

Comparaison de séries statistiques

1. Vocabulaire

Pour un caractère quantitatif continu, on regroupe les valeurs dans des intervalles appelés *classes*.

- L'**amplitude** d'une classe est la différence entre les bornes de l'intervalle, c'est la taille de l'intervalle.
- Le **centre de classe** est la valeur située au milieu de l'intervalle, on le note x_i .

2. Indicateurs de position

- **Mode, classe modale**

Pour un caractère quantitatif continu :

- la **classe modale** est l'intervalle dont l'effectif est le plus élevé ;
- le **mode** est le centre de la classe modale.

- **Moyenne, médiane**

La **moyenne** est la somme de toutes les valeurs divisée par l'effectif total :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + \dots + n_kx_k}{N}$$

La **médiane** est la valeur du caractère qui coupe la série (dont les valeurs sont rangées dans l'ordre croissant) en deux parties de même effectif : la moitié des individus a une valeur inférieure ou égale à la médiane. Autrement dit, 50 % des individus ont une valeur inférieure ou égale à la médiane.

- **Quartiles**

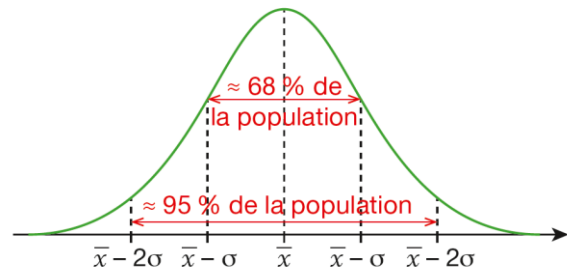
Pour une série dont les valeurs sont rangées dans l'ordre croissant :

- le **1^{er} quartile**, noté Q_1 , est la valeur du caractère qui correspond au 1^{er} quart de la population : un quart des individus ont une valeur inférieure ou égale au premier quartile. Autrement dit, 25 % des individus ont une valeur inférieure ou égale au premier quartile ;
- le **3^e quartile**, noté Q_3 , est la valeur du caractère qui correspond au 3^e quart de la population : trois quarts des individus ont une valeur inférieure ou égale au troisième quartile. Autrement dit, 75 % des individus ont une valeur inférieure ou égale au troisième quartile.

3. Indicateurs de dispersion

- L'**étendue** est la différence entre la plus grande valeur et la plus petite valeur de la série.
- L'**écart type** est un indicateur de dispersion autour de la moyenne : plus il est faible, plus les valeurs sont resserrées autour de la moyenne.

Dans le cas d'une répartition des valeurs de la série en courbe de Gauss, on a le graphique ci-contre.



- L'**écart interquartile** est la différence entre le 1^{er} et le 3^e quartile, c'est-à-dire : $Q_3 - Q_1$.

4. Diagramme en boîte à moustaches

Un **diagramme en boîte à moustaches** est une représentation graphique qui regroupe les informations suivantes pour une série de données : minimum, Q_1 , médiane, Q_3 , maximum.

