Statistiques descriptives

I Caractéristiques de position d'une série statistique

1) Série statistique

Voici les séries de notes obtenues par 3 élèves :
Paul : 4 ; 6 ; 18 ; 7 ; 17 ; 12 ; 12 ; 18
Mickael : 13 ; 13 ; 12 ; 10 ; 12 ; 3 ; 14 ; 12 ; 14 ; 15
Maryse : 15 ; 9 ; 14 ; 13 ; 10 ; 12 ; 12 ; 11 ; 10

2) Moyenne - Rappel

Définition 1
On note $\bar{x}$ la moyenne d'une série statistique.
La moyenne d'une série statistique est le quotient de la somme des valeurs de la série, par l'effectif total de la série.

Dans l'exemple précédent,
$\bar{x}_{Paul} =$

$\bar{x}_{Mickael} =$

$\bar{x}_{Maryse} =$

3) Médiane - Rappel

Pour déterminer la médiane d'une série statistique, il faut ranger les valeurs de la série par ordre croissant.

Définition 2
On note M_e la médiane d'une série statistique.
La médiane d'une série statistique est la valeur telle qu'au moins 50% des valeurs de la série sont inférieures ou égale à M_e et l'autre moitié des valeurs supérieures ou égales à M_e

Dans l'exemple précédent,

M_e Paul :

M_e Mickael :

M_e Maryse :

4) Quartiles

Pour déterminer les quartiles d'une série statistique, il faut ranger les valeurs de la série par ordre croissant.

Définition 3

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

Dans l'exemple précédent,

Paul :

Mickael :

Maryse :

II Caractéristiques de dispersion d'une série statistique

1) Etendue - Rappel

Définition 4
On note e l'entendue d'une série statistique.
L'étendue d'une série statistique est la différence entre la plus grande valeur et la plus petite valeur de la série.
$e = x_{max} - x_{max}$

Dans l'exemple précédent,

$e_{Paul} =$

$e_{Mickael} =$

$e_{Maryse} =$

2) Ecart interquartile

Définition 5
L'écart interquartile noté E_Q d'une série statistique est la différence entre le troisième quartile et le premier quartile.
$E_Q = Q_3 - Q_1$

Dans l'exemple précédent,

$E_Q Paul =$

$E_Q Mickael =$

$E_Q Maryse =$

3) Interprétations

$\bar{x}_{Paul} =$ $M_e(Paul) =$ $e_{Paul} =$ $Q_1(Paul) =$ $Q_3(Paul) =$ $E_Q(paul) =$

$\bar{x}_{Mickael} =$ $M_e(Mickael) =$ $e_{Mickael} =$ $Q_1(Mickael) =$ $Q_3(Mickael) =$ $E_Q(Mickael) =$

$\bar{x}_{Maryse} =$ $M_e(Maryse) =$ $e_{Maryse} =$ $Q_1(Maryse) =$ $Q_3(Maryse) =$ $E_Q(Maryse) =$

.

III Cas de pondération d'une série statistique

1) Série statistique

Taille des élèves d'une classe de seconde du lycée de Chateaurenard (en cm) :
174 - 160 - 161 - 166 - 177 - 172 - 157 - 175 - 162 - 169 - 160 - 165 - 170 - 152 - 168 - 156 - 163 - 167 - 169 - 158 - 164 - 151 - 162 - 166 - 156 - 165 - 179

2) Regroupement par classe

On peut regrouper les données de cette séries par classes (intervalles) de longueur 5 cm, et calculer les fréquences (arrondies au centième) :

Tailles	[150 ; 155[	[155 ; 160[	[160 ; 165[	[165 ; 170[	[170 ; 175[	[175 ; 180[
Effectifs						
Fréquence						

3) Moyenne pondérée

Définition 6
La moyenne d'une série statistique dont les valeurs sont $x_1, x_2, ..., x_k$, et les effectifs correspondants $n_1, n_2, ..., n_k$ est égale à : $\bar{x} = \frac{n_1x_1+n_2x_2+...+n_kx_k}{n_{total}}$

Remarque
Lorsque les données sont regroupés en classes, pour calculer la moyenne on prend le milieu de chaque classe.

Dans l'exemple précédent :

4) Linéarité de la moyenne

Propriété 1
Si une série de valeurs x_i à pour moyenne $\bar{x}$, alors la série de valeur $ax_i + b$ avec a et b réels a pour moyenne $a\bar{x} + b$

Exemple 1
.

x_i	4	7	-2	Moyenne
$2x_i - 5$				

5) Mediane et quartiles

Quand les valeurs sont regroupées par classes, pour calculer la médiane et les quartiles, on rajoute une ligne dans le tableau : les effectifs cumulés.

Pour la médiane :
On effectue le calcul suivant: $\frac{EffTotal}{2}$.
* Si on trouve un nombre entier n, on regarde grâce à la ligne des effectifs cumulés la "n-ieme valeur" et la "n+1 eme valeur" et on fait la moyenne des 2.
* Si on trouve un nombre à virgule, on arrondi à l'entier supérieur n et on regarde grâce à la ligne des effectifs cumulés la "n-ieme " valeur

Pour les quartiles :
* Pour Q_1 on effectue le calcul $\frac{EffTotal}{4}$ et on arrondi si besoin à l'entier supérieur n. On regarde grâce à la ligne des effectifs cumulés la "n-ieme " valeur.
*Pour Q_3 on effectue le calcul $\frac{EffTotal}{4} \times 3$ et on arrondi si besoin à l'entier supérieur n. On regarde grâce à la ligne des effectifs cumulés la "n-ieme " valeur.

Dans l'exemple précédent :

Tailles	[150 ; 155[	[155 ; 160[	[160 ; 165[	[165 ; 170[	[170 ; 175[	[175 ; 180[	TOTAL
Effectifs							
ECC							

Calcul de Med :

Calcul de Q_1 :

Calcul de Q_3 :

6) Variance, écart type

Définition 7

1. *La variance V d'une série statistique de moyenne $\bar{x}$, dont les valeurs du caractère sont $x_1, x_2, ...x_k$ et les effectifs correspondants sont $n_1, n_2, ...n_k$: est égale à :* $V = \frac{n_1 \times (x_1 - \bar{x})^2 + n_2 \times (x_2 - \bar{x})^2 + ... + n_k \times (x_k - \bar{x})^2}{n_{total}}$
2. *L'écart-type σ d'une série statistique de variance V est égale à* $\sigma = \sqrt{V}$

Dans l'exemple précédent :