STATISTIQUES COURS

Statistiques

DURÉE ESTIMÉE

25 minutes

OBJECTIFS DU CHAPITRE

- Calculer la moyenne d'une série statistique
- Déterminer la médiane d'une série statistique
- Déterminer les quartiles d'une série statistique
- Calculer des fréquences et construire un diagramme circulaire

I - Exemple et vocabulaire

On interroge les 25 élèves d'un club sportif afin de connaître leurs âges.

Voici leurs réponses, triées par ordre croissant :

11 ; 11 ; 12 ; 12 ; 12 ; 12 ; 13 ; 13 ; 13 ; 13 ; 14 ; 14 ; 14 ;
14 ; 14 ; 14 ; 15 ; 15 ; 16 ; 16 ; 16 ; 17 ; 17 ; 17 ; 17.

- L'ensemble de ces résultats forme une **série statistique**.
- La **population** étudiée est l'ensemble des élèves du club sportif.
- Le **caractère** étudié est l'âge des élèves.
- Dans notre exemple, ce caractère peut prendre sept **valeurs** distinctes : 11 ; 12 ; 13 ; 14 ; 15 ; 16 ; 17.
- Pour chacune de ces valeurs, l'**effectif** correspond au nombre de fois où la valeur a été obtenue.
  Par exemple, l'effectif de la valeur 11 est 2, l'effectif de la valeur 12 est 4, etc.
- L'**effectif total** est le nombre d'élèves du club.
  Ici, l'effectif total est 25.

Pour présenter les résultats de manière plus pratique, on utilise fréquemment un tableau des effectifs :

âges	11	12	13	14	15	16	17
effectifs	2	4	4	6	2	3	4

II - Fréquences

Définition

La **fréquence** d'une valeur s'obtient en divisant l'effectif de cette valeur par l'effectif total :

$$\text{fréquence} = \frac{\text{effectif}}{\text{effectif total}}$$

Remarque

Les fréquences peuvent être exprimées sous forme fractionnaire, sous forme décimale ou sous forme de pourcentage.

Exemple

Si l'on reprend l'exemple du paragraphe I, la fréquence des élèves âgés de 11 ans est :

$$f = \frac{2}{25} = 0{,}08 = 8\%$$

L'ensemble des fréquences de cet exemple peut être présenté dans un tableau :

âges	11	12	13	14	15	16	17
fréquences	0,08	0,16	0,16	0,24	0,08	0,12	0,16
fréquences en %	8%	16%	16%	24%	8%	12%	16%

Propriété

La somme de toutes les fréquences est égale à 1 (c'est-à-dire 100%).

III - Moyenne

Définition

La **moyenne** d'une série statistique s'obtient en divisant la somme de toutes les valeurs de la série par l'effectif total.

Si l'on note $x_1, x_2, x_3, \cdots, x_N$ les valeurs de la série et N l'effectif total, la moyenne $\bar{x}$ vaut :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_N}{N}$$

Exemple

Sonia a obtenu les notes suivantes à ses contrôles de mathématiques :

14 ; 9 ; 12 ; 13 ; 12 ; 15

L'effectif total est de 6 notes. La moyenne vaut :

$$\bar{x} = \frac{14 + 9 + 12 + 13 + 12 + 15}{6} = 12{,}5$$

Moyenne pondérée

Lorsque les valeurs sont regroupées dans un tableau d'effectifs, on peut calculer la moyenne en multipliant chaque valeur par son effectif, puis en divisant la somme obtenue par l'effectif total.

Exemple

Si l'on reprend la série statistique de la partie I, le tableau des effectifs montre que l'âge de 11 ans est présent 2 fois, l'âge de 12 ans : 4 fois, l'âge de 13 ans : 4 fois, etc.

La moyenne vaut alors :

$$\bar{x} = \frac{11 \times 2 + 12 \times 4 + 13 \times 4 + 14 \times 6 + 15 \times 2 + 16 \times 3 + 17 \times 4}{25} = 14{,}08$$

IV - Médiane

Définition

On s'intéresse à une série statistique dont on a classé les valeurs par ordre croissant.
La **médiane** est la valeur qui partage cette série en deux groupes de même effectif.

Déterminer la médiane

On range les N valeurs par ordre croissant.

- Si N est **impair** : la médiane est la valeur située en position $\frac{N+1}{2}$.
- Si N est **pair** : la médiane est la moyenne des deux valeurs situées en positions $\frac{N}{2}$ et $\frac{N}{2}+1$.

Effectif total impair

Chiara a obtenu les cinq notes suivantes :

15 ; 12 ; 7 ; 13 ; 17.

On ordonne ces notes :

7 ; 12 ; **13** ; 15 ; 17.

L'effectif total est 5 (impair). La note médiane est en position $\frac{5+1}{2}=3$, c'est-à-dire la 3e valeur.
La médiane est **13** (il y a deux notes inférieures et deux notes supérieures).

Effectif total pair — valeurs centrales égales

Luc a obtenu les six notes suivantes (déjà triées par ordre croissant) :

11 ; 12 ; **14 ; 14** ; 15 ; 17.

L'effectif total est 6 (pair). La médiane est la moyenne des 3e et 4e valeurs : $\frac{14+14}{2}=14$.

Effectif total pair — valeurs centrales différentes

Sacha a obtenu les six notes suivantes (triées par ordre croissant) :

7 ; 12 ; **12 ; 13** ; 15 ; 16.

La médiane est la moyenne des 3e et 4e valeurs : $\frac{12+13}{2}=12{,}5$.

V - Étendue d'une série statistique

Définition

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

Exemple

Si l'on reprend l'exemple de la partie I :

âges	11	12	13	14	15	16	17
effectifs	2	4	4	6	2	3	4

L'étendue est :

$$17-11=6$$

Cela correspond à la différence d'âge entre l'élève le plus âgé et l'élève le plus jeune.

Remarque

L'étendue mesure la **dispersion** d'une série statistique. Par exemple, si l'on considère les notes d'un élève, une étendue élevée signifiera que ses résultats sont irréguliers.

VI - Quartiles

Définition

On considère une série statistique dont les valeurs sont rangées par ordre croissant.

- Le **premier quartile** Q_1 est la plus petite valeur de la série telle qu'au moins 25% des valeurs lui sont inférieures ou égales.
- Le **troisième quartile** Q_3 est la plus petite valeur de la série telle qu'au moins 75% des valeurs lui sont inférieures ou égales.

Déterminer les quartiles

On range les N valeurs par ordre croissant.

- On calcule $\frac{N}{4}$. On arrondit à l'entier supérieur si nécessaire : Q_1 est la valeur située à ce rang.
- On calcule $\frac{3 \times N}{4}$. On arrondit à l'entier supérieur si nécessaire : Q_3 est la valeur située à ce rang.

Exemple

On reprend la série des âges de la partie I (25 valeurs rangées par ordre croissant) :

11 ; 11 ; 12 ; 12 ; 12 ; 12 ; **13** ; 13 ; 13 ; 13 ; 14 ; 14 ; 14 ; 14 ; 14 ; 14 ; 15 ; 15 ; **16** ; 16 ; 16 ; 17 ; 17 ; 17 ; 17.

Pour le premier quartile : $\frac{25}{4} = 6{,}25$, on arrondit à 7.

La 7e valeur est 13, donc $Q_1 = 13$.

Pour le troisième quartile : $\frac{3 \times 25}{4} = 18{,}75$, on arrondit à 19.

La 19e valeur est 16, donc $Q_3 = 16$.

Au moins 25% des élèves ont 13 ans ou moins, et au moins 75% ont 16 ans ou moins.

Remarque

La médiane et les quartiles sont des **indicateurs de position**. Ils permettent de résumer la répartition des données d'une série statistique.

VII - Représentations graphiques

Diagramme en bâtons

Un **diagramme en bâtons** représente chaque valeur d'une série par un bâton vertical dont la hauteur est proportionnelle à l'effectif (ou à la fréquence) de cette valeur.

Exemple

Voici le diagramme en bâtons des âges des 25 élèves du club sportif (partie I) :

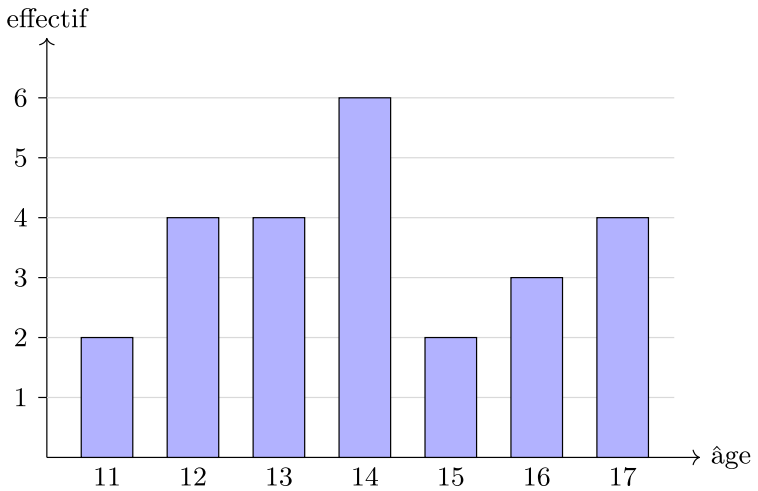


Diagramme circulaire

Un **diagramme circulaire** représente les fréquences d'une série par des secteurs angulaires. Pour une valeur de fréquence f (en pourcentage), la mesure de l'angle correspondant est :

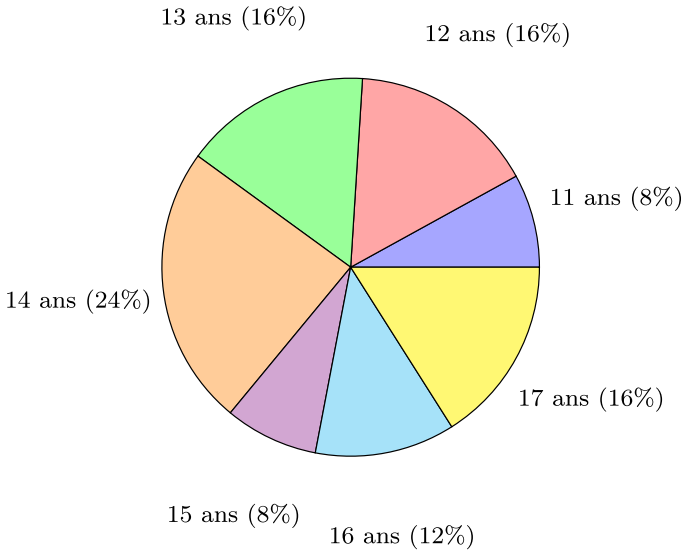
$$\text{angle} = \frac{f}{100} \times 360^\circ$$

Exemple

Si l'on reprend l'exemple de la partie I, l'âge de 14 ans a une fréquence de 24%. L'angle correspondant dans un diagramme circulaire est :

$\frac{24}{100} \times 360 = 86{,}4^\circ$

En procédant de même pour chaque âge, on obtient le diagramme circulaire suivant :



Histogramme

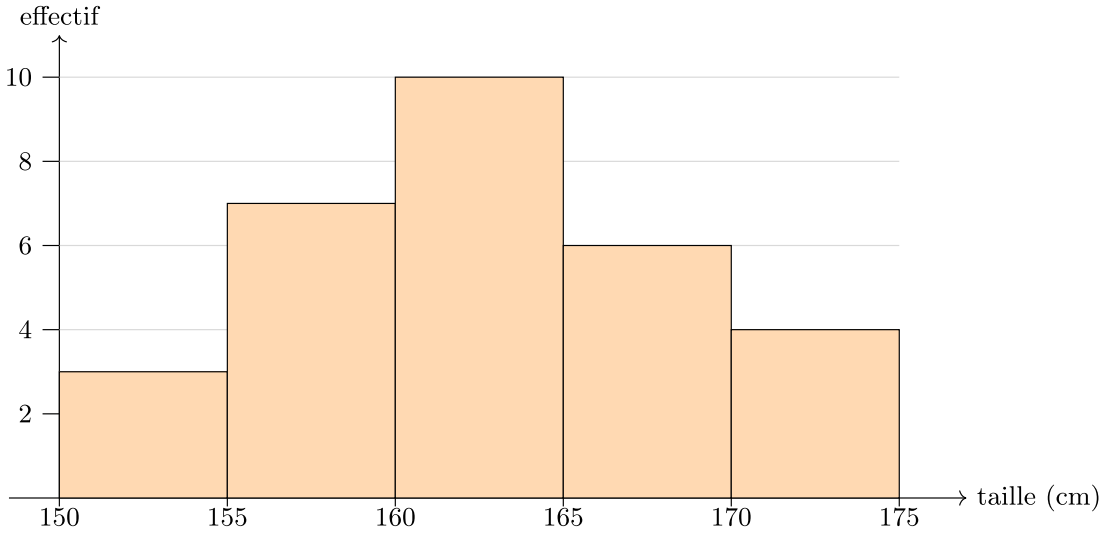
Lorsque les données sont regroupées en **classes** (intervalles de même amplitude), on les représente par un **histogramme** : des rectangles accolés dont la base correspond à la classe et dont la hauteur est proportionnelle à l'effectif de cette classe.

Exemple

On a relevé les tailles (en cm) des élèves d'une classe de 3e. Les résultats sont regroupés dans le tableau suivant :

Taille (en cm)	[150 ; 155[	[155 ; 160[	[160 ; 165[	[165 ; 170[	[170 ; 175[
effectif	3	7	10	6	4

On représente cette série par un histogramme dans lequel chaque rectangle correspond à une classe. Les rectangles sont accolés et ont tous la même largeur car les classes sont de même amplitude (5 cm).



Remarque

Contrairement au diagramme en bâtons, les rectangles d'un histogramme sont **accolés** (pas d'espace entre eux), car les classes se succèdent sans interruption.

Les questions essentielles

1. Comment calculer la moyenne d'une série statistique ?

On additionne toutes les valeurs de la série puis on divise par l'effectif total. Lorsque les données sont regroupées dans un tableau d'effectifs, on multiplie chaque valeur par son effectif avant de diviser par l'effectif total.

Voir la fiche méthode : Calculer la moyenne d'une série statistique

2. Comment déterminer la médiane d'une série statistique ?

On range les valeurs par ordre croissant. Si l'effectif total est impair, la médiane est la valeur centrale. S'il est pair, c'est la moyenne des deux valeurs centrales.

Voir la fiche méthode : Déterminer la médiane d'une série statistique

3. Comment déterminer les quartiles d'une série statistique ?

On range les valeurs par ordre croissant. Le premier quartile est la valeur au rang correspondant au quart de l'effectif total (arrondi à l'entier supérieur). Le troisième quartile est la valeur au rang correspondant aux trois quarts de l'effectif total.

Voir la fiche méthode : Déterminer les quartiles d'une série statistique

4. Comment calculer des fréquences et construire un diagramme

circulaire ?

La fréquence d'une valeur se calcule en divisant son effectif par l'effectif total. Pour le diagramme circulaire, on convertit chaque fréquence en angle en multipliant par 360°.

Voir la fiche méthode : Calculer des fréquences et construire un diagramme circulaire