

Les statistiques – Fiche de cours

1. Vocabulaire des statistiques

- Caractère

Le caractère représente le critère ou le trait sur lequel porte l'étude de la série statistique.

- Population

La population est l'ensemble des individus étudiés.

2. Les séries statistiques

- Série de valeurs

La série statistique se présente sous la forme de plusieurs valeurs :

Exemple : 10 – 12- 20 - 25

- Tableau d'effectifs

La série statistique se présente sous la forme d'un tableau présentant plusieurs effectifs pour une même valeur du caractère.

| | | | | |
|----------|-------|-------|-----|-------|
| Valeur | x_1 | x_2 | ... | x_n |
| Effectif | n_1 | n_2 | ... | n_n |

- Tableau d'effectifs en classe

| | | | | |
|---------------------|-----------------------|-----------------------|-----|-----------------------|
| Valeur | $[a ; b[$ | $[b ; c[$ | ... | $[y ; z[$ |
| Centre de la classe | $x_1 = \frac{a+b}{2}$ | $x_2 = \frac{b+c}{2}$ | | $x_n = \frac{y+z}{2}$ |
| Effectif | n_1 | n_2 | ... | n_n |

Lorsque le caractère est défini par classe on doit déterminer le centre de chaque classe.

3. Paramètres de position d'une série statistique

a. Moyenne

La moyenne d'une série statistique est définie par :

$$\text{moyenne} = \frac{\text{somme des valeurs}}{\text{somme des effectifs}}$$

Lorsque la série se présente sous la forme d'un tableau d'effectifs la moyenne est définie par :

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_n n_n}{n_1 + n_2 + \dots + n_n}$$

b. Linéarité de la moyenne

Soient x et y deux séries statistiques $\forall a \in \mathbb{R} \quad \forall b \in \mathbb{R} :$

$$\text{si } y_i = a \cdot x_i + b \text{ alors } \bar{y} = a \cdot \bar{x} + b$$

c. Médiane

La médiane Me est définie comme la plus petite valeur du caractère rangé par ordre croissant, qui donne au moins 50% de l'effectif total.

- Effectif total impair

Lorsque l'effectif total est impair la médiane est la $\frac{N+1}{2}$ n^{ième} valeur du caractère rangé par ordre croissant.

- Effectif total pair

Lorsque l'effectif total est pair la médiane est la moyenne entre la $\frac{N}{2}$ et la $\frac{N}{2} + 1$ n^{èmes} valeurs du caractère rangé par ordre croissant.

4. Paramètres de dispersion d'une série statistique

a. Ecart type

L'écart type d'une série statistique est défini par :

$$\sigma = \sqrt{\frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_n(x_n - \bar{x})^2}{N}}$$

b. Quartiles

- Premier quartile

Le premier quartile est défini comme la plus petite valeur du caractère rangé par ordre croissant, qui donne au moins 25% de l'effectif total.

- Troisième quartile

Le troisième quartile est défini comme la plus petite valeur du caractère rangé par ordre croissant, qui donne au moins 75% de l'effectif total.

- Ecart interquartile

$$I = Q_3 - Q_1$$

4. Comparaison des séries statistiques

- si $I < Me$ (série homogène) le couple $(Me; I)$ est approprié pour comparer les séries statistiques.

- si $I > Me$ (série dispersée) le couple $(\bar{x}; \sigma)$ est approprié pour comparer les séries statistiques.

5. Représentation graphique

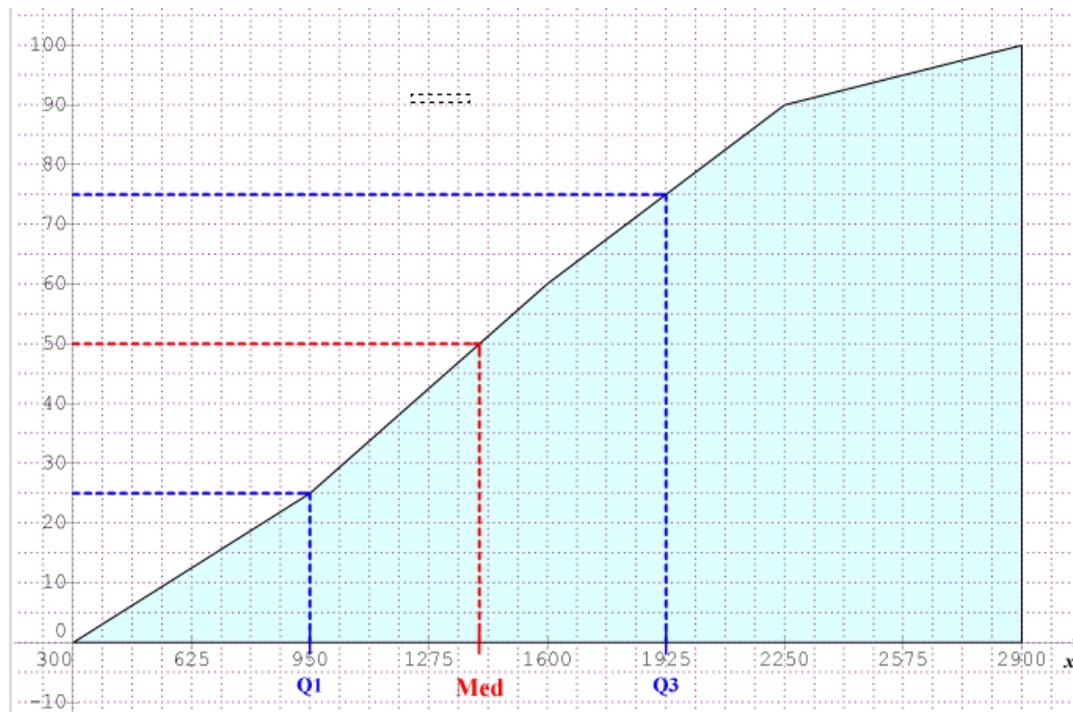
a. Diagramme en bâton et histogramme

L'histogramme permet de comparer les effectifs ; il représente une évolution.

La hauteur de chaque barre est proportionnelle à l'effectif.

b. Polygone des effectifs cumulés croissants

Lorsque le caractère est continu on peut utiliser le polygone des effectifs cumulés croissants pour déterminer Q_1 , Me et Q_3 .



c. Diagramme en boîte (diagramme à moustaches)

