Le cours – 2nd : Statistiques

Moyenne :

Cours

Valeur	x_1	x_2	…	x_p
Effectif	n_1	n_2	…	n_p

On note $N = n_1 + n_2 + \ldots + n_p$ l'**effectif total**.
La **moyenne** notée $\bar{x}$ est définie par :

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_p x_p}{N}$$

Si on note f_i la **fréquence** de la valeur x_i, on a $f_i = \frac{n_i}{N}$
et $\bar{x} = f_1 x_1 + f_2 x_2 + \ldots + f_p x_p$

Exemple

Valeur	1	2	4	5
Effectif	2	5	9	4

$N = 2+5+9+4 = 20$

$$\bar{x} = \frac{2\times1+5\times2+9\times4+4\times5}{20} = \frac{68}{20} = 3{,}4$$

$$f_1 = \frac{2}{20} = \frac{1}{10},\ f_2 = \frac{5}{20} = \frac{1}{4},\ f_3 = \frac{9}{20},\ f_4 = \frac{4}{20} = \frac{1}{5}$$

$$\bar{x} = \frac{1}{10}\times1+\frac{1}{4}\times2+\frac{9}{20}\times4+\frac{1}{5}\times5 = 3{,}4$$

Linéarité :
Si on **multiplie** toutes les valeurs par une même constante k, sans changer les effectifs, la moyenne est multipliée par k.
Si on **ajoute** à toutes les valeurs une même constante b, sans changer les effectifs, la moyenne est augmentée de b.

Variance et écart-type :

La **variance** d'une série statistique est le réel noté V, défini par :

$$V = \frac{n_1(x_1-\bar{x})^2 + n_2(x_2-\bar{x})^2 + \ldots + n_p(x_p-\bar{x})^2}{N}$$

L'écart-type d'une série statistique est le réel noté σ et défini par $\sigma = \sqrt{V}$.

$$V = \frac{2(1-3{,}4)^2+5(2-3{,}4)^2+9(4-3{,}4)^2+4(5-3{,}4)^2}{20}$$

$$V = \frac{34{,}8}{20} = 1{,}74.$$

L'écart-type est $\sigma = \sqrt{1{,}74} \simeq 1{,}32$

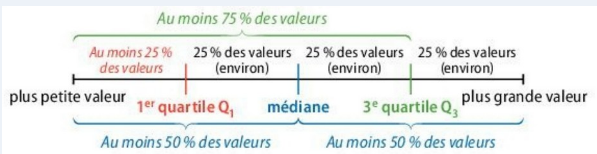
L'écart-type mesure la **dispersion** des valeurs autour de la moyenne : plus il est grand, plus les valeurs sont dispersées.

Médiane, quartiles et écart interquartile

Pour une série statistique d'effectif total N et dont les valeurs sont ordonnées par ordre croissant :
Si N est impair, la **médiane** est la valeur du terme de rang $\frac{N+1}{2}$. Si N est pair, elle est la valeur du terme de rang $\frac{N}{2}$.

Le rang du **1er quartile** Q_1 est le plus petit entier supérieur ou égal à $\frac{N}{4}$. Au moins 25 % des valeurs lui sont inférieures.

Le rang du **3e quartile** Q_3 est le plus petit entier supérieur ou égal à $\frac{3N}{4}$. Au moins 75 % des valeurs lui sont inférieures.



L'écart interquartile d'une série est la différence $Q_3 - Q_1$

Les méthodes – 2nd : Statistiques