Statistiques descriptives – 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.

- **Variables statistiques**

  - qualitative nominale : on ne peut pas ordonner le caractère selon une qualité (exemple : rouge, vert, bleu)
  - qualitative ordinale : on peut ordonner le caractère selon une qualité (exemple, faible, moyen, fort)
  - quantitative discrète : le caractère est constitué par un nombre fini de valeurs
  - quantitative continue : le caractère est constitué par des intervalles de valeurs

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. Mode

Le mode est la valeur du caractère qui correspond à l'effectif le plus important de la série statistique

Il existe des séries bimodales, trimodales (associées aux 2 ou 3 effectifs les plus importants de la série statistique

b. 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 :

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

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^{ièmes} valeurs du caractère rangé par ordre croissant.

4. Paramètres de dispersion d'une série statistique

a. Etendue

$$étendue = maximum - minimum$$

b. Ecart type

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

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

c. Quartiles

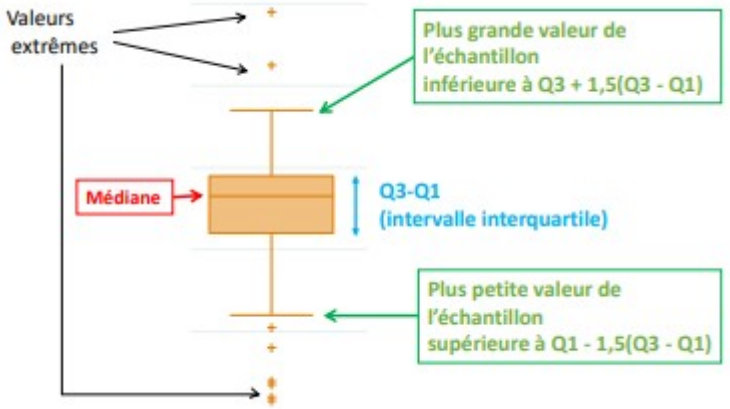
- **Premier quartile**

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**

$$I = Q_3 - Q_1$$

d. Diagramme box-plot