Chapitre IV : Statistique à une variable

I - Définitions generales

1) La statistique a pour but de collectionner des chiffres, des données, d'étudier des faits pour permettre de prendre des decisions.

2) L'ensemble sur lequel le statisticien fait des observation est appelé population statistique. Chaque élt de cette population est appelé individu statistique ou unité statistique.

3) En statistique, on etudie certaines propriétés des unités statistique que l'on appelle caractere statistique. Il existe 2 types de caractère : le caractère quantitatif et le caractere qualitatif.

4) Un caractère est dire qualitatif s'il n'est ni mesurable ni reperable (nationalité, situation matrimoniale, fidelité ...)

5) Par contre, un caractere est quantitatif s'il est mesurable ou reperable. Il existe 2 sous type de caractère quantitatif :

* le caractere est dit quantitatif discret s'il est denombrable (où il prend les valeurs de $\mathbb{N}$) Exple : nbre d'enfants dans une famille, le nbre de vehicules dans un parc automobile.

Scanné avec CamScanner

* Le caractère est dit quantitatif continue s'il peut prendre toutes les valeurs d'un intervalle de l'ensemble $\mathbb{R}$. (Poids d'un enfant à la naissance, intérêt produit par un dépôt à la banque...)

6) On appelle modalité d'une variable statistique l'ensemble des ≠ états (Dans le cas qualitatif) ou des ≠tes valeurs (Dans le cas quantitatif) que peut prendre la variable statistique. Expl: Pour la profession, il existe plusieurs modalités: ouvriers, maçons, avocats, enseignants...

7) On appelle effectif n_i (fréquence f_i) d'une variable statistique, le nbre (la proportion) associé à cette variable. On a alors $f_i = \frac{n_i}{N}$ ($\times 100$) en %.

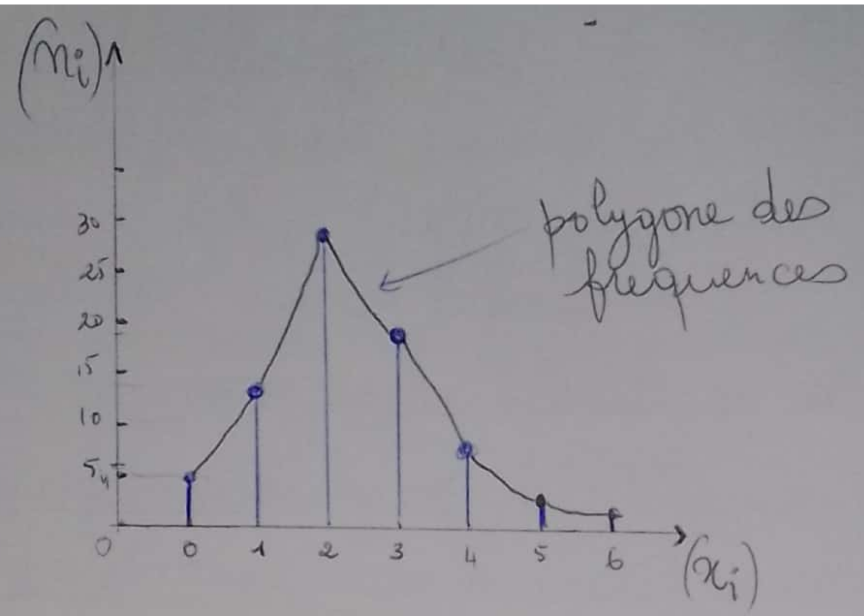
II - Représentations graphiques

1 - Diagramme à bâtons (cas discret)

On place en abscisse les valeurs de la variable statistique et en ordonnée les effectifs ou fréquences correspondants

Expl 1: Le tableau statistique suit donne la distribution du personnel d'une entreprise selon le nbre d'enfants en charge.

nbre d'enf. à charge (x_i)	0	1	2	3	4	5	6
eff (n_i)	4	15	29	18	10	3	1

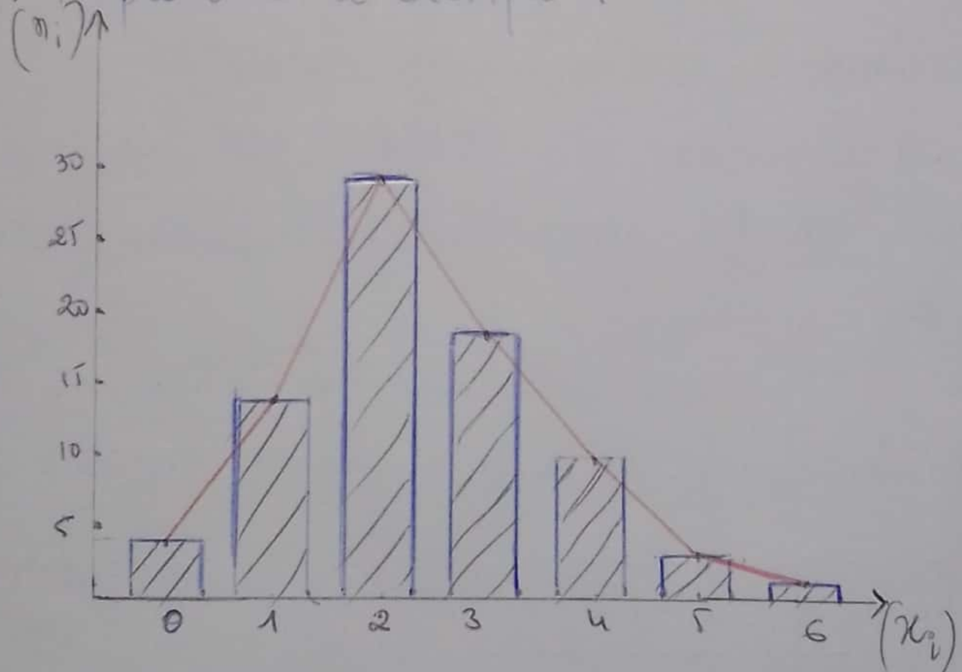


En reliant les extrémités supérieures des batons, on obtient le polygone des frequences.

2- Diagramme à bandes (cas discret)

On represente des bandes rectangulaires dont la hauteur est proportionnelle aux effectifs ou aux frequences et les bases sont aleatoires de même largeur.

Exple 2: Reprenons l'exemple 1



3_ Histogramme (cas continu)

C'est une suite de rectangles dont les bases coincident avec les classes dont les hauteurs sont proportionnelles aux effectifs ou aux frequences relatives.

On distingue 2 types d'histogramme : l'histogramme à amplitude égale et l'histogramme à amplitude inégale.

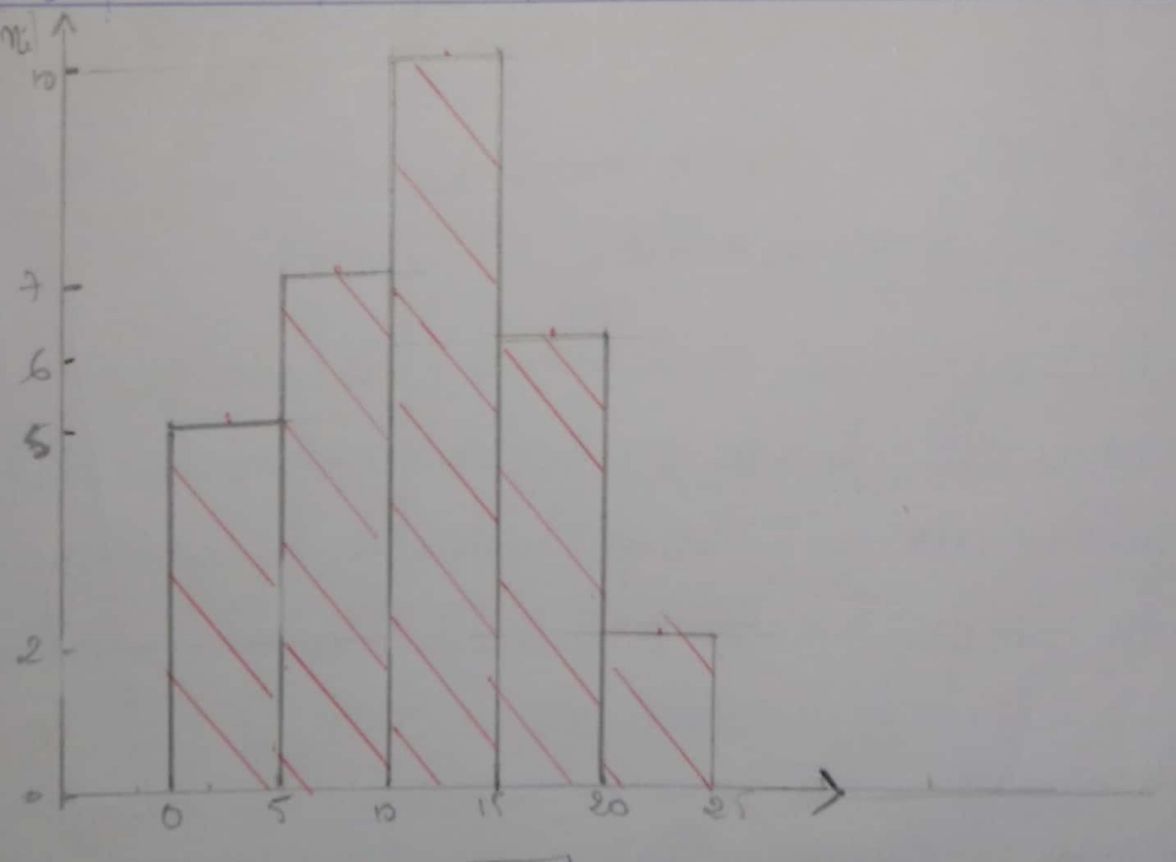
Si on a : $[b_i ; c_i[\Rightarrow$ amplitude $a_i = c_i - b_i$.

On appelle densité d'une classe, le reel $d_i = \frac{n_i}{a_i}$

ⓐ Histogramme à ampl. egales

Exemple : Soit la distribution statistique suivante :

Classes	[0 ; 5[	[5 ; 10[	[10 ; 15[	[15 ; 20[	[20 ; 25[
eff (n_i)	5	7	10	6	2

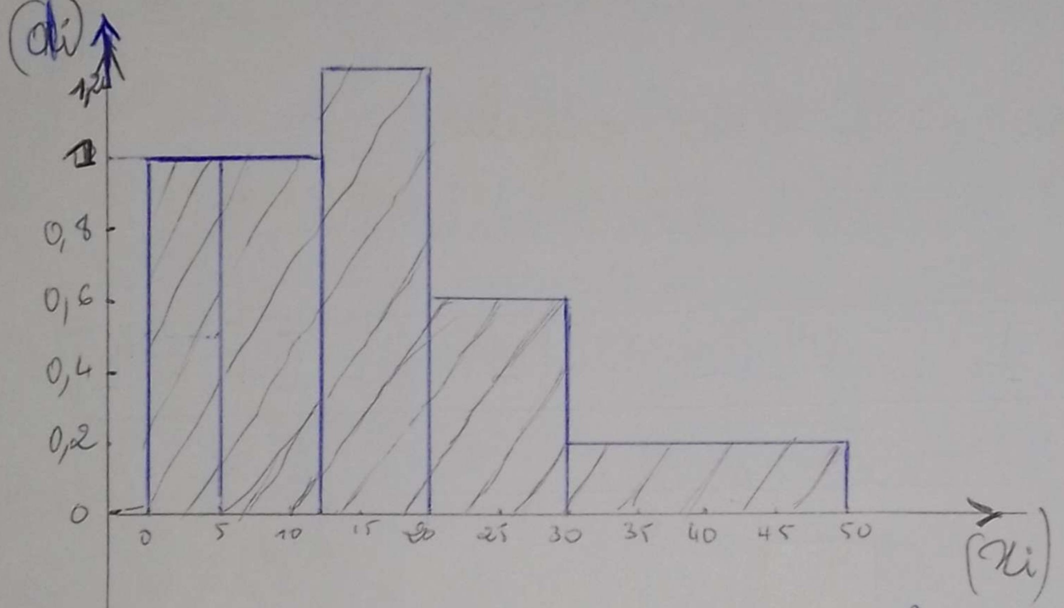


NB: On appelle classe modale de cette distribution, la classe qui a le plus gros effectif

(b) Histogramme à ampl. inégales

Exemple

Classes	[0;5[	[5;12[	[12;20[	[20;30[	[30;50[
eff (n_i)	5	7	10	6	2
densité (d_i)	1,00	1,00	1,25	0,60	0,10



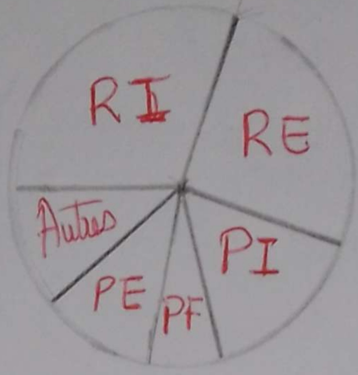
Dans ce cas, la classe modale est la classe qui a la plus grande densité.

4 – Diagramme à secteurs (circulaire)

C'est un cercle divisé en secteurs dont la surface de chacun d'eux est proportionnelle à la fréquence de chacune des modalités étudiées. L'angle réservé à chaque secteur est défini par $\alpha_i = 360° \cdot f_i$ (f_i = freq relative)

Exemple: Le tableau suit donne la répartition des ménages sénégalais sur la source d'approvisionnement en eau (Source ANSD 1985

Source	Rob Int (RI)	Rob Ext (RE)	Puits Int (PI)	Pompes Forages (PF)	Puits Ext PE	Autres
freq(%)	30,1	25,7	15,3	7	10,8	11,1
α_i(°)	108,36	92,52	55,08	25,2	38,88	39,96



5) Tableau recapitulatif

Le tableau statistique suivant donne la repartition des salaires horaires d'une entreprise en centaines de francs.

Salaire	eff(ni)	freq(fi)	C4 ECC	C5 ECD	C6 FCC	C7 FCD
[47,5 ; 52,5[	10	0,040	10	250	0,040	1
[52,5 ; 57,5[	30	0,120	40	240	0,160	0,96
[57,5 ; 60,5[	60	0,240	100	210	0,400	0,84
[60,5 ; 63,5[	72	0,288	172	150	0,688	0,6
[63,5 ; 67,5[	40	0,160	212	78	0,848	0,312
[67,5 ; 73,5[	24	0,096	236	38	0,944	0,152
[73,5 ; 80,5[	14	0,056	250	14	1	0,056
Totaux	$\Sigma n_i = 250$	1		0		0

Lecture du tableau

① Colonne 4

* 10 employés gagnent moins de 5250F à l'heure
* 172 " 6350F à " .

Colonne 6

* 4% des employés . . 5250F " "
* 84,8% " " " " . 6350F à l'heure

Colonne 5

* 240 des employés gagnent 5250^F par heure ou plus
* 38 " " " 6750^F " " " " "

Colonne 7

* 100 % " " 4750^F " " ou plus
* 15,2% 6750^F " " "

III – Caractéristiques de

1 – Moyennes

ⓐ Moyenne arithmétique

C'est la somme de toutes les données statistiques divisée par le nbre de données.

Soit la série statistique $x_i = (x_1, x_2, \dots x_n)$. La moyenne arithmétique simple de cette série est notée:

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Si les valeurs x_i sont pondérées par les effectifs n_i alors on aura (*) $\overline{X} = \frac{1}{N}\sum_{i=1}^{n} n_i x_i$ où $N = n_1 + \dots + n_n$

NB: Lorsqu'on a une variable statistique continue sur des classes $[a_i; b_i[$, le x_i dans la formule (*) est remplacée par le centre des classes $\frac{a_i + b_i}{2}$. $\overline{X} = \frac{1}{N}\sum_{i=1}^{n} n_i C_i$

Exple 1: Lors de l'examen du 1er semestre un étudiant a obtenu les notes sotes.

Matières	Math	Proba	Algo	SGBD	LC	Anglais
Notes	09,5	12	07	13	12	15
Coef	4	3	4	2	3	1

$$\overline{X} = \frac{1}{17}\left[(9{,}5\times4) + (12\times3) + (7\times4) + (13\times2) + (12\times3) + (15\times1)\right] = 10{,}52$$

Exple 2: Calculer la moyenne arithmetique

$S = \{-2; 8; 10; -4; 7; 0; 3\}$

$$\bar{x} = \frac{(-2+8+10-4+7+0+3)}{7} = 3,14$$

Exple 3. Soit la série statistique suivante qui donne le salaire des ouviers de PME SSS en milliers de frans.

Determiner la salaire moyen dans ce PME

Classes	eff (n_i)	Centre (x_i)	$n_i x_i$
[45; 50[	15	47,5	712,5
[50; 74,5[	15	62,25	933,75
[75,5; 100[	20	87,25	1745
[100; 220[	8	160	1280
[220; 300[	4	260	1040
[300; 400[	3	350	1050
Totaux	65		6761,25

$$\bar{x} = \frac{1}{N} \sum n_i x_i$$

$$= \frac{1}{65}(6761,25)$$

$$= 104,01 \times 1000$$

$$\boxed{\bar{x} = 104010\ F}$$

(b) Moyenne geometrique

On appelle moyenne statistique de la serie statistique $X = \{x_1, \dots, x_n\}$ la racine $n^{ième}$ du produit des valeurs observées.

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n} = (x_1 \cdot x_2 \cdot \ldots \cdot x_n)^{1/n}$$

Si les valeurs x_i sont ponderées par des coefs n_i, on obtient:

$$G = \sqrt[N]{x_1^{n_1} \cdot x_2^{n_2} \cdot \ldots \cdot x_n^{n_n}} \quad \text{ou } N = n_1 + n_2 + \ldots n_n$$

$$\ln G = \ln \left(x_1^{n_1} \cdot x_2^{n_2} \cdot \ldots \cdot x_n^{n_n}\right)^{\frac{1}{N}} = \frac{1}{N}\left[n_1 \ln x_1 + \ldots + n_n \ln x_n\right]$$

$$= \alpha \Rightarrow G = e^{\alpha}$$

Exple 1 : $X = \{2 ; 5 ; 1 ; 11\} \Rightarrow G = \sqrt[4]{2 \times 5 \times 1 \times 11} = \sqrt[4]{110} = (110)^{1/4}$

$G = 3,238$

$\bar{X} = \frac{1}{4}(2+5+1+11) = \frac{19}{4} = 4,75$

Exple 2 : Un particulier place un capital C_0 à un taux de 5,28% pdt 5 ans, puis au taux 6,10% pdt 3 ans et enfin au taux de 4,90% pdt 2 ans. Calculer le taux moyen de placement pdt les 10 ans.

- à la fin de la 1ère année

$C_1 = C_0 + 5,28\% C_0 = C_0(1 + 0,0528) = 1,0528 C_0$

- à la fin de la 5e année

$C_5 = C_4 + 5,28\% C_4 = (1,0528)^5 C_0$

- à la fin de la 6e année

$C_6 = C_5 + 6,1\% C_5 = 1,061 C_5 = 1,061 (1,0528)^5 C_0$

- $C_7 = C_6 + 6,1\% C_6 = (1,061)^2 (1,0528)^5 C_0$

- $C_8 = (1,061)^3 (1,0528)^5 C_0$

- $C_{10} = (1,049)^2 (1,061)^3 (1,0528)^5 C_0$

$(1+t)^{10} = (1,0528)^5 . (1,061)^3 . (1,049)^2$

$t = \sqrt[10]{(1,0528)^5 . (1,061)^3 . (1,049)^2} - 1$

c) Moyenne harmonique

L'inverse de la moyenne arithmétique des inverses des valeurs observées : $H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$

Si les valeurs x_i sont pondérées par les coefs n_i, on aura

$$H = \frac{N}{\sum \frac{n_i}{x_i}} \quad \text{avec } N = n_1 + n_2 + \dots + n_n$$

Elle est généralement utilisée dans le calcul des vitesses, du temps dans le domaine électronique, ...,

Exple : Lors du d

Un étudiant quitte son domicile à une vitesse $V_1 = 4,72$ km/h et y retourne chez lui a une vitesse $V_2 = 3,55$ km/h

Soit D la distance domicile-institut ; t_1 : tmp aller

t_2 : tmp retour

$$V = \frac{d}{t} = \frac{d+d}{t_1+t_2} = \frac{2d}{\frac{d}{V_1}+\frac{d}{V_2}} = \frac{2}{\frac{1}{V_1}+\frac{1}{V_2}} = \frac{2}{\frac{1}{4,72}+\frac{1}{3,55}} = 4,05 \text{ km/h}$$

NB : Il est des fois plus aisé de calculer au préalable $\frac{1}{H}$

$$\frac{1}{H} = \frac{1}{N}\sum_{i=1}^{n} \frac{n_i}{x_i} \Rightarrow \frac{1}{H} = \alpha \Rightarrow H = \frac{1}{\alpha}$$

Chapitre IV : Statistique à une variable

I - Définitions generales

1) La statistique a pour but de collectionner des chiffres, des données, d'étudier des faits pour permettre de prendre des decisions.

2) L'ensemble sur lequel le statisticien fait des observation est appelé population statistique. Chaque élt de cette population est appelé individu statistique ou unité statistique.

3) En statistique, on etudie certaines propriétés des unités statistique que l'on appelle caractere statistique. Il existe 2 types de caractère : le caractère quantitatif et le caractere qualitatif.

4) Un caractere est dire qualitatif s'il n'est ni mesurable ni repérable (nationalité, situation matrimoniale, fidelité ...)

5) Par contre, un caractere est quantitatif s'il est mesurable ou reperable. Il existe 2 sous type de caractère quantitatif :

* le caractere est dit quantitatif discret s'il est denombrable (où il prend les valeurs de $\mathbb{N}$) Exple : nbre d'enfants dans une famille, le nbre de vehicules dans un parc automobile.