

Vocabulaire abordé au cours du chapitre des STATISTIQUES

population

échantillon

unité statistique ou individu

caractère ou variable statistique

qualitatif - quantitatif

discret

continu

classe

diagramme en bâtons

diagramme à secteurs circulaires

histogramme

polygone des effectifs

mode et classe modale

effectif

effectifs cumulés croissants (décroissants)

fréquence

moyenne

médiane

quartile

décile

étendue

variance

écart-type

capabilité

distribution normale

Population : La population est l'ensemble des individus sur lequel portent les observations.

Échantillon : C'est la portion de population servant à l'étude.

Unité statistique ou individu : élément qui constitue la population.

Caractère : Le caractère est la propriété étudiée.

Le caractère est **qualitatif** s'il n'est pas une valeur numérique.

Le caractère est **quantitatif** s'il peut être mesuré : il est **quantitatif discret** s'il ne prend que des valeurs isolées. Il est **quantitatif continu** s'il peut prendre toutes les valeurs dans un intervalle donné.

Classe : portion de l'intervalle auquel appartiennent les valeurs de caractère.

Diagramme en bâtons : On l'utilise pour les séries à caractère discret. Les hauteurs des différents bâtons sont proportionnelles aux effectifs correspondants.

Diagramme à secteurs circulaires : On l'utilise le plus souvent dans le cas d'une variable discrète. Chaque secteur a un angle au centre de mesure proportionnelle à la fréquence de la classe correspondante exprimée en pourcentage.

Histogramme : On l'utilise pour les séries à caractère continu, lorsque les valeurs de la variable sont réparties en classes. Les aires des différents rectangles sont proportionnelles aux effectifs (aux fréquences) correspondantes.

Polygone des effectifs : est la ligne polygonale joignant les milieux de chaque côté supérieur des rectangles constituant l'histogramme.

Mode et classe modale : On appelle mode d'une série statistique à caractère discret la valeur du caractère statistique qui correspond au plus grand effectif.

On appelle classe modale d'une série statistique à caractère continu la classe qui correspond au plus grand effectif. Le mode est le centre de la classe modale.

Effectif : L'effectif d'une classe est le nombre d'éléments de la classe, on le note n_i . L'effectif total (noté N) est la somme des effectifs de chaque classe.

Le polygone des effectifs cumulés croissants est la ligne polygonale joignant les points ayant :

- pour abscisse : la borne supérieure de la classe.
- pour ordonnées : l'effectif cumulé croissant de la classe.

Fréquence : La fréquence de la n -ième classe est :

$$f_i = \frac{n_i}{N} \text{ (elle peut être exprimée en \%) ; } \sum f_i = 1$$

La médiane est la valeur du caractère qui partage l'effectif total en deux parties de même effectifs.

Dans le cas d'un caractère continu la médiane peut-être recherchée par lecture sur le polygone des effectifs cumulés. C'est l'abscisse du point ayant pour ordonnée $N/2$.

L'étendue d'une série statistique est la différence entre la plus grande et la plus petite valeur du caractère.

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

$$\bar{x} = \frac{\sum_{i=1}^{i=p} x_i n_i}{N} = \sum_{i=1}^{i=p} f_i x_i$$

x_i désigne le centre de classe.

Quartile : 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.

Décile : chaque décile partage en dix parties égales l'effectif total.

Variance : La variance V est la moyenne des carrés des écarts à la moyenne.

$$V = \frac{1}{N} \sum_{i=1}^{i=p} n_i (x_i - \bar{x})^2 = \left(\frac{1}{N} \sum_{i=1}^{i=p} n_i x_i^2 \right) - \bar{x}^2$$

Avec N : effectif total

x_i : valeur de la variable

n_i : effectif de la variable x_i

$\bar{x}$: moyenne de la série

Écart-type : L'écart-type σ (lire : sigma) est la racine carrée de la variance : $\sigma = \sqrt{V}$

Capabilité : La capabilité d'un processus, ou d'une machine, est l'aptitude de ce processus, ou de cette machine, à fabriquer des pièces bonnes. En appelant T_s et T_i respectivement limite supérieure et limite inférieure de la tolérance :

$$C_{m_k} = \frac{T_s - \bar{x}}{3\sigma} \text{ et } C_{m_k} = \frac{\bar{x} - T_i}{3\sigma}$$

Distribution normale : De nombreuses séries statistiques dont l'effectif est important ont une population distribuée suivant une loi dite normale avec une courbe des effectifs appelée courbe de Gauss.

Dans une loi normale, valeur moyenne, valeur médiane, valeur modale, sont égales.

Pour une série statistique « normalement » distribuée, il y a environ :

- 68 % de la population dans l'intervalle $[\bar{x} - \sigma ; \bar{x} + \sigma]$
- 95 % de la population dans l'intervalle $[\bar{x} - 2\sigma ; \bar{x} + 2\sigma]$
- 99% de la population dans l'intervalle $[\bar{x} - 3\sigma ; \bar{x} + 3\sigma]$