CHAPITRE 10 : Statistiques descriptives

[1 Moyenne 2](#_Toc96196814)

[1.1 Moyenne et moyenne pondérée 2](#_Toc96196815)

[1.2 Linéarité de la moyenne 2](#_Toc96196816)

[2 Variance et écart-type 2](#_Toc96196817)

[2.1 Variance 2](#_Toc96196818)

[2.2 Ecart type 4](#_Toc96196819)

[3 Médiane, quartiles et écart interquartile 5](#_Toc96196820)

[3.1 Quartiles 5](#_Toc96196821)

[3.2 Ecart interquartile 5](#_Toc96196822)

[3.3 Médiane 5](#_Toc96196823)

[3.4 Résumé : comment trouver la médiane M, le 1er quartile Q1 et le 3ème quartile Q3 6](#_Toc96196824)

CHAPITRE 10 : Statistiques descriptives

# Moyenne

## Moyenne et moyenne pondérée

* Soit une série statistique de $p$ valeurs distinctes $x\_{1}, x\_{2}, …. x\_{P}$ et d'effectifs $n\_{1}, n\_{2}, …. n\_{P}$

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | $$x\_{1}$$ | $$x\_{2}$$ |  |  |  |  | $$x\_{p}$$ |  |
| Valeurs $x\_{i}$ |  |  |  |  |  |  |  | TOTAL |
| Effectifs $n\_{i}$ |  |  |  |  |  |  |  | $$n=$$ |

L'effectif total est $n=n\_{1}+n\_{2}+…+n\_{p}$. La moyenne est

$$\overbar{x}=\frac{n\_{1}x\_{1}+n\_{2}x\_{2}+…+n\_{p}x\_{p}}{n}$$

## Linéarité de la moyenne

Si on multiplie chaque valeur par un réel $a$ et qu'on additionne au résultat un réel $b$ alors on obtient une deuxième série $ax\_{1}+b, ax\_{2}+b, …. ax\_{P}+b$. Sa moyenne est $a\overbar{x}+b$.

# Variance et écart-type

## Variance

Deux professeurs M. X et M.Y ont donné les notes suivantes à un examen :

$$X: 7, 12, 11, 13, 8, 14, 15, 15, 14, 7$$

$$Y: 11, 12, 12, 11, 10, 13, 12, 12, 11, 12$$

On souhaite comparer les deux façons de noter. Remplir les tableaux d’effectifs suivants :

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | $$x\_{1}$$ | $$x\_{2}$$ |  |  |  |  | $$x\_{p}$$ |  |
| Valeurs $x\_{i}$ |  |  |  |  |  |  |  | TOTAL |
| Effectifs $n\_{i}$ |  |  |  |  |  |  |  | $$n=$$ |
| $$n\_{i}x\_{i}$$ |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Valeurs $y\_{i}$ |  |  |  |  | TOTAL |
| Effectifs $n\_{i}$ |  |  |  |  | $$n=$$ |
| $$n\_{i}y\_{i}$$ |  |  |  |  |  |

En utilisant la formule du calcul de la moyenne

$$\overbar{x}=\frac{n\_{1}×x\_{1}+n\_{2}×x\_{2}+…+n\_{p}×x\_{p}}{n}$$

calculer la note moyenne $\overbar{x}$ pour le professeur X et la note moyenne $\overbar{y}$ pour le professeur Y.

$$\overbar{x}=\frac{…×…+ …×…+…+ …×…}{…}=… \overbar{y}=\frac{…×…+ …×…+…+ …×…}{…}=…$$

Les deux moyennes sont les mêmes, mais on se rend compte au moyen d’un diagramme en bâtons que les façons de noter des deux professeurs ne sont pas les mêmes.

1

1

x

y

o

5

0

5

5

•



Professeur X

1

1

x

y

o

5

0

5

5

•



Professeur Y

Bien que les deux séries aient la même position, visiblement le professeur Y note de façon moins dispersée. Comment calculer la dispersion ?

On peut calculer l**’étendue**

$$x\_{max}-x\_{min}=$$

$$y\_{max}-y\_{min}=$$

Mais il existe des indicateurs de dispersion qui tiennent compte de toutes les valeurs.

On peut penser à faire la moyenne des écarts de chaque valeur à la moyenne.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | $$x\_{1}$$ | $$x\_{2}$$ |  |  |  |  | $$x\_{p}$$ |  |
| Valeurs $x\_{i}$ |  |  |  |  |  |  |  | TOTAL |
| Effectifs $n\_{i}$ |  |  |  |  |  |  |  |  |
| $$n\_{i}(x\_{i}-\overbar{x})$$ |  |  |  |  |  |  |  |  |

$$\frac{n\_{1}\left(x\_{1}-\overbar{x}\right)+n\_{2}\left(x\_{2}-\overbar{x}\right)+… +n\_{p}\left(x\_{p}-\overbar{x}\right)}{n}=…$$

Les écarts négatifs compensent les écarts positifs.

* **La variance :** c’est la moyenne des carrés des écarts à la moyenne.

$$V=\frac{n\_{1}\left(x\_{1}-\overbar{x}\right)^{2}+n\_{2}\left(x\_{2}-\overbar{x}\right)^{2}+… n\_{p}\left(x\_{p}-\overbar{x}\right)^{2}}{n}$$

Exemple pour le professeur X :

$V\_{X}= …$

Inconvénient de la variance : son unité est celle du caractère étudié au carré.

## Ecart type

C’est la racine carrée de la variance. Son unité est donc la même que celle du caractère étudié. On la note avec la lettre grecque σ (sigma) ou avec la lettre s

 $σ=\sqrt{V}$

Pour le professeur X,

$$σ\_{X}=…$$

*Exercice : calculer la variance V, l'écart type* σ  *pour le professeur Y*

# Médiane, quartiles et écart interquartile

## Quartiles

Il faut faire le tableau des effectifs cumulés croissants

Reprenons les notes du professeur X

X : 7, 12, 11, 13, 8, 14, 15, 15 14, 7

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
| Valeurs *xi* | 7 | 8 | 11 | 12 | 13 | 14 | 15 | TOTAL |
| Effectifs *ni* | 2 | 1 | 1 | 1 | 1 | 2 | 2 |  *n* = 10 |
| Effectifs cumulés croissants | 2 | 3 | 4 | 5 | 6 | 8 | 10 |  |

la 8eme note est dedans

la 3eme note est dedans

* **La médiane** *M* est la valeur du caractère qui correspond à 50 % des observations. **Si l’effectif total *n* est impair**, la médiane *M* est la valeur centrale. **Si l’effectif total *n* est pair**, on convient que la médiane *M* est la moyenne des deux valeurs centrales.
* **Les quartiles** Q1 et Q3 sont les plus petites valeurs du caractère qui correspondent à **au moins** 25% et à au moins 75% des observations

## Ecart interquartile

* **L’écart interquartile** est la différence Q3 – Q1

Exemple avec le professeur X :

## Médiane

Comme *n* = 10 est pair :

$$M=\frac{5^{e}note+6^{e}note}{2} M=\frac{12+13}{2}=12,5$$

* **Premier quartile**

Il correspond à 25% des observations

$$10×\frac{25}{100}=2,5$$

Donc la 3ème note est le premier quartile. $Q\_{1}=8$

* **Troisième quartile**

Il correspond à 75% des observations

$$10×\frac{75}{100}=7,5$$

Donc la 8ème note est le troisième quartile. $Q\_{3}=14$

* L’écart interquartile est $Q\_{3}-Q\_{1}=14-8=6$

L’écart interquartile est **un paramètre de dispersion**.

## Résumé : comment trouver la médiane M, le 1er quartile Q1 et le 3ème quartile Q3

|  |  |  |  |
| --- | --- | --- | --- |
|  | Médiane M | 1er quartile Q1 | 3ème quartile Q3 |
| Série à 9 valeurs$$\frac{9+1}{2}$$ | On fait le tableau des effectifs cumulés croissants. La médiane est la 5ème valeur | $$\frac{25}{100}×9=2,25\rightarrow Q\_{1} est la 3ème valeur$$ | $$\frac{75}{100}×9=6.75\rightarrow Q\_{3} est la 7ème valeur$$ |
| Série à 14 valeurs$$\frac{14}{2}$$ | On fait le tableau des effectifs cumulés croissants. La médiane est la moyenne de la 7ème valeur et de la 8ème valeur | $$\frac{25}{100}×14=3.5\rightarrow Q\_{1} est la 4ème valeur$$ | $$\frac{75}{100}×14=10.5\rightarrow Q\_{3} est la 11ème valeur$$ |
| Série à 507 valeurs$$\frac{507+1}{2}$$ | On fait le tableau des effectifs cumulés croissants. La médiane est la 254ème valeur | $$\frac{25}{100}×507=126.75\rightarrow Q\_{1} est la 127ème valeur$$ | $$\frac{75}{100}×507=380.25\rightarrow Q\_{3} est la 381ème valeur$$ |
| Série à 628 valeurs$$\frac{628}{2}$$ | On fait le tableau des effectifs cumulés croissants. La médiane est la moyenne de la 314ème valeur et de la 315ème valeur | $$\frac{25}{100}×628=157\rightarrow Q\_{1} est la 157ème valeur$$ | $$\frac{75}{100}×628=471\rightarrow Q\_{3} est la 471ème valeur$$ |