



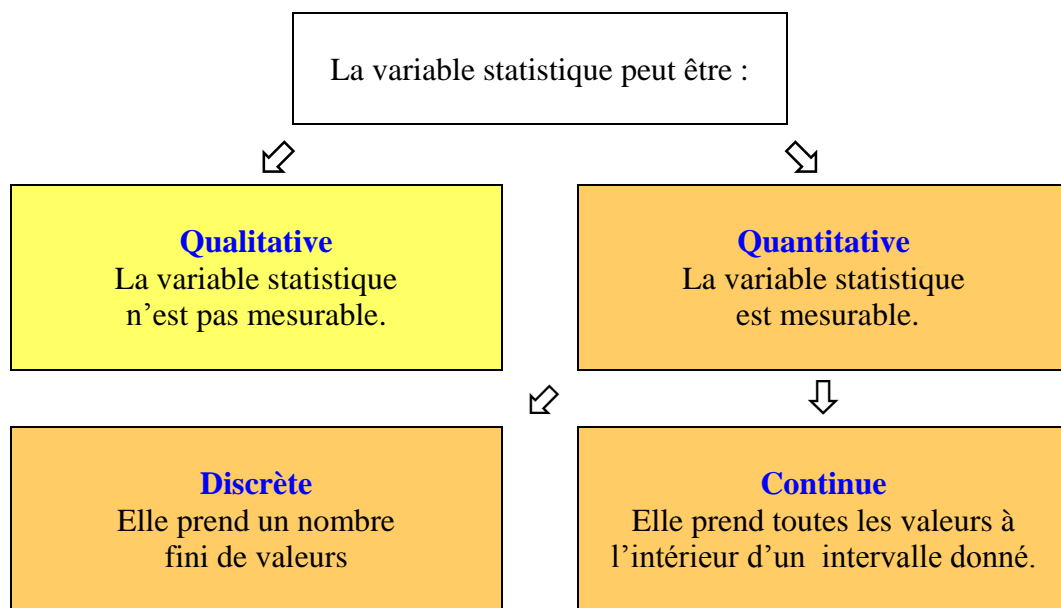
STATISTIQUES

1) Le vocabulaire utilisé en statistiques

L'ensemble sur lequel porte l'étude statistique est nommé **population** (exemple : l'ensemble des élèves d'un lycée, l'ensemble des pièces fabriquées, l'ensemble des trajets journaliers des élèves d'un LP).

Chaque élément de la population étudiée est : une **unité statistique** ou un **individu** (élève, pièce fabriquée, trajet journalier)

Le **caractère** ou **variable statistique** d'une population est la propriété sur laquelle porte l'étude statistique.



L'étude statistique d'une population par rapport à une variable continue impose de regrouper le grand nombre de valeurs en tranches ou **classes**. (classes d'âge pour une population de personnes ; classes du montant des achats pour une population de clients d'une grande surface.)

Une classe, c'est la portion de l'intervalle auquel appartiennent les valeurs de caractère.

Une série statistique associe à chaque valeur x_i du caractère le nombre d'individus correspondant, appelé **effectif partiel** et noté n_i . L'**effectif total** de la population est noté N .

La **fréquence** d'une valeur x_i du caractère est le quotient de l'effectif n_i de ce caractère par l'effectif total N : $f_i = \frac{n_i}{N}$

Remarques : ♦ La somme des fréquences est égale à 1.

- ♦ Les fréquences sont souvent exprimées en pourcentage après multiplication par 100 du rapport $\frac{n_i}{N}$.



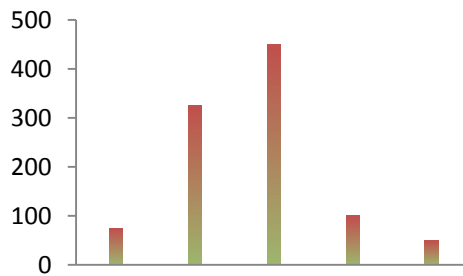
II) Différentes représentations graphiques

1) Diagramme en bâtons

On l'utilise pour les séries à caractère discret. Pour celles qui utilisent un repère cartésien :

- sur l'axe des abscisses : valeur du caractère ;
- sur l'axe des ordonnées : valeurs des effectifs ou des fréquences.

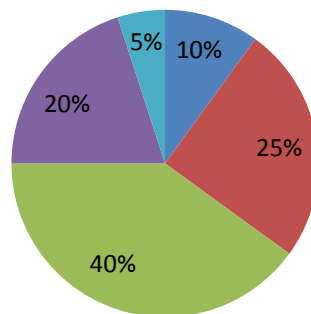
Les hauteurs des différents bâtons sont proportionnelles aux effectifs correspondants.



2) Diagramme à secteurs circulaires

On l'utilise souvent dans le cas d'une variable qualitative ou disposant de peu de valeurs.

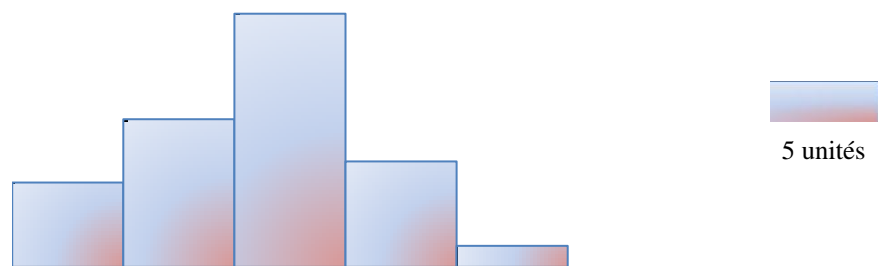
Chaque secteur a un angle au centre de mesure proportionnelle à la fréquence de la classe correspondante exprimée en pourcentage.



3) Histogramme

On l'utilise pour les séries à caractère quantitatif continu, lorsque les valeurs de la variable sont réparties en classes.

Les aires des différents rectangles sont proportionnelles aux effectifs (aux fréquences) correspondantes.





III) Indicateurs de tendance centrale

1) Calcul d'une moyenne d'une série distribuée en classes

On appelle moyenne d'une série statistique et on note \bar{x} le nombre :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{n_1 + n_2 + \dots + n_p} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N}$$

x_i désigne le centre de classe et N , l'effectif total

2) Médiane d'une série statistique

C'est la valeur (notée M_e) de la variable pour laquelle il existe, dans cette série, autant de valeurs plus grandes que de valeurs plus petites.

IV) Indicateurs de dispersion

1) Étendue

L'étendue est la différence entre la plus grande valeur et la plus petite valeur de la série.

2) Quartiles

Les trois quartiles sont les trois valeurs du caractère qui partagent la population totale en quatre parties d'effectifs égaux.

Le premier quartile Q_1 correspond à 25 % de l'effectif total.

Le deuxième quartile Q_2 correspond à la médiane (50 % de l'effectif total).

Le troisième quartile Q_3 correspond à 75 % de l'effectif total.

L'intervalle interquartile est la différence entre les quartiles extrêmes et a pour valeur $Q_3 - Q_1$.