Cours avec Exercices d'application **PROF : ATMANI NAJIB** Tronc CS
Avec solutions

STATISTIOUES

I. Population statistique /Caractère:

1) Etude d'exemples:

Regardons les exemples suivants:

Exemple :1 L'étude suivante donne les notes de 20 élèves :

$9-8-10-12-10-8-15-18-16-15-12-12-10-10-9-8-15-12-8-10$

$x_7=18$	$x_6=16$	$x_5=15$	$x_4=12$	$x_3=10$	$x_2=9$	$x_1=8$	Note
$n_7=1$	$n_6=1$	$n_5=3$	$n_4=4$	$n_3=5$	$n_2=2$	$n_1=4$	Effectif
20	**19**	**18**	**15**	**11**	**6**	**4**	Effectif cumulé

Exemple :2

Les vitesses de 150 voitures ont été détectée s sur l'autoroute entre Rabat et Casa, on a obtenu le tableau suivant :

[130,150[	[110,130[	[90,110[	[70,90[	[50,70[	Vitesse
15	**25**	**60**	***40***	**10**	Effectif
150	**135**	**110**	**50**	**10**	Effectif cumulé

[130,150[	[110,130[	[90,110[	[70,90[	[50,70[	Vitesse
15	**25**	**60**	***40***	**10**	Effectif
150	**135**	**110**	**50**	**10**	Effectif cumulé

Exemple :3

le tableau ci-dessous représente le nombre de buts par match durant la Coupe du monde de football de 2010 :

Nombre de buts x_i	0	1	2	3	4	5	6	7
Nombre de matchs n_i	7	17	13	14	8	6	0	1

Les valeurs x_i du caractère étudié sont les "nombres de buts".
Les effectifs n_i correspondants sont les "nombres de matchs".

2) Définitions :

Population statistiaque :

La population statistique est l'ensemble qui fait l'objet de l'étude. et chaque élément de cet ensemble est appelé « individu » ou « unité statistique »
Dans exemple1 la population statistique est l'ensemble des élèves.
Dans exemple2 la population statistique est l'ensemble des voitures.

Caractère : la propriété qu'on veut étudier " chez une population statistique s'appelle « le caractère » ou « la variable statistique ». " le caractère peut être quantitatif ou qualitatif.

Le caractère quantitatif est un caractère qui peut s'exprimer par des nombres

(Dans exemple1 c'est la note)

on distingue le caractère quantitatif discret et le caractère quantitatif continu.

Caractère discret :

Le caractère quantitatif discret est celui qui prend des valeurs isolées, comme les notes des élève (dans exemple1) ou le numéro du mois de naissance d'un élève par exemple.

Caractère continu :

Le caractère quantitatif continu est celui qui prend des valeurs très proches, dans ce cas les valeurs du caractère sont rassemblées dans des intervalles qu'on appelle aussi des « classes », comme les hauteurs des élèves par exemple.

2) Caractère quantitatif :

Le caractère qualitatif est un caractère qui ne peut pas s'exprimer par des nombres, comme la couleur du cheveu des élèves ou leur groupe sanguin ou leur sexe.

II. Effectif – fréquence– pourcentage

x_7	x_6	x_5	x_4	x_3	x_2	x_1	Note
n_7	n_6	n_5	n_4	n_3	n_2	n_1	Effectif

n_i est appelé « Effectif» relative à la valeur x_i

$N = n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7$ Le nombre N représente l'effectif total.

$f_i = \dfrac{n_i}{N}$ est appelé « fréquence » relative à la valeur x_i

$p_i = 100 f_i$ est appelé le pourcentage relatif à la valeur x_i

Dans exemple1

$N = n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7 = 20$

$p_1 = f_1 \times 100 = \dfrac{100}{5} = 20\%$ est appelé le pourcentage relatif à la valeur x_1

III. Paramètres de position

1) **le mode:** c'est la valeur du caractère ou la classe correspondant au plus fort effectif .

Remarque : pour déterminer le mode , il faut d'abord dresser le tableau des effectifs.

Dans exemple1: le mode est la note $x_4 = 12$

Dans exemple2: la classe modale est la classe $[90,110[$

Dans exemple3: le mode est Nombre de buts1

2) **Moyenne**

La moyenne $\bar{x}$ d'une série statistique dont les valeurs du caractère sont x_1, x_2, x_3, …, x_k et les effectifs correspondants sont n_1, n_2, n_3, …, n_k est égale à :

$$\bar{x} = \frac{x_1 \times n_1 + x_2 \times n_2 + \ldots + x_k \times n_k}{n_1 + n_2 + \ldots + n_k} = \frac{\sum_{i=1}^{k} x_i \times n_i}{\sum_{i=1}^{k} n_i}$$

Exemple :

Dans exemple1

La moyenne de NOTES est égale à :

$$\bar{x}=\frac{8\times4+9\times2+10\times5+12\times4+15\times3+16\times1+18\times1}{20}=\frac{227}{20}\approx11,35$$

Dans exemple3

La moyenne de buts par match est égale à :

$$\bar{x}=\frac{7\times0+17\times1+13\times2+14\times3+8\times4+6\times5+0\times6+1\times7}{7+17+13+14+8+6+1}=\frac{154}{66}\approx2,3$$

3) Médiane

Pour obtenir la médiane d'une série, on range les valeurs de la série dans l'ordre croissant. La médiane est la valeur qui partage la série en deux populations d'effectif égal.

Exemple : **Dans exemple3**

L'effectif total est égal à 66. La médiane se trouve donc entre la 33^e et 34^e valeur de la série.

On écrit les valeurs de la série dans l'ordre croissant :

0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 **2** **2** 2 2 2 3 3 3 3 3 …

↑

La 33^e et la 34^e valeur sont égales à 2. La médiane est donc également égale à 2.

On en déduit que durant la Coupe du monde 2010, il y a eu autant de matchs dont le nombre de buts était supérieur à 2 que de matchs dont le nombre de buts était inférieur à 2.

Remarque :La médiane

On considère une liste de N données rangées par ordre croissant.

Si la série est de taille **impaire** (N=2n+1), la *médiane* est la donnée de rang n+1.

Si la série est de taille **paire** (N=2n), la *médiane* est la demi-somme des données de rang n et de rang n+1.

IV. Paramètres de dispersion

1) Etendue:

C'est la différence entre les valeurs extrêmes.

Dans l'exemple 1, la valeur minimale est 8 et la valeur maximale est 18, donc l'étendue est égale à 18-8=10

Remarque : l'étendue est un enregistrement utile pour constater la dispersion de la série.

2) Ecart-moyen:

C'est la moyenne des écarts à la moyenne

L'écart-moyen *e* d'une série statistique de moyenne $\bar{x}$ dont les valeurs du caractère sont x_1, x_2, x_3, …, x_k et les effectifs correspondants sont n_1, n_2, n_3, …, n_k est égale à :

$$e=\frac{n_1\times\left|x_1-\bar{x}\right|+n_2\times\left|x_2-\bar{x}\right|+\ldots+n_k\times\left|x_k-\bar{x}\right|}{n_1+n_2+\ldots+n_k}=\frac{\sum_{i=1}^{k}n_i\times\left|x_i-\bar{x}\right|}{\sum_{i=1}^{k}n_i}$$

3) Variance

Définitions : - La variance *V* d'une série statistique de moyenne $\bar{x}$ dont les valeurs du caractère sont x_1, x_2, x_3, …, x_k et les effectifs correspondants sont n_1, n_2, n_3, …, n_k est égale à :

$$V=\frac{n_1\times\left(x_1-\bar{x}\right)^2+n_2\times\left(x_2-\bar{x}\right)^2+\ldots+n_k\times\left(x_k-\bar{x}\right)^2}{n_1+n_2+\ldots+n_k}=\frac{\sum_{i=1}^{k}n_i\times\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{k}n_i}$$

- L'écart-type σ d'une série statistique de variance *V* est égal à : $\sigma=\sqrt{V}$

Exemple : Dans exemple3

la variance est égale à :

$$V = \frac{7\times\left(0-\frac{7}{3}\right)^2 + 17\times\left(1-\frac{7}{3}\right)^2 + 13\times\left(2-\frac{7}{3}\right)^2 + 14\times\left(3-\frac{7}{3}\right)^2 + 8\times\left(4-\frac{7}{3}\right)^2 + 6\times\left(5-\frac{7}{3}\right)^2 + 0\times\left(6-\frac{7}{3}\right)^2 + 1\times\left(7-\frac{7}{3}\right)^2}{66}$$

$\approx 2,4646$

L'écart-type est $\sigma \approx \sqrt{2,4646} \approx 1,57$

Ainsi l'écart-type est environ égal à 1,57 buts.

Remarque :

L'écart-type exprime la dispersion des valeurs d'une série statistique autour de sa moyenne. Les valeurs extrêmes influencent l'écart-type.

Exemple :

Soit la série statistique suivante

7	2	1	Caractère
1	4	5	Effectif

La moyenne est : $m = \frac{5\times1+4\times2+1\times7}{10} = \frac{20}{10} = 2$

L'écart-moyen est égale à :

$$e = \frac{5\times|1-2|+4|2-2|+1\times|7-2|}{10} = \frac{5\times|-1|+4|0|+1\times|5|}{10}$$

$$e = \frac{5\times1+4\times0+1\times5}{10} = \frac{10}{10} = 1$$

La variance V est égale à :

$$V = \frac{5\times|1-2|^2+4|2-2|^2+1\times|7-2|^2}{10} = \frac{5\times|-1|^2+4|0|^2+1\times|5|^2}{10}$$

$$V = \frac{5\times1+4\times0+1\times25}{10} = \frac{30}{10} = 3$$

L'écart-type est égale à :

$$\sigma = \sqrt{V} = \sqrt{3}$$

Propriété : $V = \frac{n_1\times x_1{}^2 + n_2\times x_2{}^2 + \dots + n_k\times x_k{}^2}{n_1+n_2+\dots+n_k} - \left(\overline{x}\right)^2 = \frac{\sum_{i=1}^{k} n_i\times x_i{}^2}{\sum_{i=1}^{k} n_i} - \left(\overline{x}\right)^2$

Démonstration : $V = \frac{\sum_{i=1}^{k} n_i\times\left(x_i-\overline{x}\right)^2}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_i\times\left(x_i{}^2-2x_i\overline{x}+\overline{x}^2\right)}{\sum_{i=1}^{k} n_i}$

$$= \frac{\sum_{i=1}^{k} n_i\times x_i{}^2 - n_i\times 2x_i\overline{x} + n_i\times\overline{x}^2}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_i\times x_i{}^2}{\sum_{i=1}^{k} n_i} - \frac{\sum_{i=1}^{k} n_i\times 2x_i\overline{x}}{\sum_{i=1}^{k} n_i} + \frac{\sum_{i=1}^{k} n_i\times\overline{x}^2}{\sum_{i=1}^{k} n_i}$$

$$= \frac{\sum_{i=1}^{k} n_i\times x_i{}^2}{\sum_{i=1}^{k} n_i} - 2\overline{x}\frac{\sum_{i=1}^{k} n_i\times x_i}{\sum_{i=1}^{k} n_i} + \overline{x}^2\frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_i\times x_i{}^2}{\sum_{i=1}^{k} n_i} - 2\overline{x}\overline{x} + \overline{x}^2$$

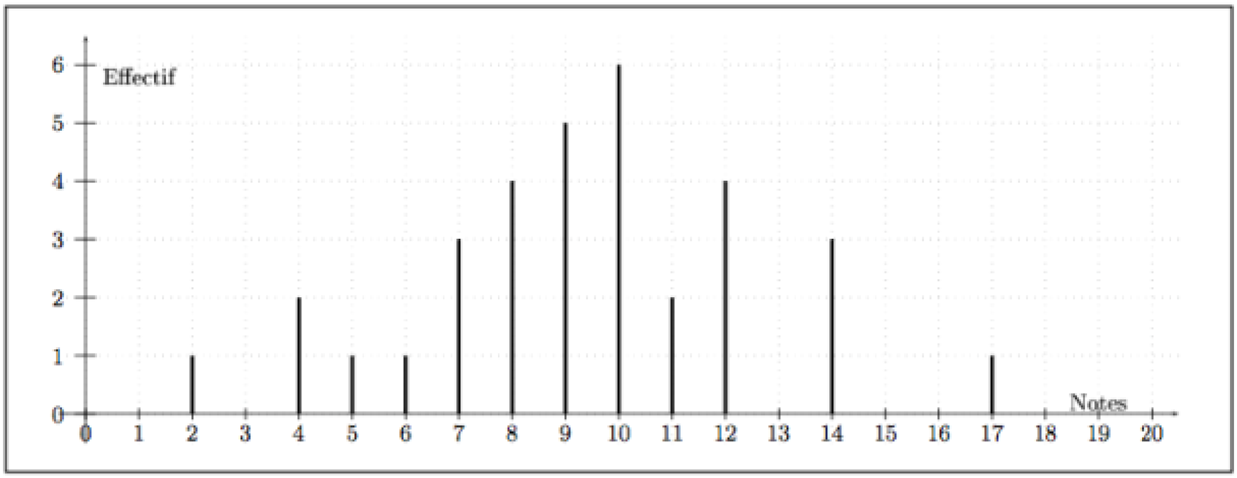
$$= \frac{\sum_{i=1}^{k} n_i\times x_i{}^2}{\sum_{i=1}^{k} n_i} - \overline{x}^2$$

V. Représentations graphiques :

1)Diagramme en bâtons

Lorsque le caractère étudié est quantitatif discret, on peut représenter la séri statistique par un diagramme en bâtons : La hauteur de chaque bâton est proportionnelle à l'effectif (ou la fréquence) associée à chaque valeur.

Exemple : Voici le diagramme en bâtons représentant une série de notes obtenues par une classe à un contrôle.



Recopiez et complétez le tableau suivant :

Notes	2	4	5	6	7	8	9	10	11	12	14	17	Total
Effectif													
pourcentage %)													

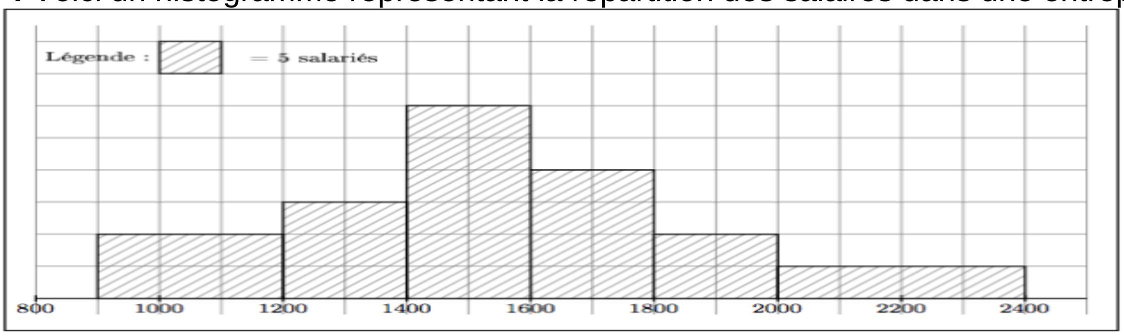
Solution :

Notes	2	4	5	6	7	8	9	10	11	12	14	17	Total
Effectif	1	2	1	1	3	4	5	6	2	4	3	1	33
pourcentage(%)	3%	6%	3%	3%	9%	12%	15%	18%	6%	12%	0%	3%	100%

2)Histogramme

Lorsque le caractère étudié est quantitatif continu, et lorsque les données sont regroupées en classes, on peut représenter la série par un histogramme : l'aire de chaque rectangle est alors proportionnelle à l'effectif (ou a la fréquence) associée à chaque classe.
Lorsque les classes ont la même amplitude, c'est la hauteur de chaque rectangle qui est proportionnelle à l'effectif.

Exemple : Voici un histogramme représentant la répartition des salaires dans une entreprise :



Recopier et compléter le tableau ci-dessous :

Salaires	[900;1200[	[1200;1400[	[1400;1600[	[1600;1800[	[1800;2000[	[2000;2400[	Total
Effectif							
Fréquence							

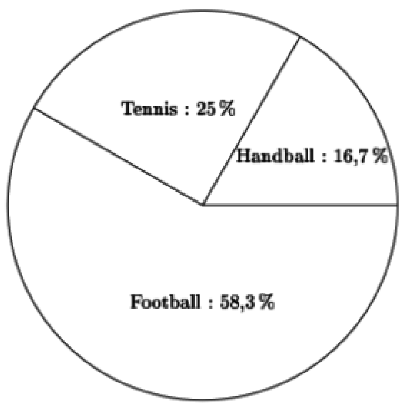
Solution

Salaires	[900;1200[	[1200;1400[	[1400;1600[	[1600;1800[	[1800;2000[	[2000;2400[	Total
Effectif	30	30	60	40	20	20	200
Fréquence	0,15	0,15	0,3	0,2	0,1	0,1	1

3)diagramme circulaire ou semi-circulaire

Lorsque le caractère est qualitatif, on représente la série par un diagramme circulaire ou semi-circulaire (camembert) : La mesure de chaque secteur angulaire est proportionnelle à l'effectif (ou à la fréquence) associée à la valeur du caractère.

Exemple : Voici un diagramme circulaire représentant la répartition des adhérents à un club sportif.



Sachant que le club compte 240 adhérents, combien d'adhérents jouent ...

- Au football ?
- Au tennis ?
- Au handball ?

Solution : On multiplie l'effectif total (240) par la fréquence de chaque caractère indiquée dans le camembert pour obtenir l'effectif du caractère. Ainsi :

- Football : 240 * 0,583 = 140
- Tennis : 240 * 0,25 = 60
- Handball : 240 * 0,167 = 40

« C'est en forgeant que l'on devient forgeron » Dit un proverbe.
C'est en s'entraînant régulièrement aux calculs et exercices

Que l'on devient un mathématicien