

1. Vocabulaire

Population :

Individu :

Effectif total :

Caractère :

Il y a deux sortes de caractères :

- les caractères (profession, marque de voiture ...).
- les caractères que l'on peut mesurer avec des nombres. On distingue les caractères quantitatifs qui ne peuvent prendre qu'un nombre fini de valeurs (notes à un devoir...) et les caractères quantitatifs dont on regroupe les valeurs par intervalles (taille, durée d'écoute...).

Pour les paragraphes suivants, on considère l'exemple suivant : 17 élèves sont séparés en deux groupes.

- Les notes obtenues à un devoir par le **groupe A** sont : 4 ; 4 ; 9 ; 9 ; 9 ; 15 ; 15 ; 17 ; 17
- Les notes obtenues à un devoir par le **groupe B** sont : 6 ; 6 ; 8 ; 8 ; 10 ; 10 ; 12 ; 12

2. Séries statistiques

a) Classement des données

► **Définition** : On appelle **série statistique** la donnée simultanée des valeurs du caractère étudié (notées x_i), rangées dans l'ordre croissant, et des effectifs (notés n_i) de ces valeurs.

Pour le groupe A de l'exemple, le regroupement des notes en série statistique est :

valeur	4	9	15	17
effectif				

Pour le groupe B de l'exemple, le regroupement des notes en série statistique est :

valeur	6	8	10	12
effectif				

► **Définition** :

L'effectif cumulé croissant d'une valeur x est la somme des effectifs des valeurs y tels que

L'effectif cumulé décroissant d'une valeur x est la somme des effectifs des valeurs y tels que

Pour le groupe A de l'exemple, le tableau des effectifs cumulés croissants et décroissants est :

valeur	4	9	15	17
effectif cumulé croissant				
effectif cumulé décroissant				

Une série statistique peut aussi être définie par les fréquences, plutôt que par les effectifs.

► **Définition** : Dans une population d'effectif total N , la **fréquence** de la valeur x_i d'effectif n_i est $f_i = \frac{n_i}{N}$.

Pour le groupe B de l'exemple, le tableau des fréquences est :

valeur	6	8	10	12
fréquence en pourcentage				

b) Représentation graphique

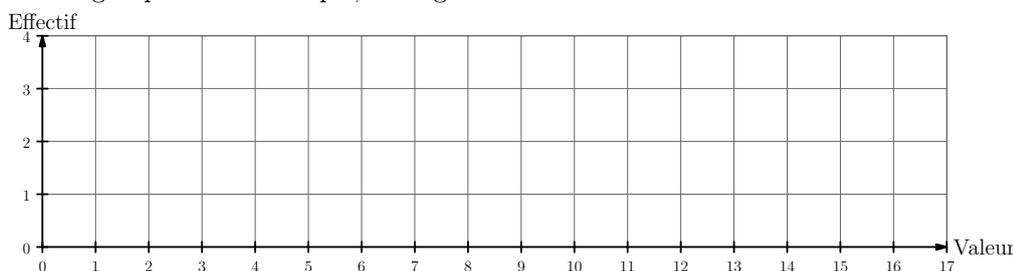
- Pour les caractères quantitatifs discrets, on utilise le :

Dans un repère orthogonal, pour chaque valeur de la série statistique on trace un trait vertical dont la hauteur correspond à l'effectif (dans l'unité choisie).

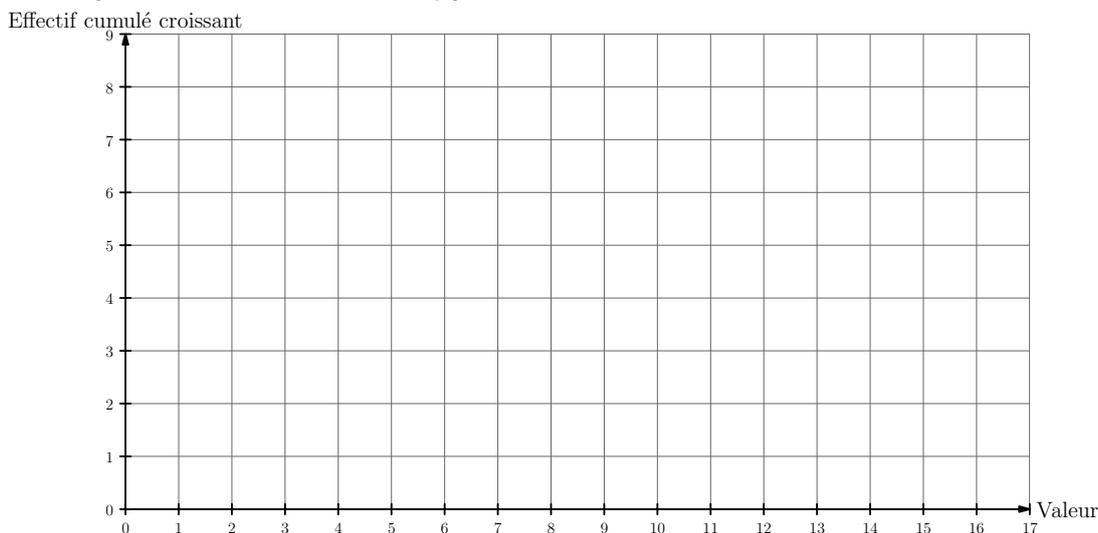
- On utilise aussi le qui se trace en reliant les points successifs de coordonnées (valeur, effectif) pour obtenir une ligne brisée.

En utilisant les effectifs cumulés croissants ou décroissants, on obtient le polygone des effectifs cumulés croissants ou décroissants.

Pour le groupe A de l'exemple, le diagramme en bâtons est :



Pour le groupe A de l'exemple, le polygone des effectifs cumulés croissants est :



3. Moyenne et écart-type d'une série statistique

a) Moyenne

► **Définition :** On appelle **moyenne** d'une série statistique d'effectif total N , le réel $\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{N}$. (k représente le nombre de valeurs prises par le caractère)

► **Remarques :**

- En utilisant les fréquences, on a : $\bar{x} = f_1x_1 + f_2x_2 + \dots + f_kx_k$.
- Si on ajoute à toutes les valeurs d'une série statistique le même nombre b ,
- Si les valeurs d'une série statistique sont multipliées ou divisées par un même nombre a ,
- Si une population d'effectif N est composée d'une partie d'effectif N_A et de moyenne \bar{x}_A et d'une autre partie d'effectif N_B et de moyenne \bar{x}_B , alors la moyenne de la population totale est telle que : $\bar{x} = \frac{N_A\bar{x}_A + N_B\bar{x}_B}{N}$

Pour le groupe A de l'exemple, la moyenne est $\bar{x}_A =$

Pour le groupe B de l'exemple, la moyenne est $\bar{x}_B =$

On peut en déduire que la moyenne globale de la classe est $\bar{x} = \frac{N_A\bar{x}_A + N_B\bar{x}_B}{N} =$

b) Écart-type

► **Définition :** La **variance** V d'une série statistique d'effectif total N est

$$V = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_k(x_k - \bar{x})^2}{N}$$

► **Définition :** L'**écart-type** σ d'une série statistique est la racine carrée de la variance : $\sigma = \sqrt{V}$. Il sert à

Pour les deux groupes de l'exemple, compléter les tableaux suivants et en déduire la variance et l'écart-type de chaque groupe :

Groupe A

valeur	4	9	15	17
écart à la moyenne				
carré de l'écart				
effectif				

Variance $V_A =$

Écart-type $\sigma_A =$

Groupe B

valeur	6	8	10	12
écart à la moyenne				
carré de l'écart				
effectif				

Variance $V_B =$

Écart-type $\sigma_B =$

► **Remarques :**

- L'écart-type est de même unité que les valeurs de la série statistique.
- Si on ajoute à toutes les valeurs d'une série statistique le même nombre b ,
- Si les valeurs d'une série statistique sont multipliées ou divisées par un même nombre $a > 0$,
- Si les valeurs d'une série statistique sont multipliées ou divisées par un même nombre $a < 0$,

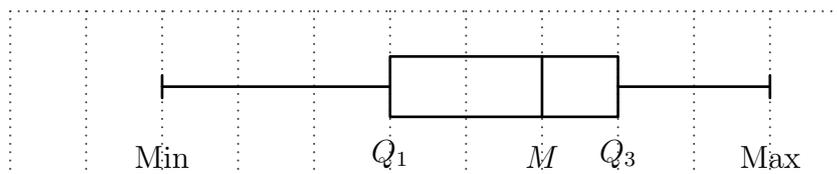
4. Médiane et écart interquartile d'une série statistique

► **Idée générale :** la médiane M d'une série statistique partage la population en deux parties de telle façon qu'au moins 50% des valeurs du caractère soient inférieurs ou égaux à M et qu'au moins 50% des valeurs du caractère soient supérieurs ou égaux à M .

► **Définition et méthode pratique**

En écrivant les valeurs du caractère par ordre croissant de telle façon que chaque valeur apparaisse un nombre de fois égal à son effectif :

- **Médiane :**
 - si l'effectif total est impair, la médiane M est
 - si l'effectif total est pair, la médiane M est
- **Quartiles :** en partageant la liste en deux sous-séries de même effectif (si l'effectif total est impair, on ne tient pas compte de la médiane)
 - le premier quartile Q_1 est
 - le troisième quartile Q_3 est
 - l'écart interquartile est défini par et sert à mesurer la dispersion des valeurs.
 - le diagramme en boîtes de la série statistique se construit de la façon suivante :

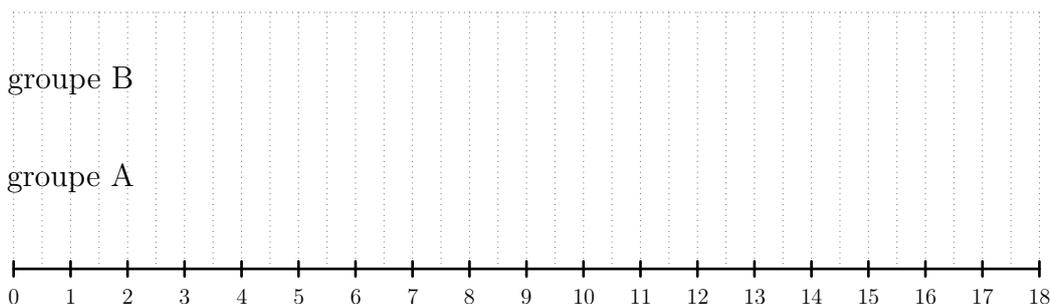


- 25% des valeurs du caractère sont comprises entre Min et Q_1 ;
- 25% des valeurs du caractère sont comprises entre Q_1 et M ;
- 25% des valeurs du caractère sont comprises entre M et Q_3 ;
- 25% des valeurs du caractère sont comprises entre Q_3 et Max.

Pour l'exemple, déterminer les valeurs de la médiane et de Q_1 et Q_3 pour les deux groupes. En déduire les diagrammes en boîtes des deux groupes dans le repère ci-dessous.

• **groupe A :** 4 ; 4 ; 9 ; 9 ; 9 ; 15 ; 15 ; 17 ; 17

• **groupe B :** 6 ; 6 ; 8 ; 8 ; 10 ; 10 ; 12 ; 12



5. Exemple d'étude d'un caractère quantitatif continu

On a posé la question suivante à 34 élèves : « Combien de temps avez-vous consulté votre téléphone portable hier ? »

Voici les résultats obtenus :

temps en minutes	[0, 30[[30, 60[[60, 120[[120, 180[
nombre d'élèves	12	8	10	4

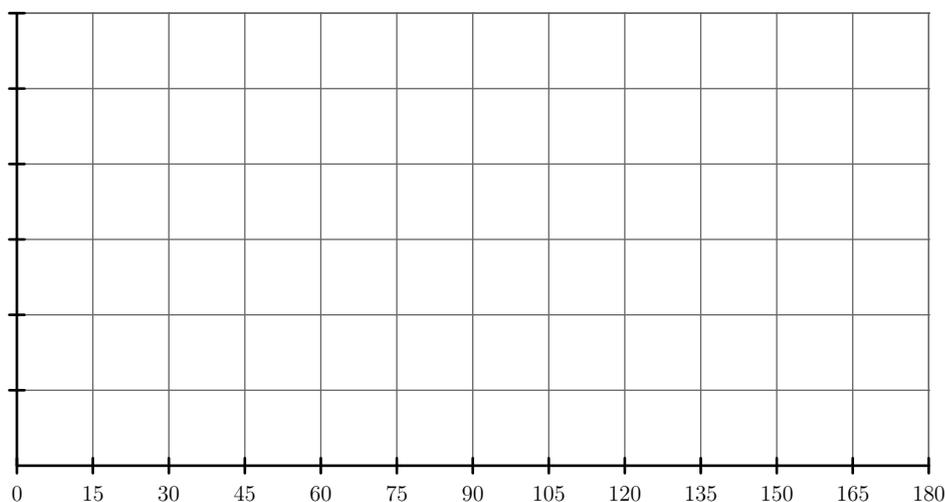
► **Remarque** : les intervalles [0, 15[, [15, 30[... sont appelés du caractère.

a) Histogramme

Pour la représentation graphique d'un caractère continu, on utilise généralement un : dans un repère orthogonal on porte en abscisse les valeurs des bornes, puis pour chaque classe on trace un rectangle dont l'aire est proportionnelle à l'effectif.

Pour l'exemple, compléter le tableau et tracer l'histogramme ci-dessous :
(unités : en abscisse 1 cm représente 15 min et 1 cm² représente 1 élève)

temps en minutes	[0, 30[[30, 60[[60, 120[[120, 180[
aire du rectangle en cm ²	12	8	10	4
largeur du rectangle en cm				
hauteur du rectangle en cm				



b) Moyenne et écart-type pour des valeurs regroupées en classes

Pour calculer la moyenne d'une série statistique d'un caractère quantitatif continu, on prend comme valeur du caractère

Pour l'exemple,

- Compléter le tableau suivant et en déduire la moyenne :

valeur (milieu de chaque intervalle)				
effectif				

Moyenne \bar{x} =

- Compléter le tableau suivant et en déduire la variance et l'écart-type :

valeur (milieu de chaque intervalle)				
écart à la moyenne				
carré de l'écart				
effectif				

Variance V =

Écart-type σ =