

TRAITEMENT DE DONNÉES STATISTIQUES

Effectuer une **étude statistique** consiste à recenser et à exploiter des informations sur un ensemble constitué d'**individus** et appelé **population**.

Etudier un **caractère** c'est examiner une particularité commune à tous les individus de la population.

Les données statistiques se présentent sous forme de nombres, de tableaux, de graphiques, de pourcentages ... apportant des renseignements sur le nombre d'éléments d'un ensemble ayant une propriété en commun.

Une **série statistique** est la donnée de la répartition des **effectifs** (et/ou des **fréquences**) selon les **valeurs** d'un **caractère** étudié dans une **population** déterminée.

VOCABULAIRE, DÉFINITIONS :

population : ensemble des individus.

effectif total : nombre total d'individus de la population étudiée.

individu : élément de la population.

caractère : particularité commune à tous les individus étudiés. Le caractère peut être :

- quantitatif : il peut être mesuré, c'est donc un nombre (taille, âge, pointure ...)
- qualitatif : il ne peut pas être mesuré (couleur des cheveux, séjour de vacances, nature du logement ...)

effectif de chaque valeur : nombre d'individus correspondant à cette valeur.

étendue : pour un caractère quantitatif, c'est la différence entre la plus grande et la plus petite des valeurs du caractère. C'est une variable de **dispersion**.

classe : Quand les valeurs du caractère (quantitatif) sont nombreuses elles sont souvent regroupées en intervalles.

Dire qu'un nombre x est un élément de la classe $[a, b[$ signifie alors que l'on a : $a \leq x < b$

Le nombre $b - a$ est l'amplitude de la classe, et le nombre $\frac{a + b}{2}$ est la valeur centrale de la classe.

effectif d'une classe : nombre d'individus pour lesquels la valeur du caractère est dans la classe.

La somme des effectifs de toutes les classes est égal à l'effectif total.

fréquence relative :

- quotient de l'effectif d'une valeur (ou d'une classe) par l'effectif total. (c'est un nombre compris entre 0 et 1)
- si on multiplie par 100 on obtient la fréquence en pourcentage.

moyenne d'une série : quotient de la somme de toutes les valeurs par l'effectif total. C'est une caractéristique de **position**

Pour calculer la moyenne d'une série dont les valeurs sont regroupées en classes, on prend la valeur centrale de chaque classe comme valeur commune à tous les individus de la classe.

moyenne pondérée : pour des nombres x_1, x_2, \dots, x_n affectés chacun d'un coefficient : c_1, c_2, \dots, c_n ,

c'est le quotient de $(c_1 \cdot x_1 + c_2 \cdot x_2 + \dots + c_n \cdot x_n)$ par la somme des coefficients $(c_1 + c_2 + \dots + c_n)$.

médiane : Soit une population ordonnée selon les valeurs croissantes d'un caractère quantitatif, la médiane de cette série statistique est la valeur qui sépare la population en deux parties de même effectif. C'est une caractéristique de **position**.

quartiles : Chacune des 3 valeurs qui séparent la population en quatre parties de même effectif (le 2^{ème} quartile c'est la médiane). Pour 25% de la population le caractère a une valeur inférieure ou égale à Q1, pour les 25% suivants la valeur est comprise entre Q1 et la médiane, pour les 25% suivants la valeur est comprise entre la médiane et Q3, pour les derniers 25% la valeur est supérieure ou égale à Q3. L'écart interquartile est égal à $Q3 - Q1$. C'est une caractéristique de **dispersion**.

représentations graphiques :

- polygone des effectifs (ou des fréquences)
- polygone des effectifs cumulés (ou des fréquences cumulées), pratique pour lire la médiane ou les quartiles.
- diagrammes en bâtons, à bandes, circulaires, semi-circulaires.
- histogrammes.

...