

Séries statistiques à une variable

1/ Tableaux statistiques et diagrammes :

Les tableaux statistiques et les diagrammes permettent d'organiser et de présenter les données recueillies. Le caractère étudié peut être **qualitatif** ou **quantitatif**. Un caractère quantitatif peut être

- **discret** : il ne peut prendre que des valeurs isolées ;
- **continu** : il peut prendre toutes les valeurs d'un intervalle.

Dans le cas d'une répartition en classes, on utilise un histogramme pour représenter graphiquement les effectifs (ou les fréquences) : les aires des rectangles sont proportionnelles aux effectifs (ou aux fréquences).

Pour tracer le **polygone des effectifs cumulés croissants**, on place les points dont :

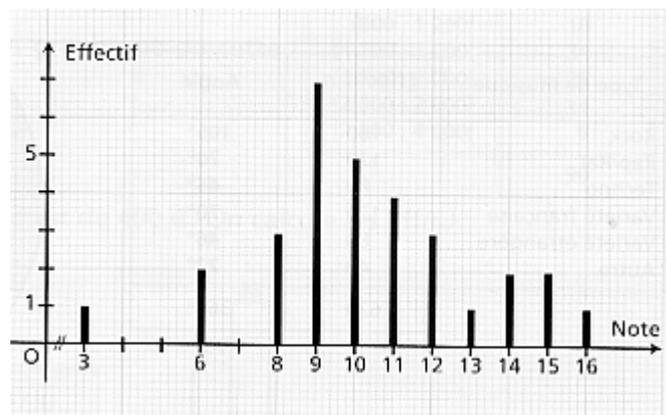
- l'abscisse est la limite supérieure d'une classe ;
- l'ordonnée est l'effectif cumulé croissant de cette classe.

Exemple : Les élèves d'une classe de Bac Pro réalisent trois enquêtes dont les informations sont données dans les tableaux suivants :

- Tableau 1 :
Notes obtenues par les 31 élèves
de la classe de Bac Pro
lors de l'évaluation de français

Note x_i	Effectif n_i
3	1
6	2
8	3
9	7
10	5
11	4
12	3
13	1
14	2
15	2
16	1
	31

Diagramme en bâtons

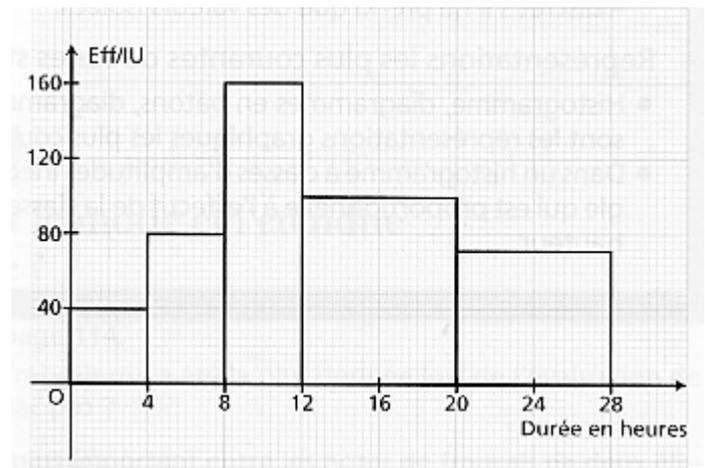


La notation est un caractère quantitatif (mesurable) discret (il prend des valeurs isolées).

- Tableau 2 : temps consacré chaque semaine par les élèves du lycée à regarder la télévision.
 - ▲ Population interrogée : les 620 élèves du lycée.
 - ▲ Intervalle unitaire (IU) : 4 heures
 - ▲ On calcule les Eff/IU. La hauteur des rectangles correspond à l'effectif par intervalle unitaire.

Histogramme

Durée h	Effectif n_i	Effectif/UI
$[0 ; 4[$	40	40
$[4 ; 8[$	80	80
$[8 ; 12[$	160	160
$[12 ; 20[$	200	100
$[20 ; 28[$	140	70
	620	

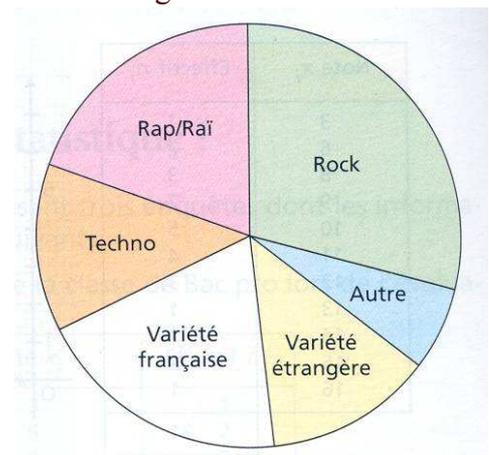


La durée est un caractère quantitatif continu (il peut prendre toutes les valeurs d'un intervalle).

- Tableau 3 : la musique préférée des élèves du lycée. On calcule les angles des secteurs correspondants aux effectifs n_i .

Type de musique	Effectif n_i	Angle
Rock	180	105°
Rap/Raï	120	70°
Techno	80	46°
Variété française	120	70°
Variété étrangère	80	46°
Autre	40	23°
	620	360°

Diagramme circulaire



Le caractère étudié est un caractère qualitatif (non mesurable).

- Mode et étendue statistique :
 - On appelle **mode d'une distribution statistique** la valeur de la variable qui a le plus grand effectif. Dans le cas d'une distribution en classes, on appelle **classe modale** la classe qui a le plus grand effectif par intervalle unitaire. Le centre de la classe modale est appelé **mode**.
 - Exemple* : dans le tableau 1 : la note qui a été le plus attribuée est 9.
9 est le mode de cette série statistique.
 - **L'étendue d'une série statistique** est la différence entre les valeurs extrêmes du caractère. Dans le tableau 1, la note minimum est 3, la maximum est 16. L'étendue est donc $16 - 3 = 13$

2/ Les indicateurs de l'analyse statistique

a. La moyenne d'une série statistique, notée \bar{x} est obtenue par :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N} = \frac{\sum_{i=1}^p n_i x_i}{N} \text{ avec}$$

- $N = \sum_{i=1}^p n_i$ et N effectif total
- x_i : valeur du caractère ou centre des classes
- n_i : effectif de x_i ou de la classe de centre x_i
- p : nombre de classes ou de valeurs différentes du caractère.

Exemple : dans le tableau 1, la note moyenne est calculée par :

$$\bar{x} = \frac{1 \times 3 + 2 \times 6 + 3 \times 8 + 7 \times 9 + 5 \times 10 + 4 \times 11 + 3 \times 12 + 1 \times 13 + 2 \times 14 + 2 \times 15 + 1 \times 16}{31} = 10,06$$

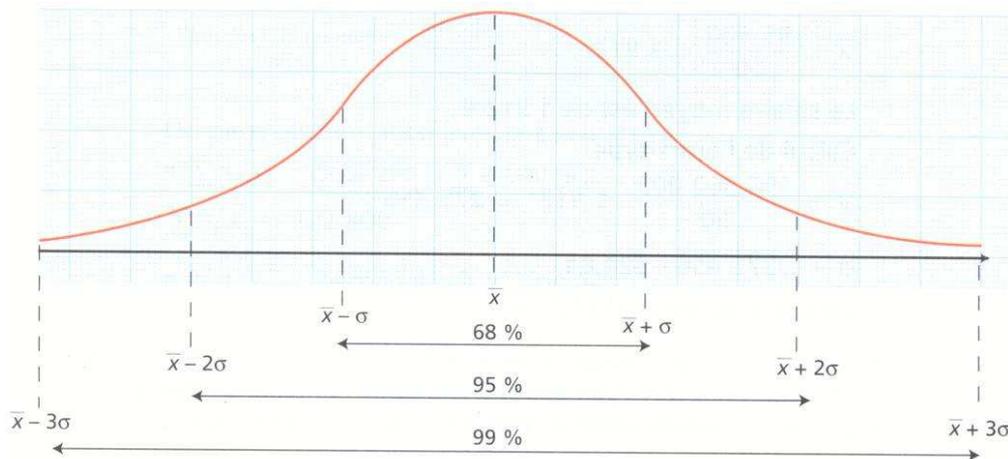
b. Variance et écart type d'une série statistique :

➤ La variance

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

➤ L'écart type $\sigma = \sqrt{V}$

- i. Plus σ est grand, plus les valeurs de la variable sont dispersées.
- ii. Lorsque la distribution statistique peut être représentée par une "courbe de Gauss", on constate la répartition suivante des effectifs autour de la moyenne \bar{x} :



- environ 68 % des valeurs appartiennent à l'intervalle $[\bar{x} - \sigma; \bar{x} + \sigma]$
- environ 95 % des valeurs appartiennent à l'intervalle $[\bar{x} - 2\sigma; \bar{x} + 2\sigma]$
- environ 99 % des valeurs appartiennent à l'intervalle $[\bar{x} - 3\sigma; \bar{x} + 3\sigma]$

Exemple : Exemple : Calcul de l'écart-type pour le tableau 2 : temps consacré chaque semaine à regarder la télévision :

Durée (h)	Effectif n_i	centre x_i	$(x_i - \bar{x})^2$	$n_i(x_i - \bar{x})^2$
[0 ; 4[40	2	144	5760
[4 ; 8[80	6	64	5120
[8 ; 12[160	10	4	640
[12 ; 20[200	16	4	800
[20 ; 28[140	24	100	14000
	620			26320

$$\text{Moyenne } \bar{x} = \frac{40 \times 2 + 80 \times 6 + 160 \times 10 + 200 \times 16 + 140 \times 24}{620} = 14 \text{ h}$$

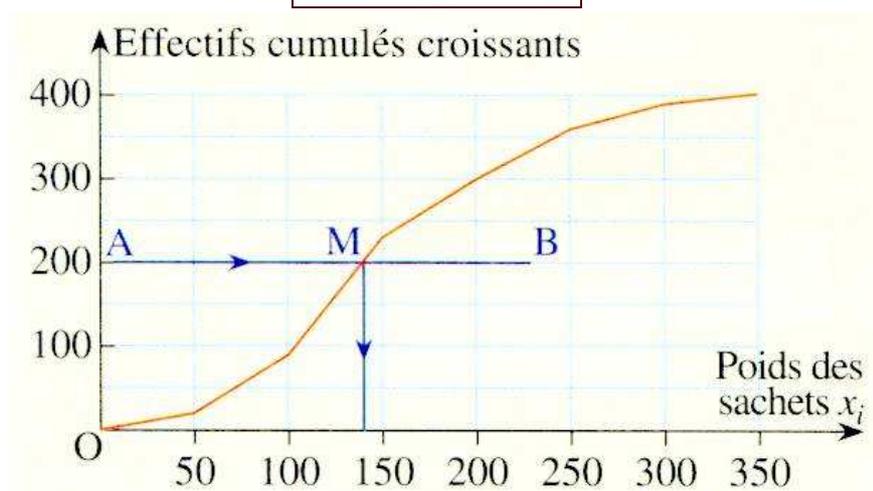
$$\text{Variance } V = \frac{26320}{620} = 42,45$$

$$\text{Ecart-type } \sigma = \sqrt{42,45} = 6,5 \text{ soit } 6 \text{ h } 30 \text{ min}$$

c. La médiane : c'est la valeur de la variable telle que le nombre de valeurs qui lui sont inférieures est égal au nombre de valeurs qui lui sont supérieures.

Exemple : La distribution statistique suivante a été établie suivant les ventes de sachets de bonbons dans un libre service :

Poids des sachets (en g) x_i	nombre de sachets n_i	Effectif cumulé croissant ECC
]0 ; 50]	20	20
]50 ; 100]	70	90
]100 ; 150]	140	230
]150 ; 200]	70	300
]200 ; 250]	60	360
]250 ; 300]	30	390
]300 ; 350]	10	400
	400	



Le rang de la médiane est $\frac{N}{2} = \frac{400}{2} = 200$. La droite (AB) d'ordonnée $\frac{N}{2}$ et parallèle à l'axe des abscisses coupe la courbe en M. L'abscisse du point M, égale à 139 est la valeur médiane de la série statistique.