Caractéristiques d'une série statistique

Une caractéristique est une grandeur qu'on utilise pour résumer une série statistique
On distingue deux sortes de caractéristique : les caractéristiques de position et les caractéristiques de dispersion.

1. Caractéristiques de positions

1.1 Le mode

1.1.1 Variable discrète

Le mode ou dominante est la valeur la plus fréquente de la variable. C'est la variable qui a le plus grand effectif.

Si une série peut posséder un seul mode on dit qu'elle est uni modale. Si elle en possède plusieurs, on dit qu'elle est plurimodale.

Le mode est définie même si la variable est qualitative

Exemple

pour la série suivantes :

5 6 6 7 7 7 8 8 9 9 9 9 9 10 10 10 10 11 11 12 12 12 14 15 16

Le mode est 9.

1.1.2 variable continue

Pour une série classée, dont les classes sont d'égal effectif, la classe modale est la classe qui correspond au plus grand effectif.

Exemple

on étudie la vitesse de 400 véhicules enregistrée par un radar lors d'un contrôle routier.

Classes:vitesse en km/h	[65 ; 80[	[80 ; 95[	[95 ; 110[	[110 ; 125[	[125;140[	140 ; 155[	[155 ; 170[
Effectifs	39	176	97	51	23	12	2

La classe modale de cette série est : [80 ; 95[.

1.2 La moyenne

1.2.1 Variable discrète

Soit x_1, x_2,x_3, ….x_p les valeurs d'une série statistiques discrète, et n_1, n_2, n_3, …,n_p les effectifs associés .

N l'effectif totale.

La moyenne de cette série est :

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_p x_p}{N}$$

ou encore

$\overline{x} = f_1 x_1 + f_2 x_2 + \ldots f_p x_p$ f_i désignant la fréquence de x_i.

1.2.2 variable continue

Dans le cas d'une variable continue, la variable est remplacée par le centre de la classe.

Exemple

Le tableau suivant représente la note obtenue par des élèves lors d'un test en maths :

Notes	**[0 ; 5[**	**[5 ; 8[**	**[8 ; 12[**	**[12 ; 15[**	**[15 ; 20]**
Effectif	10	8	12	11	9
Centre de la classe (x_i)	2,5	6,5	10	13,5	17,5

$\overline{x}$ = [(2,5*10)+(6,5*8)+(10*12)+(13,5*11)+(17,5*9) / 50
= 503/50
= 10,06

1.2.3 Propriétés de la moyenne

(p1) Si une série (x_i ; n_i) a pour moyenne $\overline{x}$, et la série (y_i ; n_i) telle que $y_i = ax_i + b$, alors

$$\overline{y} = a\,\overline{x} + b$$

(p2) Pour deux séries statistiques S_1 et S_2 d'effectifs N1 et N2,et de moyenne $\overline{x_1}$ et $\overline{x_2}$, la moyenne de la série S regroupant les deux séries S_1 et S_2 est : $\overline{x} = \dfrac{N_1\overline{x}_1 + N_2\overline{x}_2}{N_1 + N_2}$

1.3 La médiane

1.3.1 Définition

La médiane d'une série statistique est le nombre qui partage la population en deux parties de même effectif : les individus dont la valeur du caractère est inférieure à la médiane, et les individus dont la valeur est supérieure à la médiane.

1.3.2 Détermination

a) Exemples pour le cas d'une variable discrète

Pour calculer la médiane Me d'une série a caractère discret,on peut :

- Construire le tableau des effectifs cumulés croissants et décroissants ;
- déterminer la modalité x_1 correspondant au premier effectif cumulé croissant supérieur ou égal à $\dfrac{N}{2}$
- déterminer la modalité x_2 correspondant au premier effectif cumulé décroissant supérieur ou égal à $\dfrac{N}{2}$

- Calculer Me = $\frac{x_1+x_2}{2}$

Exemple : Déterminons la médiane relative aux notes suivantes :

Note	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Effectif	2	2	5	10	9	10	12	10	7	5	5	1	1	1
Effectif cumulé croissant	2	4	9	19	28	38	50	60	67	72	77	78	79	80
Effectif cumulé décroissant	80	78	76	71	61	52	42	30	20	13	8	3	2	1

Le premier effectif cumulé croissant supérieur à 40 est 50, la modalité correspondante est x_1 =10 ;

Le premier effectif cumulé décroissant supérieur à 40 est 42 ; la modalité correspondante est x_2 =10.

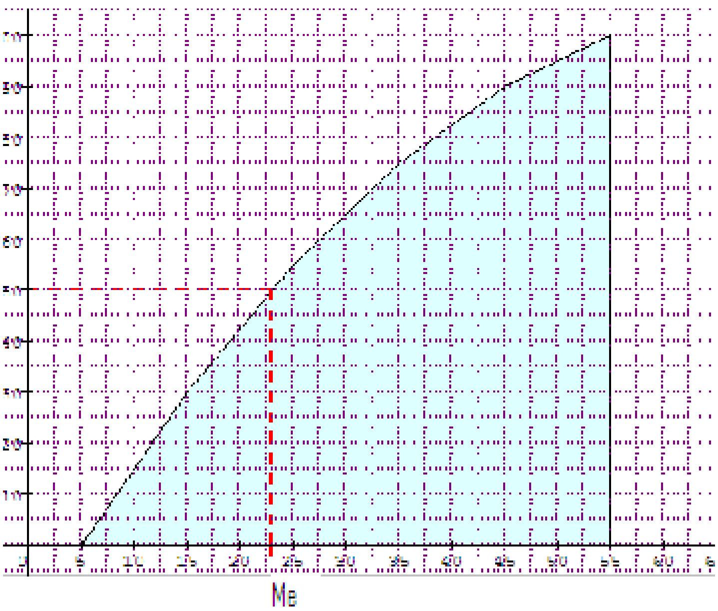
la médiane de la série est Me = $\frac{10+10}{2}$ =10

Pour déterminer la médiane Me d'une série statistique à caractère continue,on peut :

- construire le polygone des fréquences cumulées croissantes. La médiane est l'abscisse du point du polygone dont l'ordonnée est 0,5 ou50 %.
- Construire les polygones des effectifs cumulés croissants et décroissants. La médiane est l'abscisse du point d'intersection de ces deux polygones.

Exemple : Dans un magasin, sur un total de 200 paires de chaussures vendus, la répartition en millier d'Ariary est donné par le tableau suivant :

Prix(millier d'Ariary)	[5 ; 15[	[15 ; 25[	[25 ; 35[	[35;45[	[45;65[
Effectif	60	50	40	30	20
Fréquence (en%)	30	25	20	15	10
Fréquences cumulées croissantes	30	55	75	90	100



les caractéristiques de dispersion

•Définition

- La *variance* d'une série est la moyenne des carrés des écarts à la moyenne.

$$V=\frac{\sum_{i=1}^{p} n_i |x_i - \bar{x}|^2}{N}$$

- L'*écart type* d'une série est la moyenne quadratique des écarts à la moyenne. C'est la racine carrée de la variance.

Méthode de calcul

Avec des valeurs observées x_i très simples, il arrive souvent que la moyenne $\bar{x}$ soit un nombre décimal. Dans ce cas, le calcul de la variance V nécessite des calculs fastidieux.

On démontre que :

$$V=\frac{\sum_{i=1}^{p} n_i x_i^2}{N} - \bar{x}^2$$

Cette expression de la variance permet de simplifier les calculs.

Exemple

Deux nageuses, Vola et Bako ont participé à 32 compétitions (100 m nage libre). Le tableau suivant indique leur performance :

Temps en seconde	68	67	66	65	64	63	62	61
Nombre compétitions vola	0	3	8	10	7	3	0	1
Nombre compétitions Bako	2	4	6	7	7	3	2	1

- On vérifie que le temps moyen obtenu par chaque nageuse est 64,9 s
- Calculons l'écart-type de leurs performances.

Performance de Vola

Temps en seconde	68	67	66	65	64	63	62	61	Total
effectif	0	3	8	10	7	3	0	1	32
$x_i - \bar{x}$	3,1	2,1	1,1	0,1	-0,9	-1,9	-2,9	-3,9	
$(x_i - \bar{x})^2$	9,61	4,41	1,21	0,01	0,81	3,61	8,41	15,21	
$n_i (x_i - \bar{x})^2$	0	13,23	9,68	0,10	5,67	10,83	0	15,21	54,72

La variance de la série est $V_1 = \dfrac{54,72}{32} = 1,71$.

L'écart-type de la série est $\sqrt{1,71} = 1,30$

Performance de Bako

Temps en seconde	68	67	66	65	64	63	62	61	Total
effectif	2	4	6	7	7	3	2	1	32
$x_i - \bar{x}$	3,1	2,1	1,1	0,1	-0,9	-1,9	-2,9	-3,9	
$(x_i - \bar{x})^2$	19,22	17,64	7,26	0,07	5,67	10,83	16,82	15,21	92,72

La variance de la série est $V_2 = \dfrac{92,72}{32} = 2,89$. L'écart-type de la série est 1,7.

1,30 < 1,7 donc Vola est plus régulière que Bako.