

Statistiques

I - Effectifs et fréquences

Vocabulaire

Une étude statistique porte sur une **population**, dont chaque élément est un **individu**. On observe sur chacun un **caractère**, qui prend différentes **valeurs**.

Définitions

- L'**effectif** d'une valeur est le nombre d'individus qui la présentent.
- L'**effectif total** est le nombre d'individus de la population.
- La **fréquence** d'une valeur est le quotient

$$f = \frac{\text{effectif de la valeur}}{\text{effectif total}}.$$

Remarque

Une fréquence est un nombre compris entre 0 et 1 ; multipliée par 100, elle s'exprime en pourcentage. La somme de toutes les fréquences vaut 1, soit 100%.

Exemple — Les notes d'un contrôle

Note	8	10	12	14	16	Total
Effectif	2	5	8	7	3	25
Fréquence	0,08	0,2	0,32	0,28	0,12	1

La fréquence de la note 12 vaut $\frac{8}{25} = 0,32$, soit 32% des élèves.

Méthode 1 — Retrouver un effectif à partir d'une fréquence

L'effectif est le produit de la fréquence par l'effectif total :

$$\text{effectif} = f \times \text{effectif total}.$$

II - Indicateurs de position : moyenne et médiane

Définition 1

La **moyenne** d'une série est le quotient de la somme de toutes les valeurs par l'effectif total.

Formule — Moyenne d'une série donnée par un tableau d'effectifs

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{n_1 + n_2 + \dots + n_p}$$

où x_i sont les valeurs et n_i leurs effectifs.

Exemple

Avec le tableau des notes :

$$\bar{x} = \frac{2 \times 8 + 5 \times 10 + 8 \times 12 + 7 \times 14 + 3 \times 16}{25} = \frac{308}{25} = 12,32.$$

Définition 2

La **médiane** d'une série est une valeur qui partage la série rangée dans l'ordre croissant en deux groupes de même effectif : au moins la moitié des valeurs lui sont inférieures ou égales, au moins la moitié lui sont supérieures ou égales.

Méthode 2 — Déterminer la médiane

1. Ranger toutes les valeurs dans l'ordre croissant, en répétant celles qui apparaissent plusieurs fois.
2. Si l'effectif est impair, la médiane est la valeur centrale.
3. S'il est pair, la médiane est la demi-somme des deux valeurs centrales.



Fig. 1. – Sept valeurs : la médiane est la quatrième.

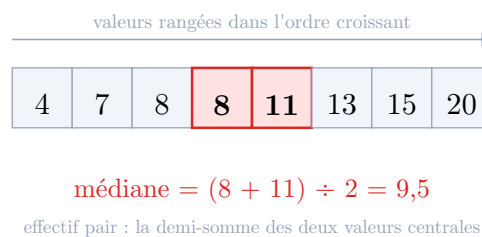


Fig. 2. – Huit valeurs : la médiane est la demi-somme de la quatrième et de la cinquième.

Remarque

Moyenne et médiane ne disent pas la même chose. Une valeur très grande tire la moyenne vers le haut sans déplacer la médiane : sur la série 1, 2, 3, 4, 200, la moyenne vaut 42 et la médiane 3.

Exemple — Interpréter une médiane

La médiane des notes du contrôle est 12 : au moins la moitié de la classe a obtenu 12 ou moins, au moins la moitié a obtenu 12 ou plus.

III - Étendue d'une série

Définition 3

L'**étendue** d'une série est la différence entre sa plus grande et sa plus petite valeur.

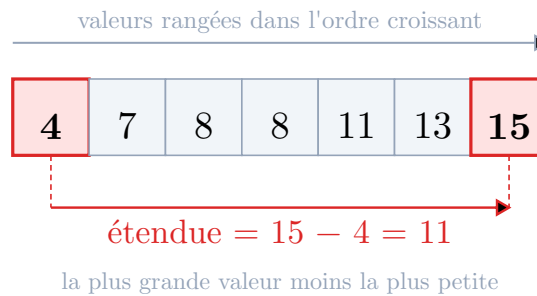


Fig. 3. – L'étendue se lit aux deux extrémités de la série rangée.

Remarque

L'étendue est un indicateur de **dispersion**, et non de position : elle dit si les valeurs sont resserrées ou étalées, pas où elles se situent.

Exemple — Deux séries de même moyenne

Série A	Série B
9, 10, 10, 10, 11	2, 5, 10, 15, 18
moyenne 10, étendue 2	moyenne 10, étendue 16

Même moyenne, même médiane, mais des séries très différentes : seule l'étendue le montre.

IV - Regroupement en classes et histogrammes

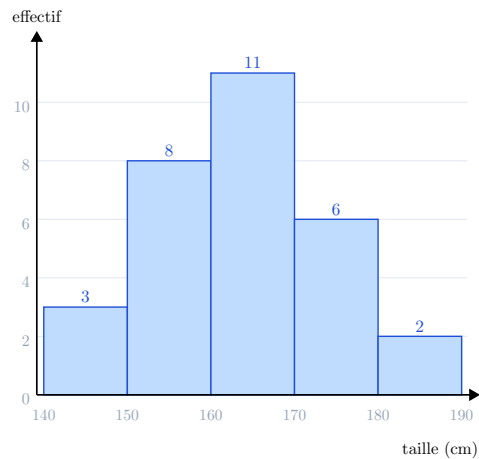
Quand le caractère prend un grand nombre de valeurs — une taille, une durée, une masse — les lister une à une n'apprend rien. On les regroupe alors en **classes**.

Définition 4

Une **classe** est un intervalle de valeurs. Son **amplitude** est sa largeur. Quand toutes les classes ont la même amplitude, les effectifs sont directement comparables.

Définition 5

Un **histogramme** représente une série regroupée en classes : à chaque classe correspond un rectangle, dont la largeur est celle de la classe et la hauteur est proportionnelle à l'effectif. Les rectangles sont **contigus**.



cinq classes de même amplitude (10 cm) : les rectangles se touchent

Fig. 4. – Les tailles de 30 élèves, regroupées en cinq classes de 10 cm.

Un histogramme n'est pas un diagramme en bâtons : les bâtons sont séparés parce que les valeurs sont isolées, les rectangles se touchent parce que les classes se suivent sans trou.

Méthode 3 — Construire un histogramme

1. Choisir des classes de même amplitude couvrant toutes les valeurs.
2. Compter l'effectif de chaque classe.
3. Graduer l'axe horizontal par les bornes des classes.
4. Tracer un rectangle par classe, de hauteur proportionnelle à l'effectif.

Remarque

Une fois la série regroupée, les valeurs individuelles sont perdues : la moyenne se calcule alors avec le centre de chaque classe, et ne donne qu'une valeur approchée.

V - Lire et interpréter une représentation

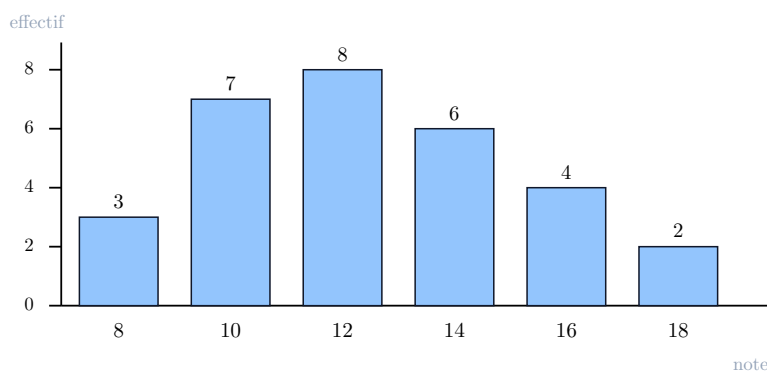


Fig. 5. – Le diagramme en barres : la hauteur de chaque barre est proportionnelle à l'effectif.

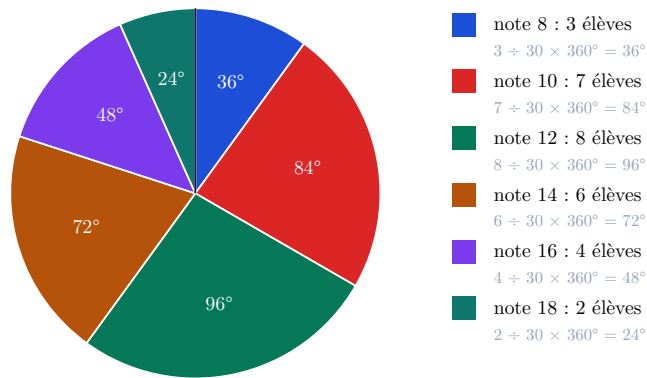


Fig. 6. – Le diagramme circulaire : l'angle de chaque secteur est proportionnel à l'effectif, le tour complet valant 360° .

Méthode 4 — Choisir une représentation

- Un **diagramme en bâtons** convient à un caractère prenant peu de valeurs isolées.
- Un **diagramme circulaire** met en évidence des parts d'un tout.
- Un **histogramme** convient à un caractère regroupé en classes.

Méthode 5 — Lire une série sur un diagramme

Tous les indicateurs se lisent sur n'importe quelle représentation, à condition de reconstituer d'abord le tableau des effectifs : c'est de lui que viennent moyenne, médiane et étendue.

Exemple — Une lecture critique

Un diagramme circulaire montre qu'un candidat obtient un quart des voix ; il ne dit rien du nombre de votants. Un même quart peut représenter 25 voix ou 250 000.

VI - Traitement de données à l'aide d'un tableur

Protocole — Les fonctions utiles

Fonction	Ce qu'elle renvoie
=SOMME (B2 : B8)	la somme des cellules de la plage
=MOYENNE (B2 : B8)	la moyenne des valeurs
=MEDIANE (B2 : B8)	la médiane des valeurs
=MAX (B2 : B8) - MIN (B2 : B8)	l'étendue
=NB (B2 : B8)	l'effectif total

Protocole — Calculer des fréquences

1. Placer les valeurs en colonne A et les effectifs en colonne B.
2. En B9, écrire =SOMME(B2:B8) pour l'effectif total.
3. En C2, écrire =B2/\$B\$9, puis recopier la formule vers le bas.
4. Le symbole \$ fige la cellule B9 pendant la recopie.

Remarque

Pour une série donnée par un tableau d'effectifs, =MOYENNE ne convient pas : il faut écrire =SOMMEPROD(A2:A8;B2:B8)/SOMME(B2:B8), qui multiplie chaque valeur par son effectif.