

Statistiques

Vocabulaire

Voici les moyennes de mathématiques d'une classe de troisième :

13,5 ; 13 ; 02 ; 04 ; 04,5 ; 03,5 ; 03 ; 14 ; 18,5 ; 17 ; 09,5 ; 14 ; 03 ; 13,5 ; 17

On appelle **population** l'ensemble des **individus** étudié, ici des élèves.

Le **caractère** étudié est **quantitatif** : une note sur 20 en mathématiques.

Le **caractère** peut être **qualitatif** : la couleur des yeux, le sexe...

L'**effectif** d'une série statistiques est le nombre total d'individus.

Cette série statistique est donnée sous forme **exhaustive**, on connaît toutes les notes.

Un sére statistique peut-être donnée sous forme d'un **regroupement en classes**.

Classe	[0 ; 5 [[5 ; 10 [[10 ; 15 [[15 ; 20]
Effectif	6	1	5	3

La représentation en classe fait perdre des informations. Pour calculer la moyenne il faut faire la moyenne des centres de classe pondérés par les effectifs

Moyenne arithmétique

Notons x_1, x_2, \dots, x_n les valeurs du caractère d'une série statistique.

La **moyenne arithmétique** est $M = \frac{x_1 + x_2 + \dots + x_n}{n}$

Pour le premier exemple on obtient :

$$M = \frac{13,5 + 13 + 2 + 4 + 4,5 + 3,5 + 3 + 14 + 18,5 + 17 + 9,5 + 14 + 3 + 13,5 + 17}{15} = \frac{150}{15} = 10$$

Moyenne pondérée

Pour le second exemple, les centres de classe sont 2,5 ; 7,5 ; 12,5 et 17,5.

La moyenne des centres de classe pondérée par les effectifs est :

$$M = \frac{6 \times 2,5 + 1 \times 7,5 + 5 \times 12,5 + 3 \times 17,5}{6 + 1 + 5 + 3} = \frac{137,5}{15} \approx 9,17$$

L'écart constaté montre la perte d'informations dans le regroupement en classes.

Fréquences

On calcule la **fréquence** qui correspond à une classe en calculant le quotient suivant :

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

Une **fréquence** est un nombre compris entre 0 et 1, on peut l'écrire en pourcentage.

Classe	[0 ; 5 [[5 ; 10 [[10 ; 15 [[15 ; 20]	Total
Effectif	6	1	5	3	15
Fréquence	$\frac{6}{15} = 0,4$	$\frac{1}{15} \approx 0,067$	$\frac{5}{15} \approx 0,333$	$\frac{3}{15} = 0,2$	$\frac{15}{15} = 1$
Fréquence en %	40 %	7 %	33 %	20 %	100 %

Étendue

L'**étendue** est la différence entre la valeur maximale du caractère et sa valeur minimale. L'**étendue** permet de mesurer la dispersion d'une série statistique.

Ici, la note minimale est 02 et la note maximale est 18,5. L'étendue est 18,5-2=16,5. C'est un premier indicateur de la dispersion des notes de mathématiques.

Médiane

La **médiane** est une valeur du caractère qui partage la série classée dans l'ordre croissant en deux séries d'effectifs égaux.

La série précédente est constituée de 15 valeurs. On peut partager un effectif de 15 individus en deux classes de 7. Le 8ème sépare la population en deux.

Classons la série précédente dans l'ordre croissant :

02 ; 03 ; 03 ; 03,5 ; 04 ; 04,5 ; 09,5 ; 13 ; 13,5 ; 13,5 ; 14 ; 14 ; 17 ; 17 ; 18,5

Première partie de la population

Médiane

Seconde partie de la population

Cela signifie que la moitié des élèves a une note inférieure à 13 et l'autre moitié une note supérieure à 13. L'écart avec la moyenne arithmétique montre la dispersion des notes.

Si la série statistique a un effectif pair, par exemple 14 :

02 ; 03 ; 03 ; 03,5 ; 04 ; 04,5 ; 09,5 ; 13 ; 13,5 ; 13,5 ; 14 ; 14 ; 17 ; 18,5

Première partie de la population

Seconde partie de la population

Médiane

Toutes valeurs entre 9,5 et 13 est une médiane.

On prend souvent la moyenne des deux ici $\frac{9,5+13}{2} = 11,25$