

Chapitre 4: Statistiques

1 Le vocabulaire

→ **Population** : Une population est un ensemble de personnes ou d'objets, appelés individus, définis par une propriété commune.

Exemple : Les élèves d'une classe.

→ **Caractère** : Pour une population choisie, on peut étudier un caractère de ses individus.

Exemple : La taille des élèves, le nombre de frères et soeurs, la couleur des yeux...

→ **Caractère quantitatif** : Un caractère est dit quantitatif lorsqu'il est possible de le mesurer en associant un nombre à chaque individu. Ce caractère est dit **continu** lorsque les nombres qui le mesurent peuvent prendre toutes les valeurs d'un intervalle et est dit **discret** dans le cas contraire.

Exemple : La taille des élèves est un caractère quantitatif continu et le nombre de frères et soeurs est un caractère quantitatif discret.

→ **Caractère qualitatif** : Tout caractère non-quantitatif est dit qualitatif.

Exemple : La couleur des yeux.

Dans la suite du cours nous allons, pour illustrer les diverses définitions, utiliser la série statistique composées des notes à l'épreuve de mathématiques du baccalauréat des élèves d'une classe de terminale

→ La population étudiée est une classe de terminale.

→ Le caractère étudié est la note à l'épreuve de mathématiques du baccalauréat, il est quantitatif discret.

2 Présentation d'une série statistique

2.1 Effectifs et effectifs cumulés croissants

Le tableau ci-dessous donne les notes des élèves :

Note x_i	5	6	7	8	9	10	11	12	13	14	15	16	17
Effectif n_i	1	3	2	0	3	6	8	4	1	2	0	0	2

Si on note N l'effectif total, $N = 32$.

A partir de ce tableau, on peut dresser le tableau des effectifs cumulés croissants (E.C.C) :

Note $\leq x$	5	6	7	8	9	10	11	12	13	14	15	16	17
E.C.C	1	4	6	6	9	15	23	27	28	30	30	30	32

A l'aide du tableau, on peut affirmer que 15 élèves ont une note inférieure ou égale à 10.

2.2 Fréquences et fréquences cumulées croissantes

A partir des effectifs, on obtient les fréquences. Par exemple, $f_6 = \frac{4}{32} = 0,125$.

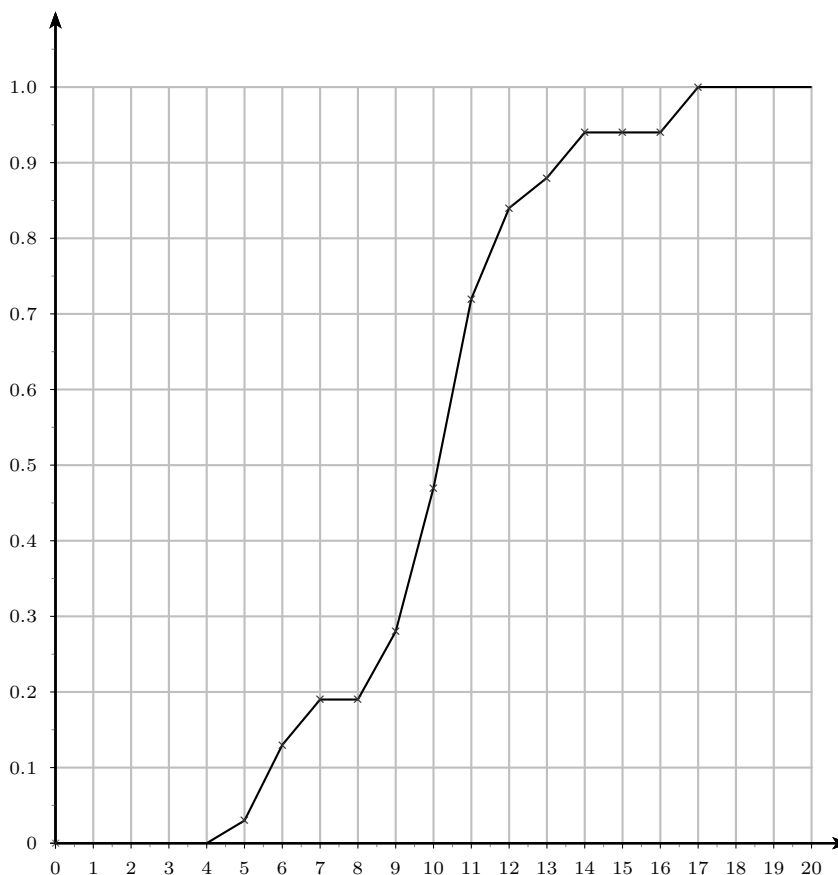
Note x_i	5	6	7	8	9	10	11	12	13	14	15	16	17
Fréquence f_i	0,03125	0,09375	0,0625	0	0,09375	0,1875	0,25	0,125	0,03125	0,0625	0	0	0,0625

La somme des fréquences est toujours égale à 1.

A partir de ce tableau, on peut dresser le tableau des fréquences cumulées croissantes (F.C.C) :

Note $\leq x$	5	6	7	8	9	10	11
F.C.C.	0,03125	0,125	0,1875	0,1875	0,28125	0,46875	0,71875
Note $\leq x$	12	13	14	15	16	17	
F.C.C.	0,84375	0,875	0,9375	0,9375	0,9375	1	

On peut ainsi tracer la courbe des fréquences cumulées croissantes :



3 Paramètres d'une série statistique

3.1 Moyenne

Définition:

La moyenne d'une série statistique est le nombre, noté $\bar{x}$, défini par :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n n_i x_i = \frac{n_1 x_1 + \dots + n_p x_p}{N} \quad \text{avec} \quad N = n_1 + n_2 + \dots + n_p$$

Exemple:

La moyenne des notes à l'épreuve de mathématiques du baccalauréat des élèves de cette classe de terminale est :

$$\bar{x} = \frac{5 \times 1 + 6 \times 3 + \dots + 17 \times 2}{32} = \frac{335}{32} \simeq 10,5$$

Remarque:

On peut aussi calculer la moyenne d'une série statistique en utilisant les fréquences :

$$\bar{x} = \sum_{i=1}^n f_i x_i = f_1 x_1 + \dots + f_p x_p \quad \text{avec} \quad f_i = \frac{x_i}{n_i}$$

3.2 Médiane et quartiles

Définition:

La médiane d'une série statistique est le nombre, noté M_e , tel que :

- 50% au moins des individus ont une valeur du caractère inférieure ou égale à M_e ;
- 50% au moins des individus ont une valeur du caractère supérieure ou égale à M_e ;

Remarque:

Le liste des N valeurs étant rangée par ordre croissant, chacune figurant un nombre de fois égal à son effectif :

→ Si N est impair, la médiane est la valeur de rang $\frac{N+1}{2}$;

→ Si N est pair, la médiane est la demi-somme des valeurs de rang $\frac{N}{2}$ et $\frac{N}{2} + 1$.

Exemple:

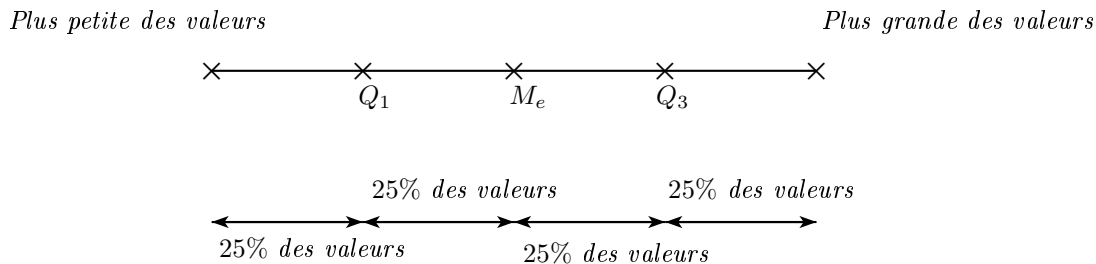
La série des notes à l'épreuve de mathématiques du baccalauréat des élèves de cette classe de terminale comporte 32 valeurs, qui est un nombre pair. La médiane est donc donnée par la demi-somme de la 16^{ème} valeur et de la 17^{ème} valeur :

$$M_e = \frac{11 + 11}{2} = 11$$

Définition:

La liste des N valeurs est rangée dans l'ordre croissant :

- Le premier quartile est la plus petite valeur Q_1 de la liste telle qu'au moins un quart des valeurs de la liste sont inférieures ou égale à Q_1 .
- Le troisième quartile est la plus petite valeur Q_3 de la liste telle qu'au moins trois quart des valeurs de la liste sont inférieures ou égale à Q_3 .
- L'écart interquartile est le nombre $Q_3 - Q_1$.



Exemple:

Pour la série des notes à l'épreuve de mathématiques du baccalauréat des élèves de cette classe de terminale, on a :

- L'étendue est $17 - 5 = 12$.
- $\frac{32}{4} = 8$ donc $Q_1 = 9$.
- $\frac{32 \times 3}{4} = 24$ donc $Q_3 = 12$.
- L'écart interquartile est $Q_3 - Q_1 = 3$.