

# Chapitre :

# Statistiques



## I. Rappels pour une variable

---

### 1. Médiane et écart interquartile

Lorsque l'on a ordonné une série statistique par ordre croissant des valeurs, on peut déterminer la médiane et les quartiles en regardant les fréquences cumulées croissantes (ou f.c.c.) :

#### Définition

- La **médiane**  $Me$  est la plus petite valeur pour laquelle on atteint ou dépasse 50% ;
- Le **premier quartile**  $Q_1$  est la plus petite valeur pour laquelle on atteint ou dépasse 25% ;
- Le **troisième quartile**  $Q_3$  est la plus petite valeur pour laquelle on atteint ou dépasse 75% ;

Exemple Imaginons que l'on ait le tableau statistique suivant :

Valeurs	3	5	6	8	11	12	15
f.c.c. (en %)	7	23	36	52	72	84	100

Alors  $Me = 8$ ,  $Q_1 = 6$  et  $Q_3 = 12$ .

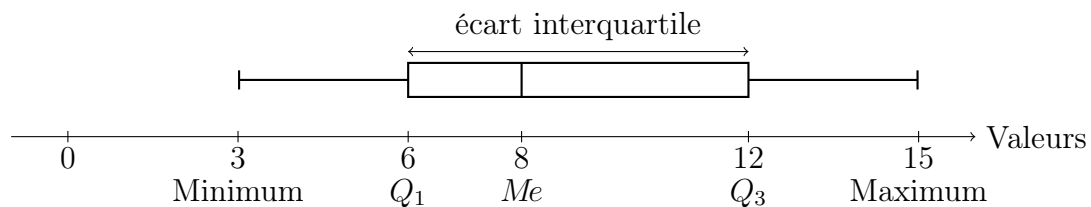
Il existe une autre règle pour la médiane, quand on écrit toutes les valeurs dans l'ordre croissant :

- S'il y a un nombre impair de données, c'est la valeur centrale ;
- S'il y a un nombre pair de données, c'est la moyenne des deux valeurs centrales ;

#### Définition

On appelle **intervalle interquartile** l'intervalle  $[Q_1; Q_3]$  et **écart interquartile** le nombre  $Q_3 - Q_1$ .

On peut représenter une série statistique à une variable par un diagramme en boîte (aussi appelé boîte à moustache) de la manière suivante, en reprenant les valeurs de l'exemple plus haut :



La droite est graduée. En particulier, il faut respecter une échelle choisie au départ.

► **Exercices : (en DM) activité 1p120, 1p130, 21p132 et 25p132**

## 2. Moyenne et écart-type

⊗ **Activité** : 2p120 (moyenne, écart-type) (voir page 128 pour l'utilisation de la calculatrice)

**Définition** Si les données statistiques sont résumées dans un tableau comme ci-dessous :

Valeurs	$x_1$	$x_2$	$\dots$	$x_p$	Total
Effectifs	$n_1$	$n_2$	$\dots$	$n_p$	$N$

alors la moyenne est définie par :

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

**Propriété** | L'écart-type (déterminé à l'aide de la calculatrice), est un nombre positif noté  $\sigma$  qui mesure la dispersion des valeurs autour de la moyenne.

► **Exercices** : 7,8,11p130

► **Exercices** : 30p133,32p134

# II. Nuage de points

---

⊗ **Activité** : 3p124

## 1. Définitions

On considère une population sur laquelle on étudie deux caractères  $X$  et  $Y$ , qui prennent des valeurs quantitatives. Pour cela, on associe à chaque individu un couple de valeurs  $(x_i; y_i)$  correspondant aux valeurs prises par l'individu pour les caractères  $X$  et  $Y$ .

Les valeurs sont en général données dans un tableau. On représente ces données sur un graphique, par ce que l'on appelle un **nuage de points**. Il s'agit de placer dans un repère les points de coordonnées  $(x_i; y_i)$ .

**Définition** Soit  $(X; Y)$  une série statistiques double. Si l'on note  $\bar{x}$  (resp.  $\bar{y}$ ) la moyenne des  $x_i$  (resp.  $y_i$ ), Alors le **point moyen** est le point  $G$  de coordonnées  $(\bar{x}; \bar{y})$ .

► **Exercices** : 12p130, 13,14p131

► **Exercices** : 35,36p134

## 2. Ajustement affine

**Définition** On cherche en général à savoir s'il existe un lien entre les deux séries  $X$  et  $Y$ . Effectuer un **ajustement** consiste à chercher s'il existe une fonction dont la courbe s'approche le plus possible du nuage de points.

Si le nuage de point est à peu près rectiligne, on parle d'**ajustement affine**. Nous verrons à la section suivante comment obtenir les paramètres d'un tel ajustement, avec une droite qui passe par le point moyen.

► **Exercice** : 15p131

### **Cas particulier : la méthode des moindres carrés**

La méthode des moindres carrés permet de faire un ajustement affine et d'obtenir ainsi une droite.

La droite obtenue minimise la somme des carrés des distances verticales à la droite.

Elle a également comme particularité de passer par le point moyen (voir la figure plus bas).

**Méthode** Pour obtenir l'équation de cette droite, de la forme  $y = ax + b$ , on utilise la calculatrice, qui donne les valeurs de  $a$  et  $b$ .

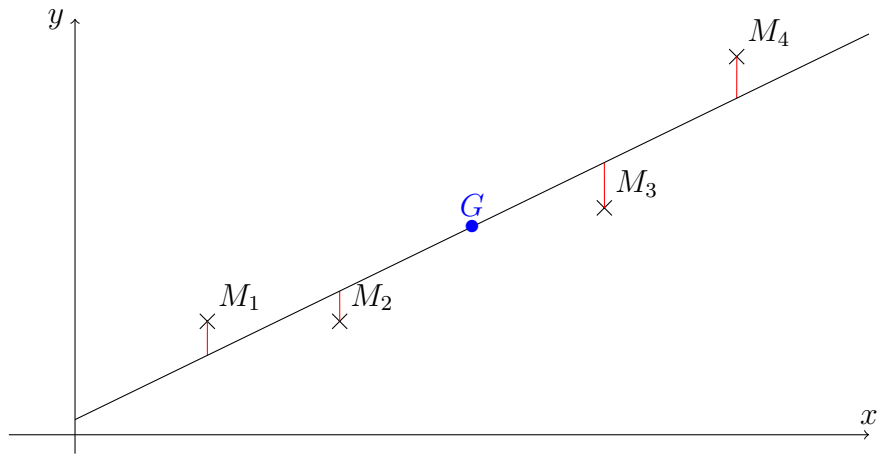
Voir page 129.

► **Exercices** : 19p131

► **Exercice** : 47,48,50,51p137

Cette méthode n'est cependant pas la seule permettant de faire un ajustement affine.

► **Exercices** : 40,41p135 (en DM?)



La droite passe par  $G$  et réduit la somme des carrés des longueurs rouges