

BTS – Statistiques à une variable

L'objectif de ce chapitre est l'étude de séries statistiques à une variable à l'aide notamment de deux couples d'indicateurs : le couple médiane / écart interquartile et le couple moyenne / écart type.

I Médiane et quartiles

Définition (Médiane d'une série discrète)

On appelle médiane de la série et on note Me tout nombre tel que :
au moins 50% des valeurs de la série sont inférieures ou égales à Me ,
et au moins 50% des valeurs de la série sont supérieures ou égales à Me .

Méthode de détermination

Si l'effectif total est impair, la médiane est la valeur centrale de la série ordonnée.

Si l'effectif total est pair, on choisit pour médiane la demi-somme des deux valeurs centrales de la série ordonnée.

Exercice 1

Déterminer la médiane.

1.	Valeurs x_i	35	36	37	38	39	40
	Effectifs n_i	3	4	2	1	1	2
	ECC						

2.	Valeurs x_i	5	11	27	33	42
	Effectifs n_i	4	3	1	5	3
	ECC					

Définition

On considère une série statistique à caractère quantitatif discret.

Le premier quartile Q_1 est la plus petite valeur de la série telle qu'au moins 25 % des valeurs 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 valeurs soient inférieures ou égales à Q_3 .

L'intervalle interquartile est $[Q_1; Q_3]$. Il contient environ 50 % des valeurs.

L'écart interquartile est le nombre $Q_3 - Q_1$.

Méthode pour déterminer les quartiles :

Soit N l'effectif total de la série. On range les N valeurs dans l'ordre croissant.

On calcule $\frac{N}{4}$, ce qui donne, en arrondissant si besoin à l'entier supérieur, le rang de Q_1 .

Pour Q_3 on procède de même en remplaçant $\frac{N}{4}$ par $\frac{3N}{4}$.

Remarque

La médiane est un indicateur de position.

L'écart interquartile est un indicateur de dispersion.

Exercice 2

Déterminer Q_1 et Q_3 .

Valeurs x_i	35	41	46	65	81
Effectifs n_i	5	3	2	3	2
Effectifs cumulés croissants					

Diagramme en boîte

Ce type de diagramme permet de visualiser facilement plusieurs indicateurs : Min, Q_1 , Me , Q_3 , Max.

Exercice 3

Voici un relevé des températures durant le mois de février :

Température observée	-4	-3	-2	0	2	3	4	5	6	7	8	9
Nombre de jours	1	2	1	1	1	1	2	3	4	5	4	3
Effectifs cumulés croissants												

1. Compléter les effectifs cumulés croissants. L'effectif total est $N = \dots$

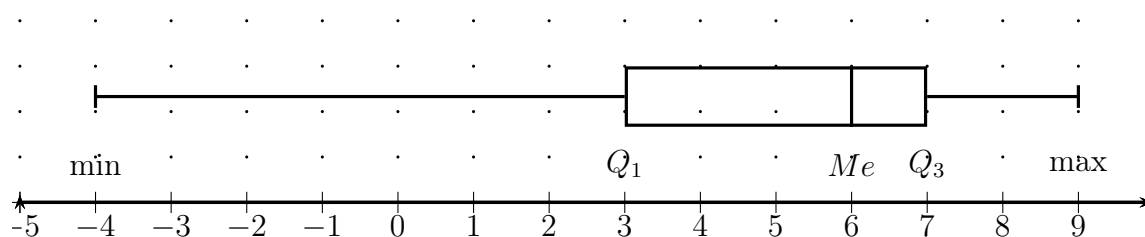
2. $\frac{N}{4} = \dots$, donc Q_1 est la $\dots^{\text{e}}$ valeur : $Q_1 = \dots$

3. $\frac{3N}{4} = \dots$, donc Q_3 est la $\dots^{\text{e}}$ valeur : $Q_3 = \dots$

4. Détermination de la médiane Me

.....

On obtient le diagramme en boîte (ou boîte à moustache) suivant :



Remarque

La droite graduée est indispensable.

Pour comparer des séries rapidement, on représente leur diagramme en boîte avec une seule droite graduée.

II Moyenne et écart type

Pour sa suite, on note $x_1, x_2, \dots, x_p$ les valeurs et $n_1, n_2, \dots, n_p$ les effectifs respectifs.

L'effectif total est alors $N = n_1 + n_2 + \dots + n_p$.

La fréquence de la valeur x_i est $f_i = \frac{n_i}{N}$.

Définition (moyenne)

La moyenne (pondérée) est le nombre noté $\bar{x}$ défini par :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N} \quad \text{ou} \quad \bar{x} = f_1x_1 + f_2x_2 + \dots + f_px_p$$

Définition

1. La variance de la série statistique est le nombre positif V défini par :

$$V = \frac{n_1(\bar{x} - x_1)^2 + n_2(\bar{x} - x_2)^2 + \cdots + n_p(\bar{x} - x_p)^2}{N} = \frac{1}{N} \sum_{i=1}^p n_i(\bar{x} - x_i)^2$$

2. L'écart-type est le nombre positif noté σ défini par $\sigma = \sqrt{V}$.

Exercice 4

On considère la série statistique suivante.

Valeurs x_i	7	11	13	15
Effectifs n_i	1	4	3	2

1. Calculer la moyenne $\bar{x}$ de cette série.
2. Déterminer la variance V de la série.
3. En déduire l'écart-type σ de la série.

Remarque

La moyenne est un indicateur de position.

L'écart type est un indicateur sur la dispersion des valeurs par rapport à la moyenne.

Plus précisément, plus l'écart-type est grand, plus les valeurs ont tendance à être éloignées de la moyenne.

III Utilisation de la calculatrice pour les statistiques

	Texas	Casio	Numworks
Entrer les données	Stat, Edit x_i dans L_1 n_i dans L_2	Menu, Statistiques x_i dans List1 n_i dans List2	Menu, Statistiques, données x_i dans V1 n_i dans N1
Obtenir les indicateurs	Stats, CALC, Stats 1-Var Xliste : L_1 ListeFreq : L_2	CALC, SET 1Var X list : List1 1Var Freq : List2	onglet Stats

Exercice 5

Voici un tableau présentant les salaires dans une entreprise.

Salaires mensuels (euros)	1200	1650	2100	2400	6500
Effectifs	5	8	4	3	1
Effectifs cumulés croissants					

1. Compléter les effectifs cumulés. Donner l'effectif total.
2. Déterminer, à l'aide de la calculatrice, la moyenne et l'écart-type, la médiane et les quartiles, l'étendue.
3. Que deviennent tous ces indicateurs si l'on remplace la plus grande valeur par 500 000 ?
Que peut-on dire de la moyenne dans ce dernier cas ?

Remarque

La moyenne et l'écart type sont sensibles aux valeurs extrêmes.

Il faut donc veiller à ce que ces valeurs extrêmes ne "faussent" pas trop le sens de ces indicateurs.