

I) Médiane

Définition :

La **médiane** M_e d'une série statistique est la valeur du caractère qui partage l'effectif total en 2 parties de même effectif.

Remarque :

- Si N est impair, M_e est la valeur centrale.
Si N est pair, M_e est la moyenne des 2 valeurs centrales.

Exemple :

On considère la série : 1 ; 5 ; 6 ; 6 ; 7 ; 8 ; 9 ; 10. L'effectif total est 8 il n'y a donc pas une unique valeur centrale mais deux. On choisit donc pour médiane la moyenne entre ces deux valeurs centrales, soit les valeurs de rang 4 et 5. La médiane est donc $\frac{6+7}{2} = 6,5$.

Moyenne VS médiane

La différence entre moyenne et médiane apparaît clairement par exemple dans certaines statistiques de population des communes. Dans la région Alsace sur les 946 communes, la population moyenne par commune est de 2044 habitants et la population médiane est de 778 habitants...



La moyenne est très supérieure à la médiane, car la plupart des communes sont faiblement peuplées, mais les villes comme Strasbourg, Mulhouse, Colmar et Haguenau contiennent près de 25 % de la population d'Alsace et rendent la moyenne plus importante.

II) Quartiles

Définition :

- Le premier quartile noté Q_1 de la série statistique est la plus petite valeur telle qu'au moins 25% des valeurs lui soient inférieures ou égales;
- le troisième quartile noté Q_3 de la série statistique est la plus petite valeur telle qu'au moins 75% des valeurs lui soient inférieures ou égales.

Détermination pratique :

On suppose la série ordonnée dans l'ordre croissant des valeurs du caractère. Soit N l'effectif total.

- Si $\frac{N}{4}$ est un entier alors Q_1 est la valeur de rang $\frac{N}{4}$ et Q_3 est la valeur de rang $\frac{3N}{4}$;
- si $\frac{N}{4}$ n'est pas un entier, alors Q_1 est la valeur dont le rang suit $\frac{N}{4}$ et Q_3 est la valeur dont le rang suit $\frac{3N}{4}$.

III) Moyenne pondérée

Définition :

La **moyenne** est donnée par :

$$\frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{N}$$

Remarque :

Dans le cas d'un **caractère continu** (donc de valeurs regroupées en **classes**), on prend pour x_i le centre $\frac{a_i+b_i}{2}$ des classes $[a_i; b_i[$.

Linéarité de la moyenne

Propriétés

- Si toutes les valeurs d'une série sont multipliées par une même valeur a , alors la moyenne de cette nouvelle série est aussi multipliée par a .
- Si l'on ajoute une même valeur b à toutes les valeurs d'une série, alors la moyenne de cette nouvelle série est augmentée de b .

IV Ecart-type

Définition :

L'écart-type s d'une série statistique est un **indicateur de dispersion** de cette série autour de la moyenne. Concrètement :

- plus l'écart-type est petit, plus les valeurs de la série sont concentrées autour de la moyenne, donc la série est homogène ;
- plus l'écart-type est grand, plus les valeurs de la série sont éloignées de la moyenne, donc la série est hétérogène ;

Le calcul de l'écart-type de la série définie par

Valeurs	x_1	x_2	x_p
Effectifs	n_1	n_2	n_p

avec $n = n_1 + n_2 + \dots + n_p$

est donné par :

$$s = \sqrt{\frac{n_1(x_1 - m)^2 + n_2(x_2 - m)^2 + \dots + n_p(x_p - m)^2}{n}}$$

Étendue_{rappel}

Définition :

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

V) L'écart interquartile

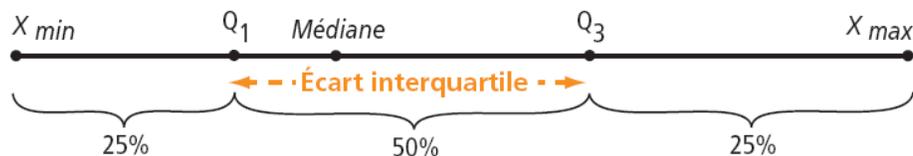
C'est un critère de dispersion de la série.

L'écart interquartile est $Q_3 - Q_1$.

Remarque :

L'écart interquartile d'une série mesure la dispersion autour de la médiane. Il contient au moins 50% des valeurs de la série. L'écart interquartile n'est pas influencé par les valeurs extrêmes de la série.

L'intervalle interquartile est $[Q_1 ; Q_3]$.



Choix des indicateurs

En général, on choisit un indicateur de position et un indicateur de dispersion :

- **Moyenne, étendue**

Le défaut du premier couple est qu'il est sensible aux valeurs extrêmes. On dit que ces paramètres ne sont pas robustes.

- **Médiane et intervalle interquartile**

Ce couple **médiane – intervalle interquartile** est plus robuste et est souvent représenté graphiquement par un **diagramme en boîte**, appelé parfois « **boîte à moustaches** » ou « **boîte à pattes** ».

Ce couple est moins sensible aux valeurs extrêmes.