

1 Notation

La somme de n nombres numérotés de 1 à n peut s'écrire :

$$x_1 + x_2 + x_3 + x_4 + \cdots + x_{n-1} + x_n$$

mais cette écriture est longue et les pointillés ne sont pas satisfaisants.

On écrira, pour faire moins long et éviter les pointillés, cette somme à l'aide du symbole Sigma :

$$\sum_{i=1}^n x_i$$

Exemples :

1. $\sum_{i=0}^n x_i = x_0 + x_1 + x_2 + \cdots + x_{n-1} + x_n$
2. $\sum_{i=1}^{n-1} i = 1 + 2 + 3 + 4 + \cdots + (n-2) + (n-1)$
3. $\sum_{i=0}^n i^2 = 0^2 + 1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2$

2 Quelques généralités et rappels

On note $(x_i; n_i)_{i \in \mathbb{N}}$ la série statistique ci-dessous :

Rappels : L'effectif n_i est le nombre de fois où apparaît la valeur x_i dans la série.

valeurs x_i	x_1	x_2	x_3	x_4	x_{k-2}	x_{k-1}	x_k
Effectifs n_i	n_1	n_2	n_3	n_4	n_{k-2}	n_{k-1}	n_k

2.1 Effectif total d'une série

Définition :

L'effectif total N de la série statistique est la somme de tous les effectifs ou le nombre de valeurs total dans cette série :

$$N = \sum_{i=1}^k n_i = n_1 + n_2 + n_3 + \cdots + n_{k-1} + n_k$$

2.2 Fréquence d'apparition d'une valeur

Définition :

La fréquence d'apparition d'une valeur x_i est la proportion de cette valeur par rapport à l'effectif total.

$$\text{Fréquence par rapport à 1 : } f_i = \frac{\text{Effectif de la valeur}}{\text{Effectif total}} = \frac{n_i}{N}$$

$$\text{Fréquence par rapport à 100 : } F_i = \frac{100n_i}{N}$$

Propriétés :

$$S_f = \sum_{i=1}^k f_i = f_1 + f_2 + f_3 + \dots + f_{k-1} + f_k = 1$$

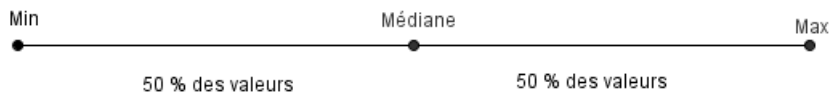
$$S_F = \sum_{i=1}^k F_i = F_1 + F_2 + F_3 + \dots + F_{k-1} + F_k = 100$$

3 Critère de position

3.1 Médiane

Définition :

La médiane d'une série statistique est la valeur qui partage cette série en deux séries de même effectif.



Si M_e est la médiane de la série statistique, alors :

50 % des valeurs de la série sont inférieures ou égales à M_e
 50 % des valeurs de la série sont supérieures ou égales à M_e

Méthode pour trouver la médiane :

Il faut commencer par classer la série dans l'ordre croissant.

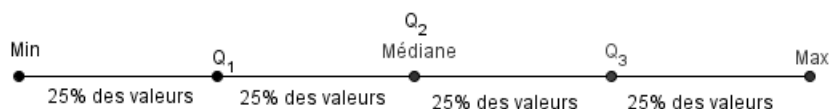
▷ Si N est impair alors la médiane est la $\frac{N+1}{2}$ ième valeur de la série.

▷ Si N est pair alors la médiane est la $\frac{N}{2} + 1$ ième valeur de la série.

3.2 Quartiles

Définition :

Les quartiles d'une série statistique sont les valeurs qui partagent cette série en quatre séries de même effectif.



Si Q_1 est le premier quartile et Q_3 le troisième de la série statistique, alors :

25 % des valeurs de la série sont dans $[Min, Q_1]$

50 % des valeurs de la série sont dans $[Q_1, Q_3]$

25 % des valeurs de la série sont dans $[Q_3, Max]$

Méthode pour trouver les quartiles :

Il faut commencer par classer la série dans l'ordre croissant.

On utilisera une méthode approximative mais qui donnera des résultats significatif pour des séries à grands effectifs. (Autement il suffit de couper en deux les deux séries $[Min, M_e]$ et $[M_e, Max]$)

Calculer $\frac{N}{4}$ et on note a l'entier supérieur à $\frac{N}{4}$.

Calculer $\frac{3N}{4}$ et on note b l'entier supérieur à $\frac{3N}{4}$.

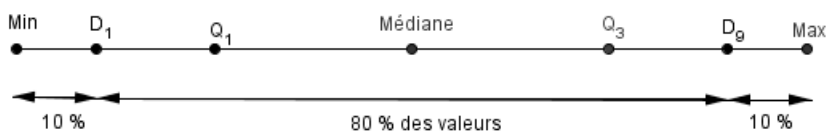
▷ Q_1 est la a ième valeur de la série statistique.

▷ Q_3 est la b ième valeur de la série statistique.

3.3 Déciles

Définition :

Les déciles d'une série statistique sont les valeurs qui partagent cette série en dix séries de même effectif.



Si D_1 est le premier décile et D_9 le neuvième de la série statistique, alors :

10 % des valeurs de la série sont dans $[Min, D_1]$

80 % des valeurs de la série sont dans $[D_1, D_9]$

10 % des valeurs de la série sont dans $[D_9, Max]$

Méthode pour trouver les Déciles :

Il faut commencer par classer la série dans l'ordre croissant.

On utilisera une méthode approximative mais qui donnera des résultats significatif pour des séries à grands effectifs.

Calculer $\frac{N}{10}$ et on note a l'entier supérieur à $\frac{N}{10}$.

Calculer $\frac{9N}{10}$ et on note b l'entier supérieur à $\frac{9N}{10}$.

▷ D_1 est la a ième valeur de la série statistique.

▷ D_9 est la b ième valeur de la série statistique.

3.4 Moyenne

Définition :

La moyenne arithmétique de la série statistique est le nombre :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k x_i \times n_i = \frac{x_1 n_1 + x_2 n_2 + x_3 n_3 + \dots + x_{k-1} n_{k-1} + x_k n_k}{N}$$

ou

$$\bar{x} = \sum_{i=1}^k x_i \times f_i = x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_k f_k$$

Propriétés de la moyenne :

1. Si \bar{x} est la moyenne d'un groupe d'effectif N_1 et \bar{y} la moyenne d'un groupe d'effectif N_2 alors la moyenne \bar{z} de la série constituée de l'ensemble des deux groupes est :

$$\bar{z} = \frac{N_1 \bar{x} + N_2 \bar{y}}{N_1 + N_2}$$

2. Si \bar{x} est la moyenne d'une série (x_i, n_i) alors la moyenne de la série $(ax_i + b, n_i)$ est :

$$\bar{y} = a\bar{x} + b$$

3. **Moyenne élaguée :**

Quand une valeur aberrante, correspondant à une erreur de mesure ou à une situation exceptionnelle, est présente dans une série, elle influence considérablement la valeur moyenne. Une moyenne calculée après avoir enlevé certaines valeurs est appelée **Moyenne élaguée**.

3.5 Modes

Définition :

Les modes d'une série sont les valeurs ayant le plus grand effectif.

4 Critère de dispersion

4.1 Etendue

Définition :

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

$$Et = Max - Min$$

4.2 Ecart Inter-Quartiles

Définition :

L'écart inter-quartiles est la différence entre Q_3 et Q_1

$$E_Q = Q_3 - Q_1$$

L'intervalle inter-quartiles est l'intervalle entre Q_1 et Q_3

$$I_Q = [Q_1, Q_3]$$

4.3 Ecart Inter-Déciles

Définition :

L'écart inter-déciles est la différence entre D_9 et D_1

$$E_D = D_9 - D_1$$

L'intervalle inter-déciles est l'intervalle entre D_1 et D_9

$$I_D = [D_1, D_9]$$