

## 0.1 Statistiques

Les statistiques permettent d'étudier un caractère d'une population. Le nombre d'éléments de la population s'appelle l'effectif global (ou l'effectif total) que l'on notera  $N$  dans ce chapitre. Pour une valeur de caractère donnée, l'effectif est le nombre d'éléments correspondant à cette valeur.

### 0.1.1 Tableau d'effectif

Une série statistique est l'ensemble des résultats d'une étude : valeurs du caractère et effectifs correspondants. On range ces valeurs dans un tableau d'effectif où les valeurs du caractère (le nombre d'heures de jeu vidéos) sont rangées en ordre croissant. Une série statistique est l'ensemble des résultats d'une étude : valeurs du caractère et effectifs correspondants. On donne les exemples sur la même série statistique. On interroge un groupe d'élève de lycée :

«A combien d'heures de jeux vidéos jouez-vous du lundi au vendredi hors vacances ?»

Voici leur réponses :

0-1-1-1-2-2-2-2-4-4-5-10-10-10-12-15-15-15-18-18-20-20-20-20-24

Elles ont été regroupées dans un tableau d'effectif ci-dessous.

Nombre d'heures $x_i$	0	1	2	4	5	10	12	15	18	20	24
Effectif $n_i$	1	3	4	2	1	3	1	3	2	4	1
<u>ECC</u>											

Figure 1: Tableau d'effectif avec effectifs cumulés croissants (ECC)

### 0.1.2 Moyenne

On calcule la moyenne pondérée de ses effectifs dite coefficientée. Elle est notée  $\bar{x}$  ou  $m$  et se calcule avec la formule suivante pour  $p$  valeurs différentes du caractère :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N} = \frac{\sum_{n=1}^p n_ix_i}{N}$$

Valeur $x_i$	1	3	5	7	10
Effectif $n_i$	4	2	5	1	7
<u>ECC</u>					

Figure 2: Tableau d'effectif de la série 1

Calculer les paramètres statistiques pour la série 1 et la série «temps de jeux videos».

### 0.1.3 Médiane

La médiane d'une série, dont les valeurs du caractère sont discrètes et rangées dans l'ordre croissant, est la valeur partageant la population en deux groupes de même effectif. Si l'effectif de la série est pair on prendra la moyenne des deux valeurs centrales de la série pour la médiane.

**Propriété :** Au moins 50 % des valeurs de la série sont plus petites que la médiane.

**Déterminez la médiane de la série 1 :**

Déterminez la médiane de la série «temps de jeux vidéos»:

### 0.1.4 Quartiles

1. Le rang du 1er quartile noté  $Q_1$  d'une série de  $N$  valeurs est le plus petit entier supérieur ou égal à  $\frac{N}{4}$ . Il suffit d'identifier la valeur du quartile correspondante avec la ligne des ECC.

**Propriété :** Au moins 25 % des valeurs de la série sont plus petites que  $Q_1$ .

2. Le rang du 3e quartile noté  $Q_3$  est le plus petit entier supérieur ou égal à  $3 \times \frac{N}{4}$ .

**Propriété :** Au moins 75 % des valeurs de la série sont plus petites que  $Q_3$ .

#### Ecart interquartile

- L'écart interquartile est la différence  $Q_3 - Q_1$ . C'est un indicateur de dispersion.
- L'intervalle interquartile est l'intervalle  $[Q_1; Q_3]$ .

**Propriété :** Au moins 50 % des valeurs de la série sont comprises dans l'intervalle interquartile.

Calculez les quartiles et l'écart interquartile de chacune des séries

### 0.1.5 Diagramme en boîte

Un diagramme en boîte résume seulement quelques indicateurs de position du caractère étudié (médiane, quartiles, minimum, maximum ou déciles). Ce diagramme est utilisé principalement pour comparer un même caractère dans deux populations de tailles différentes.

Il s'agit de tracer un rectangle allant du premier quartile au troisième quartile et coupé par la médiane. Il est alors possible de comparer rapidement plusieurs séries.

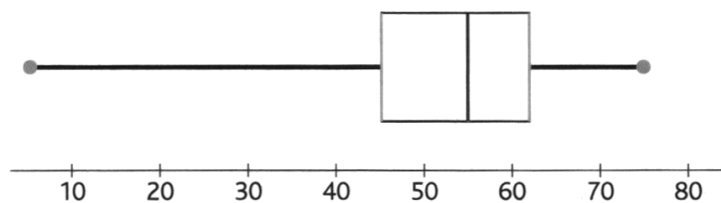


Figure 3:  $Min = 5$ ,  $Q_1 = 45$ ,  $Me = 55$ ,  $Q_3 = 62$ ,  $Max = 75$

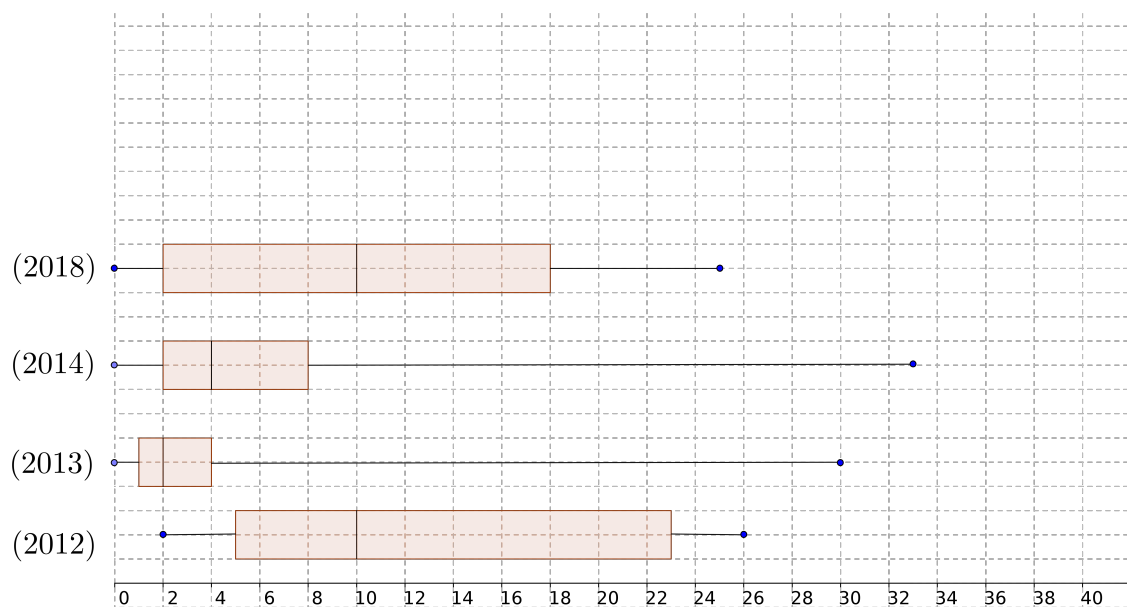
**Exemple :** Compléter avec le diagramme en boîte de votre série *Nombre d'heures de jeux vidéos*.

### 0.1.6 Écart-type

Cet indicateur noté  $\sigma$  s'utilise en couple avec la moyenne qui n'est pas toujours un indicateur pertinent pour résumer une série statistique quand elle est analysée seule. Il est la racine carrée de la variance. La variance est la moyenne arithmétique des écarts à la moyenne élevés au carré :

$$Var(x) = \sigma^2 = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{N} = \sum_{n=1}^p \frac{n_i(x_i - \bar{x})^2}{N}$$

Il suffit alors de prendre la racine carrée.



**Exemple :** Complétez le tableau suivant afin de déterminer l'écart-type de la série 1.

Valeur $x_i$	1	3	5	7	10
Effectif $n_i$	4	2	5	1	7
ECC					
$x_i - m$					
$(x_i - m)^2$					
$ni(x_i - m)^2$					

**Exemple :** Déterminez l'écart-type de la série «jeux vidéos »à l'aide de la calculatrice: