

STATISTIQUES

www.mathsbook.fr

I - RAPPELS DE STATISTIQUES

1 - RAPPELS DE BASE ET LANGAGE STATISTIQUE

En langage statistique	En langage mathématiques	En langage courant
Population	Ensemble P	Ensemble des objets, des individus, mesures étudiées
Individu	Élément de P	Objet, individu, mesure
Caractère Caractère quantitatif Caractère qualitatif	Application f $f : P \rightarrow \mathbb{R}$ $f : P \rightarrow A(\text{Anoninclusdans}\mathbb{R})$	Aspect des objets étudiés Les données sont numériques Les données ne sont pas numériques
Modalité Caractère quantitatif discret Caractère qualitatif continu	Image par f des éléments de P $f(P)$ sont des valeurs isolées $f(P)$ est un intervalle	Valeur que peut prendre le caractère Valeurs isolées que l'on peut dénombrer Toute valeur d'un intervalle
Classe	Partie de l'ensemble des modalités	Regrouper des modalités
Effectif Effectif total Effectif cumulé croissant	n_i $N = n_1 + n_2 + \dots + n_k$ $n_1 + n_2 + \dots + n_p$	Nombre d'individu ayant la même modalité Nombre total d'objets Somme des effectifs
Mode <u>Caractère discret</u> Mode <u>Caractère continu</u> Classe modale Mode		Valeur du plus grand effectif Valeur du plus grand effectif Centre de la classe modale
Fréquence	$F = \frac{n_i}{N}$	Rapport : effectif d'une valeur par effectif total
Etendue		Différence entre les valeurs extrêmes

2 - MÉDIANE

Définition de la médiane : C'est la valeur du caractère qui permet de partager la population N en deux groupes de même effectifs. On distingue deux cas : celui d'un caractère quantitatif discret et celui d'un caractère quantitatif continu.

Cas d'un caractère quantitatif discret :

- Si N est impair : la médiane est la valeur du caractère observé au rang $\frac{N+1}{2}$.
- Si N est pair : la médiane n'est pas définie, mais on convient de prendre pour médiane la moyenne des caractères observés au rang $\frac{N}{2}$ et $\frac{N}{2} + 1$.

Cas d'un caractère quantitatif continu : on construit la courbe des fréquences cumulées et la médiane est l'antécédent de 0,5.

3 - MOYENNE ARITHMÉTIQUE PONDÉRÉE

Définition de la moyenne arithmétique pondérée : La moyenne arithmétique pondérée, que l'on note \bar{x} , est donnée par la formule suivante :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{n_1 + n_2 + \dots + n_k} = \frac{1}{N} \times \sum_{i=1}^{i=k} n_i x_i$$

Avec $N = n_1 + n_2 + \dots + n_k$ et n_i l'effectif de la valeur x_i .

II - PROPRIÉTÉS DE LA MOYENNE

Propriétés de la moyenne :

- Soit une série statistique (x_1, x_2, \dots, x_n) de N éléments, réparties en p sous groupes ayant respectivement n_1, n_2, \dots, n_p éléments (tels que $n_1 + n_2 + \dots + n_p = N$) et de moyennes respectives $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_p$.

Alors la série statistique a pour **moyenne** :

$$m = \frac{n_1\bar{y}_1 + n_2\bar{y}_2 + \dots + n_p\bar{y}_p}{N}$$

- Soient a et b deux réels fixés et une série statistique (x_1, x_2, \dots, x_n) de N éléments et de moyenne \bar{x} . Alors la série statistique (y_1, y_2, \dots, y_n) , avec $y_i = ax_i + b$ pour i variant de 1 à n , a pour **moyenne** :

$$\bar{y} = a\bar{x} + b$$

III - VARIANCE ET ÉCART TYPE

Définition de la variance : La **variance** de la série statistique est :

$$V = \frac{\sum_{i=1}^{i=n} n_i(x_i - m)^2}{N}$$

Avec :

$$N = \sum_{i=1}^{i=n} n_i$$

Définition de l'écart type : L'**écart type** de la série statistique est la racine carrée de la variance :

$$S = \sqrt{V}$$

IV - QUARTILE ET DÉCILE

Définitions : Soit une série statistique (x_1, x_2, \dots, x_n) de taille n .

- Les **quartiles** partagent la série statistique en quatre parties de même effectif.
Le premier quartile Q_1 est la plus petite valeur de la série telle qu'au moins 25% des données soient inférieures ou égales à Q_1 .
Le troisième quartile Q_3 est la plus petite valeur de la série telle qu'au moins 75% des données soient inférieures ou égales à Q_3 .
- L'**intervalle interquartile** est l'intervalle $[Q_1; Q_3]$.
- L'**écart interquartile** est $Q_3 - Q_1$.
- Les **déciles** partagent la série en 10 parties de même effectif.

V - DIAGRAMME EN BOÎTE

Définition d'un diagramme en boîte : Un **diagramme en boîte** permet de représenter une série statistique au moyen d'une boîte rectangulaire sur laquelle sont indiqués les informations suivantes :

- Le premier quartile Q_1 et le troisième quartile Q_3 ,
- Le premier décile D_1 et le troisième décile D_3 ,
- La médiane Me .

Voici sa représentation :

