Statistiques

Une étude statistique se base sur « **une population** », dont les éléments sont « **des individus** ».
On choisit d'étudier un **aspect** de ces individus, qui est « **le caractère** ».
Par exemple, un groupe (population) d'étudiants (individus) qui préparent le CRPE peut être étudié sur les caractères suivants : années d'études, situation familiale, etc.

1) Les caractères

On distingue deux types de caractères :

– les caractères quantitatifs : nombre de...
→ quantitatifs discrets : les valeurs sont des nombres entiers
→ quantitatifs continus : les valeurs sont des nombres réels, souvent approchés par des décimaux

– les caractères qualitatifs : nature de...

2) Effectif et fréquence

L'effectif d'une valeur d'un caractère = **nombre d'individus** de la population étudiée qui a cette valeur.
La fréquence d'une valeur d'un caractère = **quotient** de l'effectif par l'effectif total (souvent en %).

Ex : Voici le nombre de frères et sœurs des élèves d'un groupe de chant :
1 ; 2 ; 5 ; 2 ; 3 ; 0 ; 1 ; 0 ; 4 ; 2
L'effectif de la série est de 10.
L'effectif des individus qui ont 1 seul frère ou 1 seule sœur est de 2.
La fréquence de cette valeur (avoir 1 seul frère ou 1 seule sœur) est de 2/10, soit 20%.

3) Caractéristiques de position

a. Moyenne

La moyenne d'une série statistique = **addition de toutes les données ÷ nombre total de données**

La moyenne est la **valeur unique** que devraient avoir **tous les individus** de la population étudiée pour que le **total des valeurs soit inchangé**.
La moyenne est toujours **comprise** entre la **valeur minimale** et la **valeur maximale** de la série.

b. Moyenne pondérée

La moyenne pondérée = **addition des (valeurs× coefficient)÷somme des coefficients**

Ex : Lucie a eu 18 en français (coef. 3), 10 en maths (coef. 1) et 12 en allemand (coef. 2).
Sa moyenne pondérée est donc de : $\frac{18 \times 3 + 10 \times 1 + 12 \times 2}{3+1+2} = 14,7.$

c.Médiane

La **médiane** d'une série statistique est le nombre tel que, lors cette série est rangée dans **l'ordre croissant**, il y a **autant dedonnées supérieures** à la médiane **que dedonnées inférieures**.

Si le nombre de données est impair, la médiane est une de ces données.
Si le nombre de données est pair, ce n'est pas le cas, sauf si les deux valeurs centrales sont égales.

Ex : La médiane de la série 5 ; 6 ; 8 ; 12 ; 13 ; 54 ; 62 est 12 (il y a 3 valeurs avant et 3 valeurs après).
La médiane de la série 5 ; 6 ; 8 ; **12 ; 13** ; 54 ; 62, 70 est compris entre 12 et 13. C'est donc 12,5.
La médiane de la série 5 ; 6 ; 8 ; **12 ; 12** ; 54 ; 62, 70 est 12.

3) Caractéristiques de dispersion

a.Étendue

Étendue d'une série statistique = **plus grande des données – plus petite des données.**
C'est donc **l'écart** entre la donnée la plus grande et la donnée la plus petite.

Ex : L'étendue de la série 5 ; 6 ; 8 ; 12 ; 13 ; 54 ; 62 est égale à **62–5 = 57.**

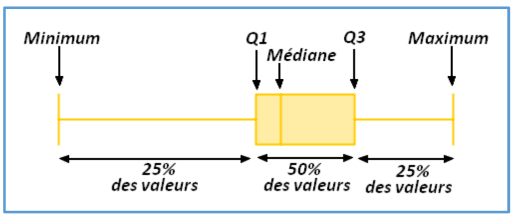
b.1er et 3^{e} quartile

Si les données statistiques d'une série sont rangées dans l'ordre croissant, alors :
– **le 1er quartile (Q1)**= le plus petit élément des données, tel qu'au moins **25 % des données sont inférieures ou égales à Q1.**
– **le 3^{e}quartile (Q3)** = le plus petit élément des données, tel qu'au moins **75 % des données sont inférieures ou égales à Q3**.

Ex : **Le 1er quartile** de la série 5 ; 6 ; 8 ; 12 ; 13 ; 54 ; 62 ; 81 est 6.
Le 3^{e} quartile de cette même série est donc54.

c.Diagramme en boîte à moustaches

Ce diagramme sert à représenter sur un même plan **la valeur minimale** et **la valeur maximale** de la série, ainsi que la **médiane** et les **1er et 3^{e} quartiles.**
L'écart entre les quartiles correspond à l'étendue de la série dont on a enlevé 25 % des plus petites valeurs et 25 % des plus grandes valeurs.



Méthode

1) Calculer la médiane d'une série statistique

• Ranger les nombres de la série dans l'ordre croissant.
• Déterminer l'effectif total de la série (nous l'appellerons N).
• Si N est impair, alors la médiane est donnée par $\frac{N-1}{2}+1$. Le nombre obtenu correspond au « rang » qu'aura la médiane dans la série. C'est le nombre qui se situe « au milieu » de la série.
• Si N est pair, alors la médiane est le nombre qui se trouve « au milieu » de $\frac{N}{2}$ et de $\frac{N}{2}+1$.

Ex : La médiane de la série 5 ; 6 ; 8 ; 12 ; 13 ; 54 ; 62 est 12
En effet, N = 7, donc $\frac{7-1}{2}+1$ = 4. La médiane sera au 4e rang de la série.

La médiane de la série 5 ; 6 ; 8 ; 12 ; 13 ; 54 ; 62, 70 est comprise entre 12 et 13. C'est donc 12,5.
En effet, N = 8, donc $\frac{8}{2}$ = 4 et $\frac{8}{2}+1$ = 5. La médiane sera le nombre qui qui se trouve « entre » le nombre du rang 4 et le nombre du 5, soit entre 12 et 13. Donc 12, 5.

La médiane de la série 5 ; 6 ; 8 ; 12 ; 12 ; 54 ; 62, 70 est 12, puisque les nombres de rangs 4 et 5 sont les mêmes.

2) Déterminer le 1er et le 3e quartile d'une série statistique

• Ranger les nombres de la série dans l'ordre croissant.
• Déterminer l'effectif total de la série (nous l'appellerons N).
• Si N est divisible par 4, alors le 1er quartile est donné par le quotient $\frac{N}{4}$. Le nombre obtenu correspond au rang du 1er quartile. Le 3e quartile sera donc égal à ce quotient multiplié par 3.
• Si N n'est pas divisible par 4, alors on détermine le plus petit entier supérieur au quotient de $\frac{N}{4}$. Cet entier le rang du 1er quartile. Le plus petit entier supérieur à $3 \times \frac{N}{4}$ sera le rang du 3e quartile.

Ex : Trouver le 1er et le 3e quartile de la série : 5 ; 6 ; 8 ; 12 ; 13 ; 54 ; 62.
• N = 7, donc $\frac{7}{4}$ = 1,75. On prend le plus petit entier supérieur à 1,75, soit 2.
Le 1er quartile est au deuxième rang de la série : **il s'agit du nombre 6.**
• Pour trouver le 3e quartile, il suffit de multiplier 1,75 par 3. On obtient 5,25. Le plus petit entier supérieur est 6. Le 3e quartile se trouve donc au sixième rang de la série : **c'est le nombre 54.**

Trouver le 1er et le 3e quartile de la série : 5 ; 8 ; 8 ; 16 ; 13 ; 60 ; 62 ; 75.
• N = 8, donc $\frac{8}{4}$ = 2.
Le 1er quartile est au deuxième rang de la série : **il s'agit du nombre 8.**
• Pour trouver le 3e quartile, il suffit de multiplier 2 par 3. On obtient 6. Le 3e quartile se trouve donc au sixième rang de la série : **c'est le nombre 60.**