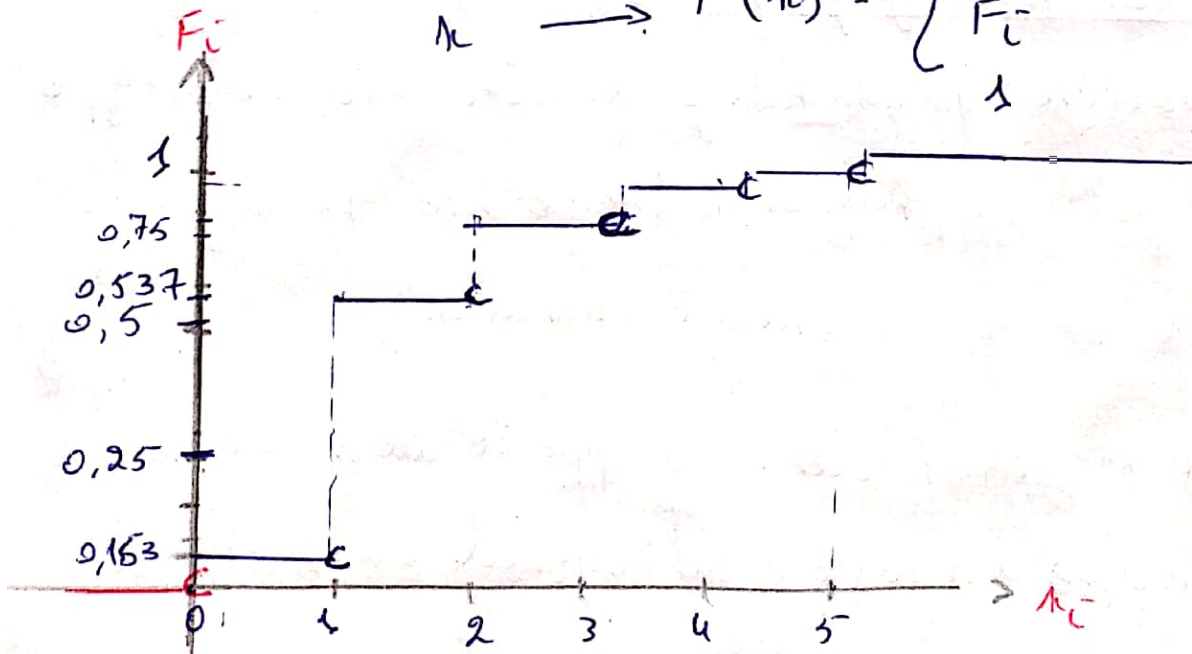
le diagramme en bâtons.

2 - le diagramme cumulatif (intégrale):

La Courbe Cumulative est la représentation graphique de la fonction Cumulative F définie par:

$$F: \mathbb{R} \rightarrow [0, 1]$$

$$n \rightarrow F(n) = \begin{cases} 0 & \text{si } n < n_1 \\ F_i & n_i < n < n_{i+1} \\ 1 & n \geq n_K \end{cases}$$



L'escalade théorique de Tendance Centrale.

1. **Médiane M** : est la valeur de la variable statistique qui partage la population en 2: $F(M) = 1/2$.

Rq: pour déterminer M en 2 façons à le faire:

1: à partir de la colonne des n_i en encadre $\frac{N}{2}$ est la valeur correspondante des n_i est M .

2: à partir de la colonne des F_i en encadre 1/2. la valeur correspondante des n_i est M .

dans notre Exp: $M = 1$.

2. **Les quartiles**: chacune des trois valeurs Q_1, Q_2, Q_3 qui divisent les éléments d'une distribution statistique en quatre parties d'effectif égale.

1^o quartile Q_1 : $F(Q_1) = 0,25$.

<	>
25%	75%

2^o quartile Q_2 : $F(Q_2) = 0,5$.

<	>
50%	50%

$$Q_2 = M.$$

3^o quartile Q_3 : $F(Q_3) = 0,75$.

75%	25%
-----	-----

Rq: d'une manière analogue à celle de M entre Q_1 et Q_3

(7)

dans notre Exp:

Donc

$$0,153 < 0,25 < 0,537.$$

$$Q_1 = 1$$

$$0,537 < 0,75 < 0,767$$

$$Q_3 = 2.$$

3. le Mode M_0 : Le Mode M_0 est la valeur de la variable statistique qui correspond au plus grand effectif. dans notre Exp $M_0 = 1$.

Rq: le Mode n'est pas unique.

4. la Moyenne arithmétique:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^K n_i \cdot m_i = \sum_{i=1}^K \frac{n_i}{N} m_i = \sum_{i=1}^K f_i m_i$$

ou

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N m_i$$

Exp: Le nombre moyenne d'enfants dans les 26 familles

$$\text{est: } \bar{x} = \frac{1}{26} \sum_{i=1}^{26} n_i m_i = \frac{43}{26} = 1,65$$

\approx 2 enfants par famille.

⑧

Un caractère qualitatif:

Soit une population N comprenant N individus pour chacun desquels on a fait une observation. Concernant un caractère K .

Celui-ci comporte les modalités $M_1, \dots, M_i, \dots, M_K$ pour procéder au classement des unités statistiques.

Suivant ces diverses modalités il est donc nécessaire de donner les définitions suivantes: Effectif ou fréquence absolue, fréquence relative.

Série statistique: Les couples $(M_i, n_i) = 1 \dots K$ forment une suite appelée la série statistique.

$(M_1, n_1), \dots, (M_K, n_K)$.

Tableau statistique: Le Tableau statistique décrivant une population N suivant un caractère qualitatif K est de la forme suivante:

Modalités	Effectifs	fréquences
M_1	n_1	f_1
\vdots	\vdots	\vdots
M_i	n_i	f_i
\vdots	\vdots	\vdots
M_K	n_K	f_K
Total	N	1.

Les Modalités sont classées dans l'ordre non croissant des fréquences.

Exemple: l'étude matrimoniale de 40 employés de la société N. G. a conduit au résultat suivant:

Tableau Matériel:

Modalités	effectifs	fréquences
Marié	20	0,5
Célibataire	12	0,3
Divorcé	6	0,15
Veuf	2	0,05
Total	40	1.

Le Tableau Matériel donne une synthèse des données Matérielles mais les graphes l'exprime d'une façon plus claire.

Représentation graphique:

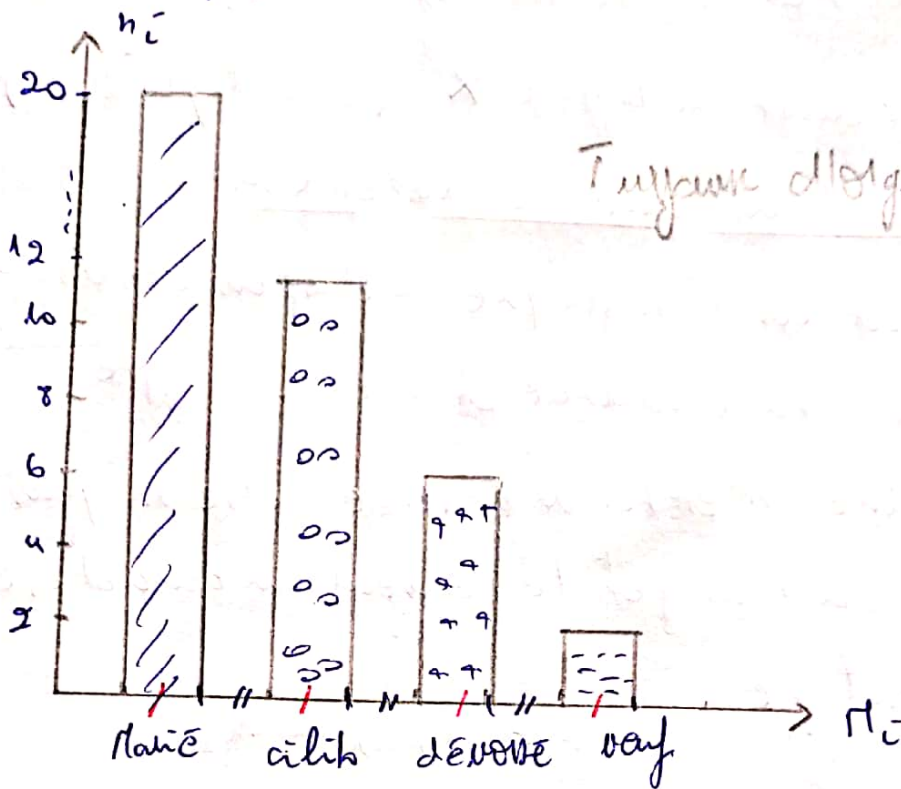
pour un caractère qualitatif on a 2 types de représentation graphique. sont célèbres: les Diagrammes d'orgue et les secteurs angulaires.

1: Diagramme d'orgue: à chaque modalité M_i correspond un rectangle dont la hauteur est égale à sa fréquence relative

(10)

f_i au son Effectif n_i . Les Rectangles ont des largeurs égales et sont séparés de a deux fois des distances égales.

Exp:

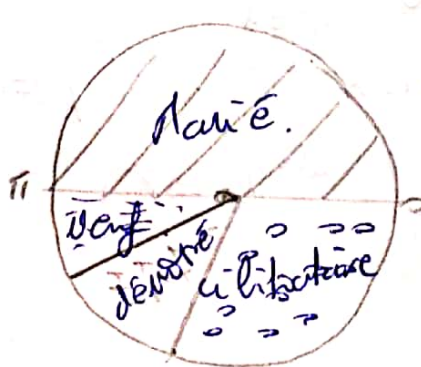


z : secteurs angulaires: chaque Modalité M_i est associée à un angle θ_i proportionnel à la fréquence relative f_i

c - a - d $\theta_i = f_i \times 360^\circ$

Exp:

M_i	Marié	célibataire	divorcé	veuf
θ_i	180°	108°	54°	18°



secteur angulaire.

(11)

Les caractéristiques de tendance centrale sont:

le médian m , Les quartiles Q_1, Q_2, Q_3 , le mode M_0 ,
la moyenne arithmétique \bar{x} (on a déjà vu).

Les caractéristiques de dispersion:

Les indicateurs statistiques de dispersion usuels sont:
l'étendue, la variance et l'écart-type.

1- l'étendue: c'est la différence entre la plus grande
valeur et la plus petite valeur du caractère donnée
par la quantité

$$E = M_{\max} - M_{\min}$$

l'étendue permet de donner une première idée de la
dispersion des observations.

2- la variance: on appelle variance de cette série
statistique X le nombre:

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

ou bien:
$$= \frac{1}{N} \sum_{i=1}^K n_i m_i^2 - \bar{x}^2$$

3- l'écart-type: la quantité: $\sigma_x = \sqrt{\text{Var}(X)}$

Il permet de mesurer la dispersion d'une série statistique
autour de la moyenne.