

Cours

Cours 1 – Vocabulaire

Une **étude statistique** porte sur l'observation d'un **caractère** commun aux **individus** d'une **population**.

Le caractère est **quantitatif** s'il est mesurable (nombre).

Le caractère est **quantitatif discret** s'il ne prend que des valeurs isolées.

Le caractère est **quantitatif continu** s'il peut prendre toutes les valeurs comprises dans un intervalle donné.

Le caractère est **qualitatif** s'il n'est pas mesurable.

Les données d'une étude statistique sont regroupées dans un **tableau** à simple entrée.

Sur la première ligne (ou colonne) du tableau on indique les différentes **valeurs du caractère (modalités)**.

Sur la deuxième ligne (ou colonne) du tableau on indique l'**effectif** (nombre de fois où cette valeur est obtenue).

La somme de tous les effectifs est l'**effectif total** et est noté N .

Dans le cas d'un caractère quantitatif continu, on peut regrouper les valeurs dans des **classes**. Les classes sont des intervalles de la forme $[a; b]$. Les valeurs de la classe sont supérieures ou égales à a et strictement inférieures à b . Dans ce cas, l'écart entre les deux bornes $A = b - a$ est l'**amplitude de classe A**, le **centre de classe** est le milieu de la classe et est donné par la relation $(a+b)/2$.

Cours 2 – Fréquence

Pour calculer la **fréquence** f d'une valeur du caractère, on divise l'effectif de la valeur (ou de la classe) par l'effectif total puis on multiplie par 100. Le résultat est exprimé en pourcentage.

$$f = \frac{\text{effectif}}{\text{effectif total}} \times 100$$

Le total des fréquences doit être de 100 %.

Remarque :

On peut également exprimer la fréquence sous la forme d'un nombre décimal compris entre 0 et 1. Dans ce cas, la formule utilisée est :

$$f = \frac{\text{effectif}}{\text{effectif total}}$$

Cours 3 – Représentation graphique d'une série statistique à une variable

Diagramme en bâtons

Il permet de représenter une série statistique dont le caractère est qualitatif ou quantitatif discret.

Les valeurs du caractère sont représentées sur l'axe des abscisses et les effectifs sur l'axe des ordonnées. Les bâtons sont séparés les uns des autres. Les hauteurs des bâtons sont proportionnelles aux effectifs (ou pourcentages).

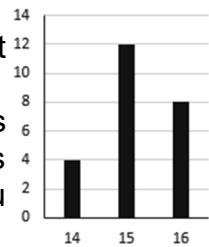


Diagramme circulaire (ou à secteurs)

Il permet de représenter une série statistique dont le caractère est qualitatif.

Chaque secteur représente une valeur du caractère. La mesure de l'angle au centre de chaque secteur est proportionnelle aux effectifs (ou pourcentages).



Diagramme en colonnes (ou histogramme)

Il permet de représenter une série statistique dont le caractère est quantitatif continu.

Les valeurs du caractère sont représentées sur l'axe des abscisses et les effectifs sur l'axe des ordonnées. Les colonnes sont accolées les unes aux autres. Si les classes ont la même amplitude, les hauteurs des colonnes sont proportionnelles aux effectifs (ou pourcentages).

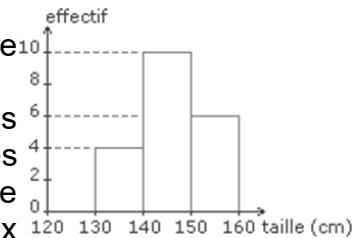
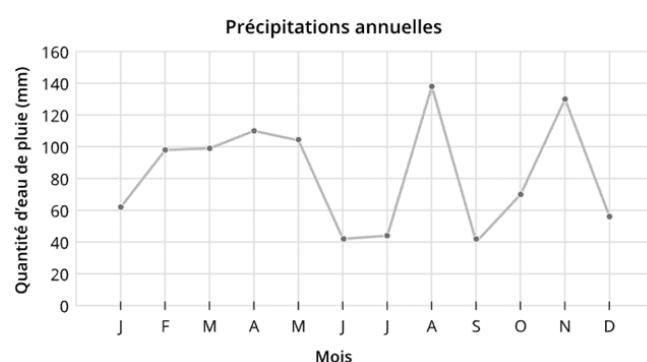


Diagramme à ligne brisée

Il permet de représenter une série statistique dont le caractère est quantitatif et varie en fonction du temps.

L'axe horizontal est gradué en unités de temps et l'axe vertical présente les valeurs du caractère. Les données sont représentées par des points reliés les uns aux autres.



Cours 4 – Indicateurs statistiques

Les **indicateurs statistiques** sont des nombres calculés à partir des données d'une série statistique.

Pour les calculer on utilise la calculatrice ou le tableur. Ils ont la même unité que le caractère étudié. Ils permettent de comparer des séries statistiques entre elles.

Ils sont de deux natures différentes : les **indicateurs de position** et les **indicateurs de dispersion**.

Cours 5 – Indicateurs de position

Ils permettent de mesurer le **centrage** des données d'une série statistique.

Mode (ou classe modale)

C'est la valeur ou la classe la plus fréquente.

Médiane (notée Med)

C'est la valeur qui sépare la série statistique en deux groupes de même effectif.

50 % des valeurs sont inférieures ou égales à la médiane.

50 % des valeurs sont supérieures ou égales à la médiane.

Premier quartile (noté Q_1)

C'est la valeur qui sépare la série statistique en deux groupes tels que :

25 % des valeurs sont inférieures ou égales à Q_1 .

75 % des valeurs sont supérieures ou égales à Q_1 .

Troisième quartile (noté Q_3)

C'est la valeur qui sépare la série statistique en deux groupes tels que :

75 % des valeurs sont inférieures ou égales à Q_3 .

25 % des valeurs sont supérieures ou égales à Q_3 .

Moyenne (notée \bar{x} et lu x barre)

C'est la somme de toutes les valeurs divisée par le nombre de valeurs.

La moyenne est plus sensible que la médiane aux valeurs extrêmes

Cours 6 – Indicateurs de dispersion

Ils permettent de mesurer le **l'étalement** des données d'une série statistique.

Étendue (notée E)

C'est la différence entre le maximum et le minimum d'une série statistique $E = Max - Min$.

Elle mesure la dispersion totale de la série statistique.

Écart interquartile (noté $Q_3 - Q_1$) :

Il mesure la dispersion de 50 % des valeurs autour de la médiane.

Écart type (noté σ et lu *sigma*) :

Il mesure la dispersion autour de la moyenne.

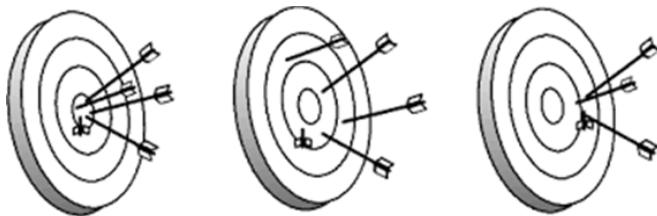
Cours 7 – Comparer des séries statistiques

Pour comparer des séries statistiques il faut :

- Comparer les centrages.
- Comparer les dispersions.

On ne peut comparer que deux indicateurs identiques entre eux. Pour comparer des séries statistiques on peut utiliser les couples :

- Médiane / Écart interquartile.
- Moyenne / Écart type



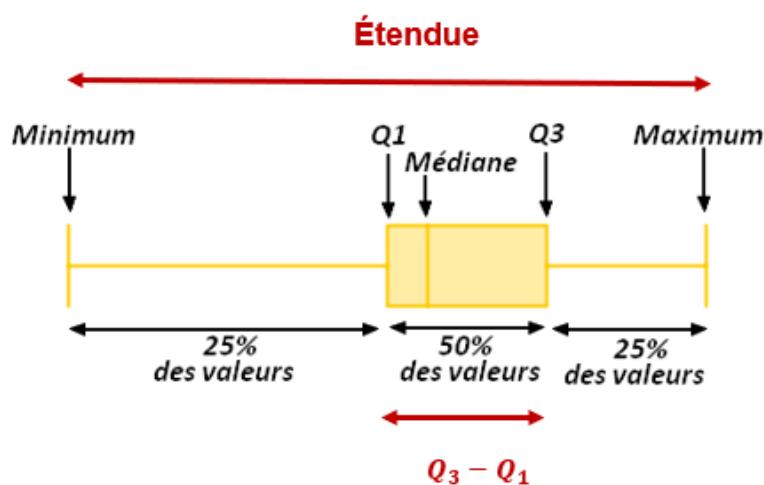
Cours 8 – Diagramme en boîte à moustaches

C'est un diagramme qui regroupe plusieurs indicateurs :

- Minimum *Min.*
- Premier quartile Q_1 .
- Médiane *Med.*
- Troisième quartile Q_3 .
- Maximum *Max.*

Entre la minimum et le maximum on retrouve l'étendue.

Entre le premier et le troisième quartile on retrouve l'écart interquartile.



À l'aide des diagrammes en boîte à moustaches on peut comparer visuellement des séries statistiques entre elles.