

Recopier et compléter ce tableau avec des **définitions précises, rigoureuses et en français correct.**

Les notions sont celles du collège et concernent les séries statistiques à une variable... Ne recopiez pas au hasard des définitions trouvées dans des documents ou sur le net et n'ayant aucun lien avec ce programme.

Un exemple est donné pour illustrer quelques définitions.

Savoir le vocabulaire descriptif:

La population est l'ensemble des éléments ou individus sur lequel porte l'étude statistique, sur lequel on effectue des observations, des classements...

Un échantillon est une partie ou un sous-ensemble de la population.

Remarque: le choix de l'échantillon est important pour la validité de l'étude statistique.

Le caractère est le sujet sur lequel porte l'étude statistique. Un caractère prend différentes valeurs.

Remarque: Pour une même population, on peut définir de nombreux caractères.

L'effectif est le nombre d'individus prenant une ou des valeurs du caractère.

L'effectif cumulé est la somme des effectifs jusqu'à une valeur donnée incluse du caractère.

Une série statistique est qualitative lorsque les valeurs prises par le caractère ne sont pas des valeurs numériques (On ne peut pas faire d'opérations (somme, produit, ...) sur ces valeurs.

Exemple: couleur des yeux,

Une série statistique est quantitative lorsque le caractère prend des valeurs numériques.

Exemples: la taille des personnes dans un groupe; le nombre de pièces par appartement

Une série statistique quantitative est discrète lorsque les valeurs prises par le caractère étudié sont isolées.

Exemple: le nombre de pièces par appartement

Une série statistique quantitative est continue lorsque le caractère peut prendre toutes les valeurs d'un intervalle.

Exemple: la taille des personnes dans un groupe

savoir le vocabulaire et calculer:

L'étendue d'une série est donnée par la différence entre les valeurs extrêmes prises par le caractère.

étendue= valeur maximale – valeur minimale

Le mode est la valeur du caractère ayant le plus fort effectif. (C'est la valeur du caractère la plus fréquente. Une série statistique peut avoir plusieurs modes)

La moyenne (moyenne pondérée) est la somme de toutes les valeurs (comptées autant de fois qu'elle apparaît dans la série) divisée par l'effectif total. On a: $\bar{x} = \frac{n_1 \times x_1 + n_2 \times x_2 + \dots + n_p \times x_p}{n_1 + n_2 + \dots + n_p}$ où les x_i représentent les valeurs prises par le caractère et les n_i leur effectif.

Pour déterminer la médiane et les quartiles, la série étant ordonnée dans l'ordre croissant des valeurs du caractère,

La médiane est la valeur qui partage la population en deux parties égales: 50% de l'effectif prend une valeur du caractère inférieure ou égale à la médiane et 50% de l'effectif prend une valeur supérieure ou égale à la médiane.

Pour calculer la médiane, on range la série par ordre croissant des valeurs x_i et on a:

si l'effectif total N est impair, la médiane est donnée par la valeur de rang $\frac{N+1}{2}$

si l'effectif total N est pair, la médiane est donnée par la demi-somme (moyenne arithmétique) des valeurs de rang $\frac{N}{2}$ et $\frac{N}{2}+1$.

Le premier quartile Q_1 est la plus petite valeur de la série telle que 25 % (un quart) au moins des effectifs ont une valeur inférieure ou égale à Q_1 .

Important : Dans cette définition, le quartile Q_1 est une valeur de la série

(il existe une autre définition où Q_1 n'est pas nécessairement une valeur de la série et que certains logiciels utilisent).

En pratique : méthode de détermination de Q_1 :

on range la série dans l'ordre des valeurs croissantes.

On divise l'effectif N par 4 (25% de l'effectif). ($q_1 = \frac{N}{4}$).

Si q_1 est un entier, Q_1 est la valeur du $q_1^{\text{ième}}$ élément.

Si q_1 n'est pas un entier, on prend l'entier q_1' juste au-dessus, et Q_1 est la valeur du $q_1'^{\text{ième}}$ élément.

Le troisième quartile Q_3 est la plus petite valeur de la série telle que 75 % (trois-quarts) au moins des effectifs ont une valeur inférieure ou égale à Q_3 .

Important : Dans cette définition, le quartile Q_3 est une valeur de la série

En pratique : méthode de détermination de Q_3 :

on range la série dans l'ordre des valeurs croissantes.

On multiplie l'effectif N par $\frac{3}{4}$ (75% de l'effectif). ($q_3 = \frac{3N}{4}$).

Si q_3 est un entier, Q_3 est la valeur du $q_3^{\text{ième}}$ élément.

Si q_3 n'est pas un entier, on prend l'entier q_3' juste au-dessus, et Q_3 est la valeur du $q_3'^{\text{ième}}$ élément.

La fréquence d'une valeur est le quotient de l'effectif de cette valeur par l'effectif total.

On note $f_i = \frac{n_i}{N} = \frac{\text{effectif de la valeur } x_i}{\text{effectif total}}$

Remarque: la moyenne $\bar{x} = f_1 \times x_1 + \dots + f_p \times x_p$

La fréquence cumulée croissante est la somme des fréquences **jusqu'à** la valeur donnée incluse. On classe les valeurs x_i dans l'ordre croissant; $f_1 + f_2 + \dots + f_k$ donne la fréquence d'individus prenant une valeur inférieure ou égale à la valeur x_k .

Lorsqu'on a calculer les fréquences cumulées (et réaliser le digramme des fréquences cumulées), la détermination des quartiles et de la médiane est immédiate.

savoir le vocabulaire et construire:

Le diagramme en bâtons est utilisé lorsque les valeurs sont discrètes ou qualitatives.

En abscisses, on place les valeurs du caractère, en ordonnées les effectifs.

Un bâton représente alors l'effectif de la valeur. La **hauteur** du bâton est **proportionnelle** à l'effectif (ou à la fréquence) de la valeur du caractère représenté.

Le diagramme circulaire est construit en créant des secteurs circulaires dans un disque de façon à ce que les **angles** des secteurs sont **proportionnels** aux effectifs (ou aux fréquences).

Un histogramme est construit en juxtaposant des rectangles de façon à ce que les **aires** des rectangles soient **proportionnelles** aux effectifs (ou aux fréquences)

Exemple: Les 35 élèves d'une classe ont obtenu les notes suivantes à un test :

9	11	8	5	15	9	14
7	11	5	6	8	6	10
11	8	8	7	13	4	13
5	5	6	4	10	5	7
4	11	4	6	9	7	4

1) Écrire la série dans l'ordre des valeurs croissantes et donner l'effectif de chaque note ainsi que l'effectif cumulé croissant.

notes	4	5	6	7	8	9	10	11	13	14	15
effectif	5	5	4	4	4	3	2	4	2	1	1
effectif cumulé	5	10	14	18	22	25	27	31	33	34	35

2) La **fréquence** de la note 11 est $\frac{4}{35}$ *(La fréquence est un nombre entre 0 et 1)*

La distribution de fréquences d'une série statistique est la donnée des valeurs prises par la série et la fréquence de chacune.

notes	4	5	6	7	8	9	10	11	13	14	15	Total
fréquences	$\frac{5}{35}$	$\frac{5}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{3}{35}$	$\frac{2}{35}$	$\frac{4}{35}$	$\frac{2}{35}$	$\frac{1}{35}$	$\frac{1}{35}$	1

La **moyenne** de cette série.

$$\bar{x} = \frac{4 \times 5 + 5 \times 5 + 6 \times 4 + 7 \times 4 + 8 \times 4 + 9 \times 3 + 10 \times 2 + 11 \times 4 + 13 \times 2 + 14 \times 1 + 15 \times 1}{35} = \frac{275}{35} \approx 7,86$$

(En ajoutant les 35 notes, la somme est 275)

La **médiane** est la 18^{ième} valeur: $Me = 7$

(La population est partagée en deux parties égales. Il y a autant de notes inférieures à 7 que de notes supérieures à 7)

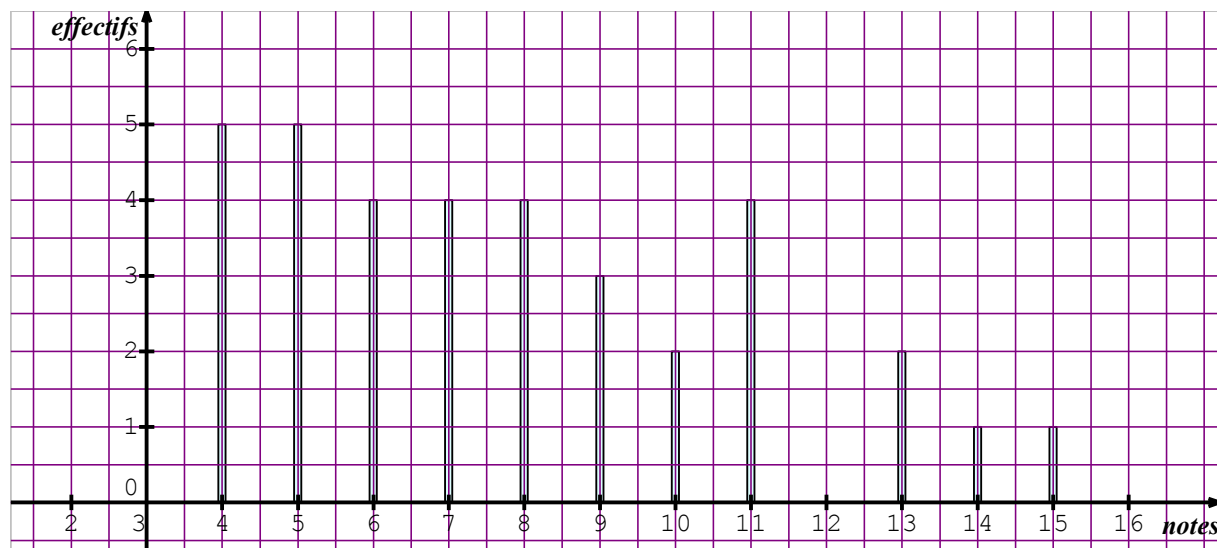
Le **premier quartile** est la neuvième valeur: $Q_1 = 5$

(5 est la plus petite valeur de la série telle qu'au moins 25 % des données sont inférieures ou égales à cette valeur)

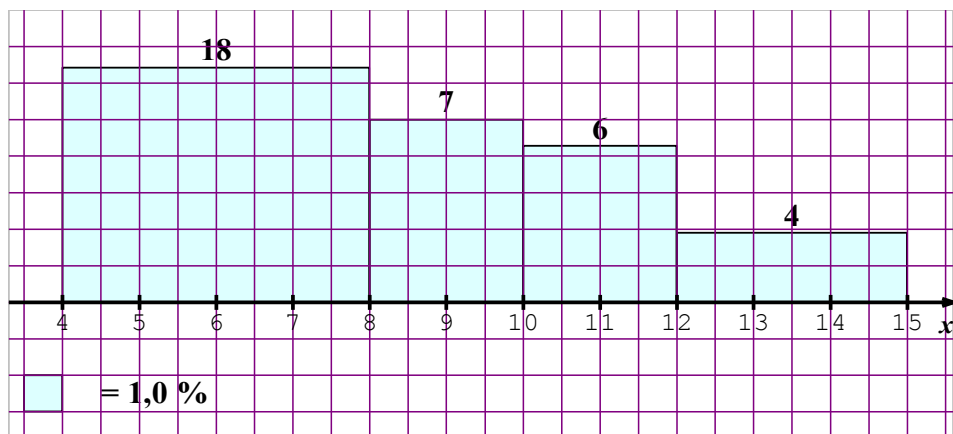
Le **troisième quartile** est la 27^{ième} valeur: $Q_3 = 10$

(10 est la plus petite valeur de la série telle qu'au moins 75 % des données sont inférieures ou égales à cette

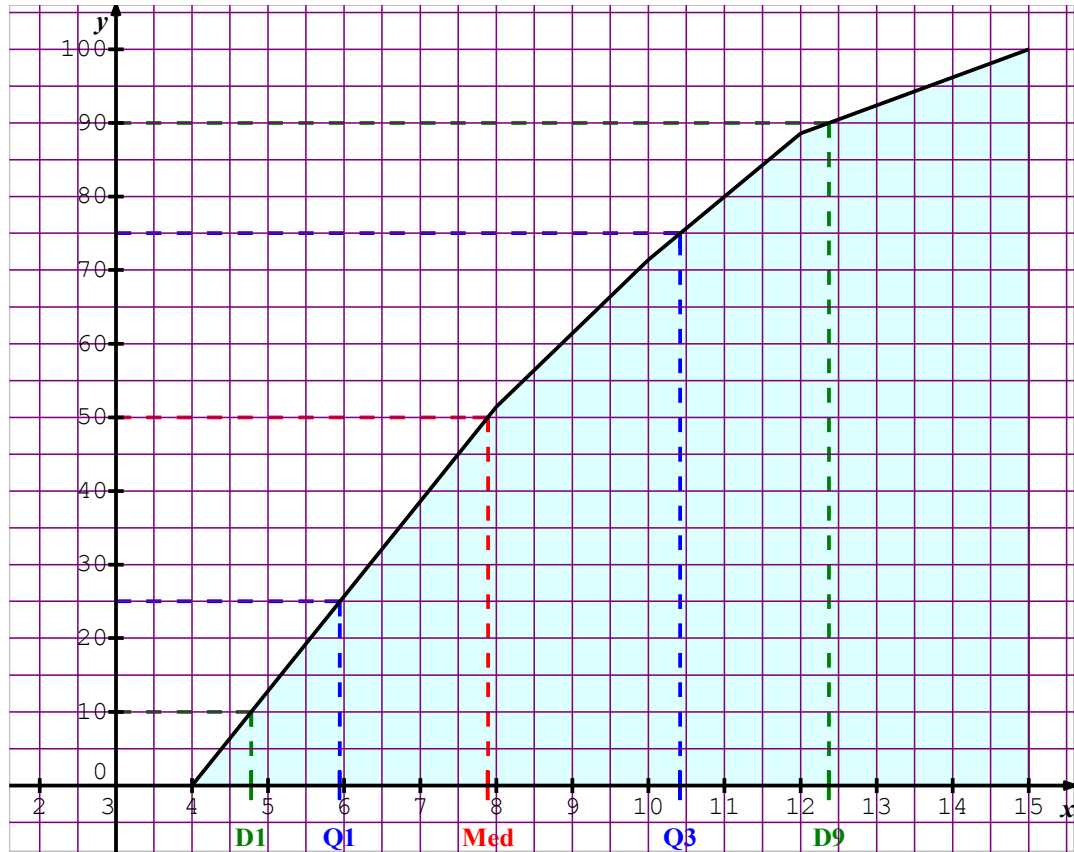
valeur)

Diagramme en bâtons**Regroupement par classes**

notes	[4; 8[[8; 10[[10; 12[[12; 15]
effectifs	18	7	6	4
effectifs cumulés	18	25	31	35
fréquences cumulées	$\frac{18}{35}$	$\frac{25}{35}$	$\frac{31}{35}$	1

et histogramme

et diagramme des fréquences cumulées



Voir livre : pages 244 et suivantes ...