Statistiques descriptives

Position (**moyenne, médiane, quartiles**) et dispersion (**écart-type, écart interquartile**) : deux familles d'indicateurs à ne pas confondre.

Définition — quartiles

Le **premier quartile** Q_1 est la plus petite valeur de la série telle qu'au moins $25\ \%$ des valeurs lui soient inférieures ou égales ; le **troisième quartile** Q_3 est la plus petite valeur telle qu'au moins $75\ \%$ des valeurs lui soient inférieures ou égales. L'**écart interquartile** est $Q_3 - Q_1$: il contient la moitié centrale des valeurs.

Méthode — calculer Q_1 et Q_3

Série ordonnée de N valeurs : Q_1 est la valeur de rang $\lceil N/4 \rceil$, Q_3 celle de rang $\lceil 3N/4 \rceil$ (on arrondit à l'entier supérieur). Pour $N = 10$: rangs 3 et 8.

Définition — variance, écart-type

$V = \frac{1}{N}\sum n_i (x_i - \overline{x})^2$ et $\sigma = \sqrt{V}$. L'écart-type mesure la dispersion autour de la moyenne ; il a la même unité que les données.

Propriété — linéarité de la moyenne

Si on ajoute b à toutes les valeurs, la moyenne augmente de b ; si on multiplie par a, la moyenne est multipliée par a et l'écart-type par $|a|$.

Définition — diagramme en boîte

Représente sur un axe : minimum, Q_1, médiane, Q_3, maximum. La boîte va de Q_1 à Q_3, les « moustaches » vont jusqu'aux extrêmes.

À retenir

La moyenne et l'écart-type sont sensibles aux valeurs extrêmes ; la médiane et l'écart interquartile ne le sont pas.

Questions fréquentes sur ce chapitre

Comment calculer les quartiles ?

On classe les valeurs ; Q_1 est la plus petite valeur telle qu'au moins $25\ \%$ des valeurs lui soient inférieures ou égales, Q_3 la plus petite telle qu'au moins $75\ \%$ le soient.

Qu'est-ce que l'écart interquartile ?

$Q_3 - Q_1$: la largeur de l'intervalle qui contient la moitié centrale des valeurs. C'est un indicateur de dispersion.

Que mesure l'écart-type ?

La dispersion des valeurs autour de la moyenne : plus il est grand, plus les valeurs sont éloignées de la moyenne.

Quelle est la différence entre indicateur de position et de dispersion ?

Moyenne, médiane et quartiles situent la série (position) ; étendue, écart interquartile et écart-type mesurent son étalement (dispersion).