Chapitre 14 : Statistiques.

I – Vocabulaire.

- Lorsque l'on mène une enquête, on effectue un relevé statistique.

On s'intéresse à une population d'individus (élèves d'une classe, pays de l'Union Européenne, animaux d'une région, …) et on en étudie une propriété commune appelée un caractère (taille des élèves d'une classe, langue officielle des pays de l'Union Européenne, régime alimentaire des animaux d'une région, …).

- Un caractère peut prendre plusieurs valeurs.

Par exemple, herbivore, carnivore, frugivore … sont des valeurs possibles du caractère « régime alimentaire » de la population « animaux d'une région ».

- Un caractère peut être :
  - quantitatif comme par exemple :

taille des élèves d'une classe ; poids des joueurs d'une équipe de rugby ; nombre de frères et sœurs des élèves d'une classe ; …

  - qualitatif comme par exemple :

le sport préféré par les élèves d'une classe ; langue officielle des pays de l'Union Européenne ; …

Exemple :
Voici les notes obtenues par les élèves d'une classe de 3ème lors d'un devoir (sur 10) :
6 ; 7 ; 2 ; 4 ; 7 ; 4 ; 10 ; 7 ; 4 ; 4 ; 10 ; 2 ; 5 ; 5 ; 4 ; 6 ; 6 ; 7 ; 6 ; 7.

- Ces notes constituent un relevé statistique.
- Population étudiée : les élèves d'une classe de 3ème.
- Caractère étudié : la note d'un devoir. Ce caractère est quantitatif.
- Valeurs du caractère : les six notes différentes obtenues : 2 ; 4 ; 5 ; 6 ; 7 ; 10.

II – Effectifs et fréquences.

Définitions :
- L'effectif d'une valeur est le nombre de fois où cette valeur apparaît.
- L'effectif total est égal à la somme des effectifs de chaque valeur.
- La fréquence d'une valeur est le quotient de son effectif pas l'effectif total :

fréquence d'une valeur = $\dfrac{effectif\ de\ la\ valeur}{effectif\ total}$.

Remarques :
- Une fréquence peut être exprimée par une fraction, un nombre décimal ou un pourcentage. La fréquence en pourcentage est obtenue en multipliant la fréquence par 100.
- Une fréquence est un nombre compris entre 0 et 1, et la somme de toutes les fréquences est égale à 1.

Exemple : On reprend l'exemple du paragraphe I et on construit un tableau d'effectifs et de fréquences pour une meilleure lecture des données.

Notes	2	4	5	6	7	10	Total
Effectifs	**2**	**5**	**2**	**4**	**5**	**2**	**20**
Fréquences	$\frac{2}{20}$ = 0,1	$\frac{5}{20}$ = 0,25	$\frac{2}{20}$ = 0,1	$\frac{4}{20}$ = 0,2	$\frac{5}{20}$ = 0,25	$\frac{2}{20}$ = 0,1	**1**
Fréquences (en %)	**10**	**25**	**10**	**20**	**25**	**10**	**100**

III – Caractéristiques de position.

1 – Moyenne d'une série statistique.

Définition : La moyenne d'une série statistique est égale à : $\frac{somme\ des\ valeurs\ du\ caractère}{effectif\ total}$.

Remarque : Lorsque la série statistique est donnée par un tableau d'effectifs, la moyenne est égale à : $\frac{somme\ des\ produits\ des\ valeurs\ par\ leurs\ effectifs}{effectif\ total}$.

Exemple : En reprenant l'exemple du paragraphe I et en notant m la moyenne, on a donc :
m = $\frac{2 \times 2 + 4 \times 5 + 5 \times 2 + 6 \times 4 + 7 \times 5 + 10 \times 2}{20} = \frac{4 + 20 + 10 + 24 + 35 + 20}{20} = \frac{113}{20}$ = 5,65.
La moyenne de cette classe de $3^{ème}$ est donc 5,65/10.

Remarque : La moyenne est une caractéristique de position.

2 – Médiane d'une série statistique.

a) Définition.

Définition : Les valeurs du caractère étant rangées dans l'ordre croissant, la médiane M d'une série statistique est la valeur du caractère qui partage la série en deux groupes de même effectif :

- au moins 50 % des valeurs lui sont inférieures ou égales ;
- au moins 50 % des valeurs lui sont supérieures ou égales.

Remarque : La médiane est une caractéristique de position.

b) Déterminer la médiane d'une série statistique.

Exemple 1 : L'effectif total est impair
Voici les notes obtenues au premier trimestre par un élève :
6 ; 7 ; 8 ; 10 ; 11 ; 13 ; 14 ; 15 ; 15 ; 15 ; 16.
L'effectif total est 11. Donc la valeur centrale est la $6^{ème}$ note.
La médiane de cette série est donc 13.
Il y a 6 valeurs inférieures ou égales à la médiane et 6 valeurs supérieures ou égales à la médiane.

Exemple 2 : L'effectif total est pair
Voici les notes obtenues au deuxième trimestre par ce même élève :
6 ; 7 ; 8 ; 8 ; 11 ; 12 ; 12 ; 13 ; 15 ; 15.
L'effectif total est 10. Donc les valeurs centrales sont les $5^{ème}$ et $6^{ème}$ notes.
La médiane de cette série est donc $\frac{11 + 12}{2}$ = 11,5.
Il y a 5 valeurs inférieures ou égales à la médiane et 5 valeurs supérieures ou égales à la médiane.

c) Déterminer la médiane d'une série statistique à partir des effectifs cumulés.

Exemple : On reprend l'exemple du paragraphe I et on construit un tableau d'effectifs et d'effectifs cumulés.

Notes	**2**	**4**	**5**	**6**	**7**	**10**
Effectifs	**2**	**5**	**2**	**4**	**5**	**2**
Effectifs cumulés	**2**	**7**	**9**	**13**	**18**	**20**

L'effectif total est 20. Donc les valeurs centrales sont les 10ème et 11ème notes.
Par lecture du tableau, on peut dire que la 10ème note est 6 et que la 11ème note est 6.
La médiane de cette série est donc $\frac{6+6}{2} = 6$.

3 – Quartiles d'une série statistique.

a) Définition.

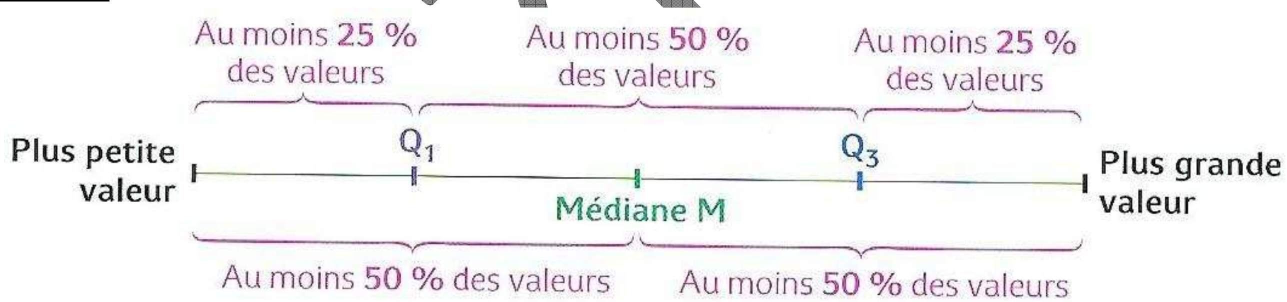
Définition : Les valeurs d'une série étant rangées dans l'ordre croissant :

- le premier quartile Q_1 d'une série statistique est la plus petite valeur du caractère telle qu'au moins 25 % (le quart) des valeurs lui sont inférieures ou égales ;
- le troisième quartile Q_3 d'une série statistique est la plus petite valeur du caractère telle qu'au moins 75 % (les trois quarts) des valeurs lui sont supérieures ou égales.

Remarques :

- Les quartiles sont des caractéristiques de position.
- Le deuxième quartile n'est autre que la médiane de la série.
- Les premier et troisième quartiles correspondent aux médianes des deux demi-séries déterminées par la médiane.

Résumé :



b) Déterminer les quartiles d'une série statistique.

Exemple 1 : L'effectif total est divisible par 4
On considère la série de valeurs : 4 ; 5 ; 8 ; 12 ; 13 ; 15 ; 17 ; 20.
L'effectif total de cette série est 8, qui est divisible par 4. La médiane est 12,5.
$\frac{8}{4} = 2$. Donc Q_1 est la 2ème valeur de la série, c'est-à-dire 5.
$\frac{3}{4} \times 8 = 6$. Donc Q_3 est la 6ème valeur de la série, c'est-à-dire 15.

Exemple 2 : L'effectif total n'est pas divisible par 4
On considère la série de valeurs : 8 ; 10 ; 11 ; 13 ; 14 ; 16 ; 17 ; 18 ; 20 ; 25 ; 30.
L'effectif total de cette série est 11, qui n'est pas divisible par 4. La médiane est 16.
$\frac{11}{4} = 2,75$. Le plus petit entier supérieur à 2,75 est 3.
Donc Q_1 est la 3ème valeur de la série, c'est-à-dire 11.
$\frac{3}{4} \times 11 = 8,25$. Le plus petit entier supérieur à 8,25 est 9.

Donc Q_3 est la 9ème valeur de la série, c'est-à-dire 20.

IV – Caractéristiques de dispersion.

1 – Étendue d'une série statistique.

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

Remarques :

- L'étendue est une caractéristique de dispersion.
- Moins l'étendue d'une série statistique est grande, moins les valeurs sont dispersées. Elles sont alors regroupées autour de la moyenne et de la médiane.

Exemple : Les élèves d'une classe ont effectué deux devoirs dont les notes ordonnées sont les suivantes :
Devoir 1 : 1 ; 4 ; 5 ; 5 ; 6 ; 8 ; 9 ; 11 ; 11 ; 11 ; 12 ; 12 ; 15 ; 15 ; 16 ; 19.
Devoir 2 : 4 ; 5 ; 6 ; 6 ; 6 ; 7 ; 9 ; 10 ; 12 ; 12 ; 13 ; 13 ; 13 ; 14 ; 15 ; 15.
Ces deux séries de notes ont la même moyenne 10 et la même médiane 11, cependant elles n'ont pas le même « profil ».
En effet, l'étendue des notes du devoir 1 est : 19 – 1 = 18 et l'étendue des notes du devoir 2 est : 15 – 4 = 11.
Donc les notes du devoir 2 sont moins dispersées que celles du devoir 1.

2 – Écart interquartile.

Définition : L'écart interquartile est la différence entre le troisième quartile et le premier quartile.

Remarque : L'écart interquartile est une caractéristique de dispersion, il correspond à l'étendue une fois que l'on a retiré les 25 % des valeurs les plus faibles et les 25 % des valeurs les plus fortes.

Exemple : On reprend l'exemple du paragraphe précédent et on détermine les premier et troisième quartiles ainsi que l'écart interquartile pour chaque devoir.
Devoir 1 : Q_1 = 5 et Q_3 = 12, donc l'écart interquartile est 7.
Devoir 2 : Q_1 = 6 et Q_3 = 13, donc l'écart interquartile est 7.