

Statistique à une variable

1 Vocabulaire

Définition 1 :

Une étude statistique porte sur un ensemble (de personnes, d'animaux, d'objets, ...) appelé population.

Chaque élément de la population est un individu.

L'aspect étudié est nommé caractère ou variable.

Les résultats obtenus après observation donnent une série statistique.

Lorsque les variables prennent des valeurs numériques (exemple : notes, tailles, âges, ...), les variables sont dites quantitatives (si la variable prend n'importe quelle valeur dans un intervalle donné, la variable est dite continue , si elle prend des valeurs isolées, la variable est dite discrète) Dans le cas contraire, les variables sont dites qualitatives (nationalité, couleurs, ...) ; les différentes possibilités du caractère sont appelées modalités (la commune de résidence pour des élèves de première fréquentant un lycée ...).

2 Effectifs et fréquences

On note X une série statistique à caractère quantitatif dans les p valeurs seront notés x_1, x_2, \dots, x_p d'effectifs respectifs n_1, n_2, \dots, n_p

Définition 2 :

L'effectif total, noté N , est le nombre d'individus de la population étudiée. On a donc :

$$N = n_1 + n_2 + \dots + n_p = \sum_{i=1}^p n_i$$

La fréquence, notée f_i , d'une valeur de la variable est le quotient de l'effectif de cette valeur par l'effectif total.

$$f_i = \frac{n_i}{N}$$

L'effectif cumulé croissant d'une valeur x_i (ou fréquence cumulée croissante) est égale à la somme des effectifs (ou fréquences) des valeurs inférieures ou égales à x_i .

3 Caractéristiques de position

On note X une série statistique à caractère quantitatif dans les p valeurs seront notés x_1, x_2, \dots, x_p d'effectifs respectifs n_1, n_2, \dots, n_p

3.1 Mode

Définition 3 :

Le mode est la valeur du caractère qui a le plus grand effectif.

Dans un cas de variable continue, on parle de classe modale, le mode étant alors le centre de la classe.

3.2 Médiane

Définition 4 :

La médiane est la valeur qui sépare la population en deux sous-ensembles de même effectif. C'est la valeur qui correspond à la fréquence cumulée croissante égale à 50 %.

3.3 Moyenne

Définition 5 :

On note X une série statistique à caractère quantitatif dans les p valeurs seront notés x_1, x_2, \dots, x_p d'effectifs respectifs n_1, n_2, \dots, n_p . La moyenne est

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N} = \frac{1}{N} \sum_{i=1}^p n_i x_i$$

4 Caractéristiques de dispersion

4.1 Etendue

Définition 6 :

L'étendue d'une série statistique est la différence entre les valeurs extrêmes : Si x_1, x_2, \dots, x_p sont rangés dans l'ordre croissant, $e = x_p - x_1$

4.2 Quartile

Définition 7 :

L'idée générale est de partager la population en quatre parties de même effectif. Étant donné une série statistique de médiane M dont la liste des valeurs est rangée dans l'ordre croissant (il s'agit de la même liste que celle qu'on utilise pour déterminer la médiane).

En coupant la liste en deux sous-séries de même effectif (Attention : quand l'effectif total est impair, la médiane ne doit pas être incluse dans les sous-séries) :

- On appelle premier quartile le réel noté Q_1 égal à la médiane de la sous-série inférieure.
- On appelle troisième quartile le réel noté Q_3 égal à la médiane de la sous-série supérieure.
- $]Q_1; Q_3[$ est appelé intervalle inter quartile. L'écart inter quartile est égal à $Q_3 - Q_1$.

4.3 Variance, écart-type

Définition 8 :
 On note X une série statistique à caractère quantitatif dans les p valeurs seront notés x_1, x_2, \dots, x_p d'effectifs respectifs n_1, n_2, \dots, n_p /
 La variance est

$$V = \frac{1}{N} \sum_{i=1}^p n_i (x_i - \bar{x})^2 = \frac{1}{N} \sum_{i=1}^p n_i x_i^2 - \bar{x}^2$$

L'écart-type est :

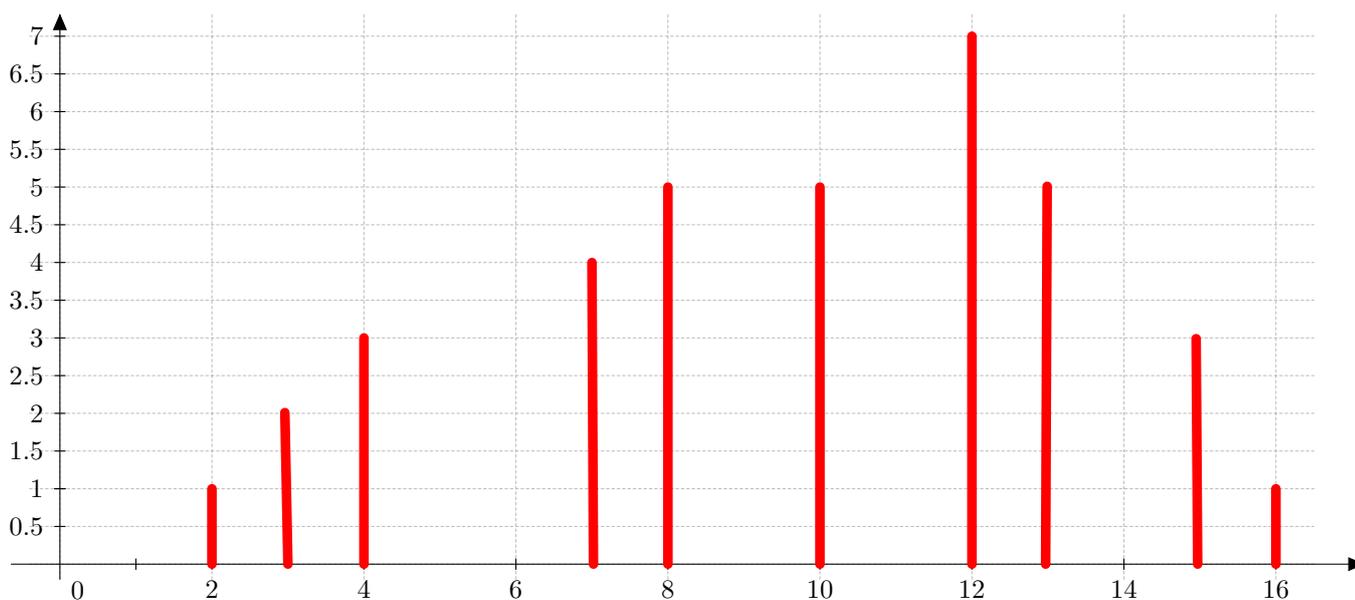
$$\sigma = \sqrt{V}$$

5 Graphisme

5.0.1 Diagramme en bâton

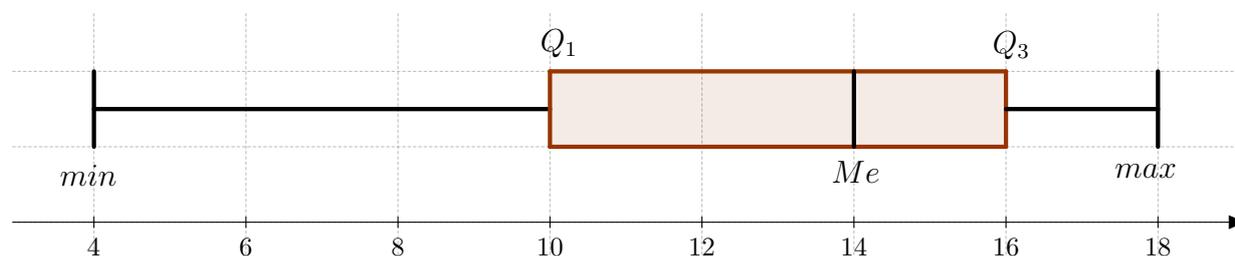
Pour l'exemple suivant :

x_i	2	3	4	7	8	10	12	13	15	16
n_i	1	2	3	4	5	5	7	5	3	1



5.1 Diagramme en boîte

Le diagramme en boîtes d'une série statistique se construit alors de la façon suivante : (les valeurs du caractère sont en abscisse - min et max représentent les valeurs minimales et maximales du caractère)



5.2 Histogramme

Lorsque le caractère étudié est quantitatif et continu, et lorsque les modalités sont regroupées en classes, on peut représenter la série par un histogramme : l'aire de chaque rectangle est alors proportionnelle à l'effectif (ou à la fréquence) associée à chaque classe. Pour l'exemple suivant :

Classe]12; 14]]14; 16]]16; 18]]18; 20]
Effectif	4	6	3	1

