

STATISTIQUES

Définitions : Série Statistique

- Lors d'un relevé de mesures effectué sur les individus d'une population, l'ensemble des données collectées constitue une série statistique.
- Le tableau d'effectif d'une série statistique à caractère quantitatif est constitué de deux lignes :
 - les différentes valeurs prises par le caractère étudié, rangé dans l'ordre croissant ;
 - les effectifs correspondants à chaque valeur de la première ligne.
- L'effectif total d'une série statistique est la somme des effectifs.

Dans la suite du chapitre, on considère une série statistique à caractère quantitatif pour laquelle le caractère étudié possède p valeurs différentes notées x_1, x_2, \dots, x_p dont les effectifs correspondants sont n_1, n_2, \dots, n_p et dont l'effectif total est N .

Définitions : Moyenne et fréquence

- La moyenne pondérée de la série statistique, notée \bar{x} est donnée par la formule :

$$\frac{n_1 \times x_1 + n_2 \times x_2 + \dots + n_p \times x_p}{N}$$

- La fréquence d'apparition de la valeur x_k est

$$x_k = \frac{n_k}{N} \quad \text{ou} \quad \frac{n_k}{N} \times 100 \text{ (\%)}$$

Propriétés :

- La somme des fréquences d'une série statistique est égale à 1 ou 100%.
- Autre formule de calcul de la moyenne d'une série : $\bar{x} = f_1 \times x_1 + f_2 \times x_2 + \dots + f_p \times x_p$
- Si toutes les valeurs d'une série sont multipliées par un même nombre a , alors la moyenne de la nouvelle série est aussi multipliée par a .
- Si on ajoute un même nombre b aux valeurs d'une série, alors la moyenne de la nouvelle série est aussi augmentée de b .

Définitions : Variance et Écart type

- La variance d'une série, notée V , est la moyenne pondérée des carrés des écarts de chaque valeur à la moyenne de la série.

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

- L'écart-type d'une série, noté σ , est la racine carrée de la variance : $\sigma = \sqrt{V}$. On dit que l'écart-type mesure la dispersion de la série autour de sa moyenne. Si les valeurs de la série sont proches de la moyenne, alors la variance et l'écart-type sont faibles.

Définitions : Médiane

- Une série statistique à caractère quantitatif est dite ordonnée après que les valeurs collectées ont été rangées dans l'ordre croissant ou décroissant.
- L'étendue désigne l'écart entre la plus grande et la plus petite des valeurs prises par le caractère.
- Dans une série statistique ordonnée, la médiane partage les valeurs prises par le caractère en deux groupes de même effectif.

Méthode : calcul de la médiane

- Si l'effectif total est impair, une valeur restera entre les deux demi-groupes. C'est la médiane.
- Si l'effectif total est pair, n'importe quelle valeur comprise entre la dernière valeur du premier groupe et la première valeur du second groupe peut être considérée comme une médiane. Le plus souvent, la moyenne de ces deux valeurs est choisie comme médiane.

Définitions : Quartiles

- Le premier quartile, souvent noté Q_1 , d'une série statistique numérique est la plus petite valeur prise par le caractère telle qu'au moins 25% des valeurs lui soient inférieures ou égales.
- Le troisième quartile, souvent noté Q_3 , d'une série statistique numérique est la plus petite valeur prise par le caractère telle qu'au moins 75% des valeurs lui soient inférieures ou égales.
- L'intervalle $[Q_1 ; Q_3]$ est appelé intervalle interquartile.
- La différence $Q_3 - Q_1$ est appelée écart interquartile.

///

Exemple : On relève la pointure des pieds de 47 personnes. On présente cette série statistique sous la forme du tableau d'effectif ci-contre.

Pointure	37	38	39	40	41	42	43	44	45	
Effectif	4	5	8	7	6	10	4	2	1	47
Fréquence	0,09	0,11	0,17	0,15	0,13	0,21	0,09	0,04	0,02	1
Fréquence en %	9%	11%	17%	15%	13%	21%	9%	4%	2%	100%

La moyenne pondérée obtenue à partir du tableau des pointures, est d'environ 40,45.

La médiane est 40.

$$V = 4,12$$

$$\sigma = 2,03$$

37	37	37	37	38	38	38	38	38	39
39	39	39	39	39	39	39	40	40	40
40	40	40	40	41	41	41	41	41	41
42	42	42	42	42	42	42	42	42	42
43	43	43	43	44	44	45			

$$Q_1 = 39 \quad Q_3 = 42$$

L'intervalle interquartile est $[39; 42]$ et l'écart interquartile est $42 - 39 = 3$.