

Statistiques

1 Vocabulaire

On utilise les statistiques pour décrire un ensemble appelé *population* constitué d'*individus* auxquels est associé un *caractère*. On distingue :

- Les *statistiques qualitatives* : le caractère étudié est d'ordre qualitatif et peut prendre différentes *modalités*.
- Les *statistiques quantitatives* : le caractère étudié est d'ordre quantitatif et peut prendre différentes *valeurs*.

Exemple 1. *On étudie la couleur des yeux en France :*

- *population* : la population française.
- *individus* : les français.
- *caractère* : la couleur des yeux.
- *modalités* : bleu, noir, vert, marron...

Il s'agit d'une série statistique qualitative.

Exemple 2. *On étudie les notes d'un Devoir Surveillé en classe de seconde :*

- *population* : la classe de seconde.
- *individus* : les élèves de la classe.
- *caractère* : la note obtenue au Devoir Surveillé.
- *valeurs* : 0 ; 0,5 ; 1 ; 1,5 ... 19,5 ; 20.

Il s'agit d'une série statistique quantitative.

L'*effectif* d'une valeur est le nombre d'individus associés à cette valeur.

La *fréquence* d'une valeur est le quotient de son effectif par l'effectif total.

Exemple. *On considère les notes d'un Devoir Surveillé en classe de seconde :*

<i>Notes x_i</i>	$x_1 = 5$	$x_2 = 7$	$x_3 = 8$	$x_4 = 10$	$x_5 = 11$	$x_6 = 13$	$x_7 = 16$	<i>Total</i>
<i>Effectifs n_i</i>	$n_1 = 3$	$n_2 = 2$	$n_3 = 2$	$n_4 = 2$	$n_5 = 7$	$n_6 = 5$	$n_7 = 3$	$\Sigma n_i = 24$
<i>Fréquences f_i</i>	$f_1 = \frac{3}{24}$	$f_2 = \frac{2}{24}$	$f_3 = \frac{2}{24}$	$f_4 = \frac{2}{24}$	$f_5 = \frac{7}{24}$	$f_6 = \frac{5}{24}$	$f_7 = \frac{3}{24}$	$\Sigma f_i = 1$

On peut regrouper les individus d'une population par *classes* :

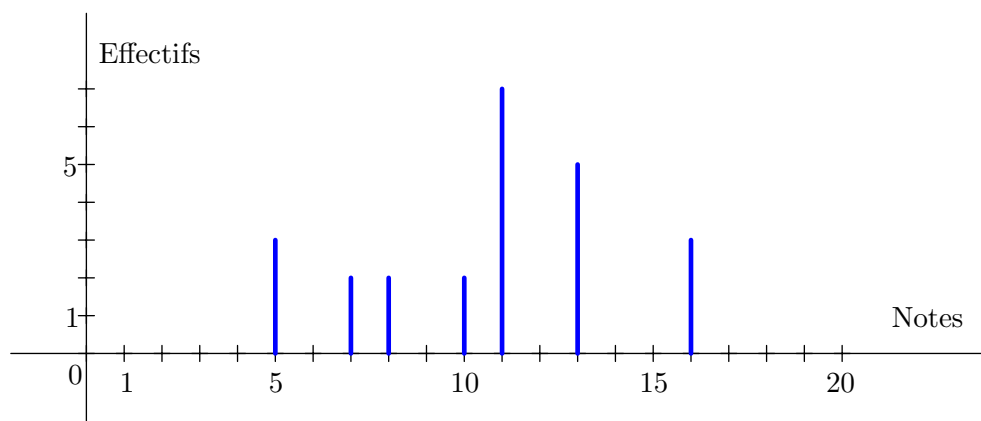
Exemple. *On considère les notes du Devoir Surveillé en classe de seconde :*

<i>Classes</i>	[0; 5]]5; 10]]10; 15]]15; 20]	<i>Total</i>
<i>Effectifs</i>	3	6	12	3	24
<i>Fréquences</i>	12,5%	25%	50%	12,5%	100%

2 Représentations graphiques

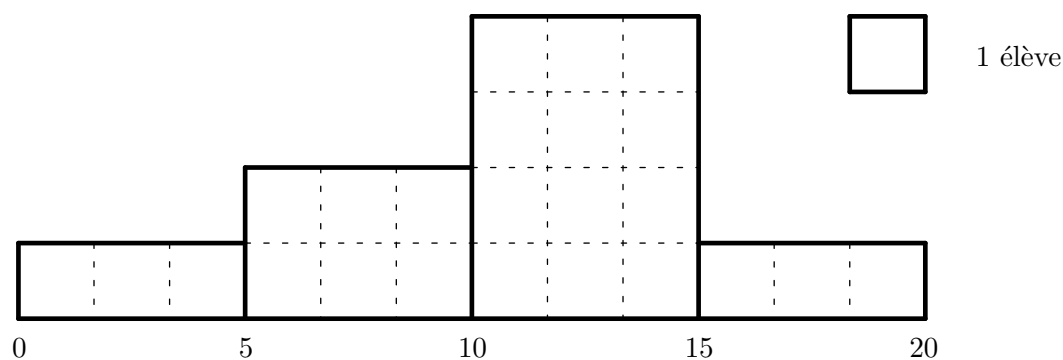
On utilise pour représenter une série statistique un *diagramme en barres*, un *diagramme en bâtons* ou bien un *diagramme circulaire*.

Exemple. *Diagramme en bâtons des notes du Devoir Surveillé en classe de seconde :*



Lorsque les individus sont regroupés par classes, on peut représenter la série statistique par un *histogramme*.

Exemple. *Histogramme des notes du Devoir Surveillé en classe de seconde :*



3 Paramètres statistiques

Définition. On note x_{min} et x_{max} les valeurs minimum et maximum d'une série statistique.

Exemple. On considère les notes du Devoir Surveillé en classe de seconde :

$$x_{min} = 5 \text{ et } x_{max} = 16.$$

Définition. On appelle étendue d'une série statistique $e = x_{max} - x_{min}$.

Exemple. On considère les notes du Devoir Surveillé en classe de seconde :

$$e = 16 - 5 = 11.$$

Définition. On appelle mode d'une série statistique la valeur la plus fréquente.

Exemple. On considère les notes du Devoir Surveillé en classe de seconde :

le mode est 11, la classe modale est $]10; 15]$.

Définition. On appelle médiane m d'une série statistique la valeur pour laquelle la moitié des individus sont associés à une valeur supérieure ou égale à m et l'autre moitié des individus à une valeur inférieure ou égale à m .

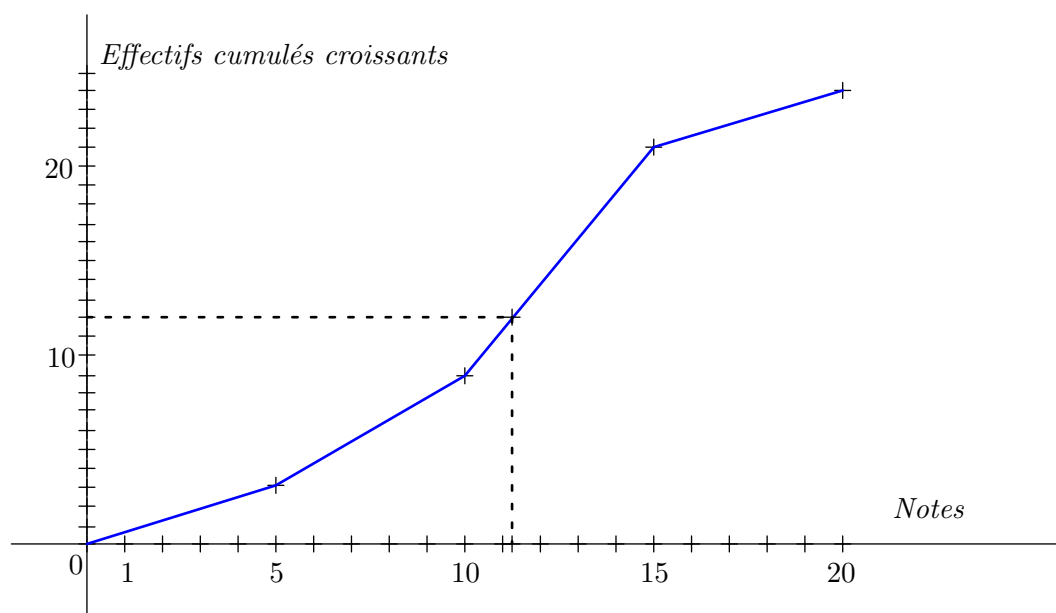
Exemple. On considère les notes du Devoir Surveillé en classe de seconde :

Il y a en tout 24 notes, on prend la moyenne des 12^{ème} et 13^{ème} notes dans l'ordre croissant, la médiane est $m = 11$.

Dans le cas d'un regroupement par classes, on construit la courbe des effectifs cumulés croissants et on cherche la valeur correspondant à la moitié de l'effectif total.

Exemple. On considère les notes du Devoir Surveillé en classe de seconde :

Classes	[0; 5]]5; 10]]10; 15]]15; 20]
Effectifs	3	6	12	3
Effectifs cumulés croissants	3	9	21	24



La médiane est $m \simeq 11,25$.

Définition. On appelle moyenne d'une série statistique de valeurs x_i et d'effectifs n_i :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots}{n_1 + n_2 + \dots} = \frac{\sum n_i x_i}{\sum n_i}$$

Exemple. On considère les notes du Devoir Surveillé en classe de seconde :

$$\bar{x} = \frac{3 \times 5 + 2 \times 7 + 2 \times 8 + \dots}{3 + 2 + 2 + \dots} = 10,625$$

Propriété. Soit une série statistique de valeurs x_i et d'effectifs n_i , alors :

$$\bar{x} = f_1x_1 + f_2x_2 + \dots = \sum f_i x_i$$

Démonstration. On note $N = n_1 + n_2 + \dots$ l'effectif total :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots}{N} = \frac{n_1}{N}x_1 + \frac{n_2}{N}x_2 + \dots = f_1x_1 + f_2x_2 + \dots$$

□

Propriété. On partage une série statistique en deux groupes disjoints d'effectifs N_1 et N_2 et de moyennes \bar{x}_1 et \bar{x}_2 . Alors la moyenne de la série est :

$$\bar{x} = \frac{N_1\bar{x}_1 + N_2\bar{x}_2}{N_1 + N_2}$$

Démonstration. ce résultat est admis.

□

Exemple. Lors d'un devoir commun en seconde, la classe de Seconde1 (25 élèves) obtient une moyenne de 9 et la classe de Seconde2 (30 élèves) obtient une moyenne de 11. La moyenne sur l'ensemble de ces deux classes est :

$$\bar{x} = \frac{25 \times 9 + 30 \times 11}{25 + 30} \simeq 10,1$$