Statistiques descriptives

Vocabulaire

Population : ensemble d'individus

Echantillon : partie de la population

Caractère : caractéristique qui définit les individus d'une population, et dont les valeurs sont différentes d'un individu à un autre de la population ; quantitatif, si ses valeurs sont numériques ; qualitatif, si ses valeurs ne sont pas numériques

Série quantitative discrète : $(x_i; n_i)$ avec valeurs, x_i; effectifs, n_i,

Effectif total :

$$n = \sum_{i=1}^{p} n_i = n_1 + n_2 + \cdots + n_p$$

Fréquence de x_i:

$$f_i = \frac{n_i}{n} \text{ avec } f_1 + f_2 + \ldots + f_p = 1$$

Moyenne :

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \cdots + n_p x_p}{n} = \sum_{i=1}^{p} f_i x_i = f_1 x_1 + f_2 x_2 + \cdots + f_p x_p$$

Variance : mesure du degré de dispersion des valeurs autour de la moyenne

$$V = \frac{n_1 (x_1 - \bar{x})^2 + \cdots + n_p (x_p - \bar{x})^2}{n}$$

Ecart-type :

$$\sigma = (s =) \sqrt{V} \text{ donc } V = \sigma^2 \ (= s^2)$$

Médiane : valeur centrale si n est impair, ou demie-somme des deux valeurs centrales si n est pair.

Q_1 : plus petite valeur de la série telle qu'au moins 25% des valeurs de la série lui sont inférieures ou égales

Q_3 : plus petite valeur de la série telle qu'au moins 75% des valeurs de la série lui sont inférieures ou égales

Ecart interquartile : différence $Q_3 - Q_1$

Etendue : différence entre la plus grande et la plus petite valeur de la série.
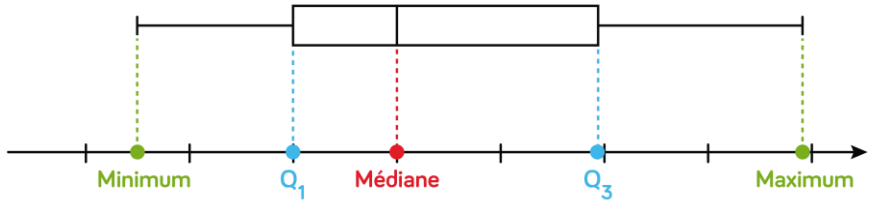
Diagramme en boîte : représentation graphique d'une série qui facilite la lecture de ses paramètres