

CHAPITRE 11

Statistiques

I- Vocabulaire associé à une série statistique

L'objectif d'une étude statistique est d'obtenir, à partir de données numériques brutes en grand nombre, des renseignements qualitatifs et quantitatifs.

Présentation du vocabulaire utilisé à partir d'un exemple.

Note obtenue	1	2	3	4	5	6	7	8	9	10
Nombre d'élèves	2	0	1	3	8	5	2	0	3	1

Vocabulaire: On appelle: - **population** l'ensemble d'objets étudié d'un point de vue statistique.
- **individu** un élément de cette population.
- **caractère** l'attribut qui prend des **valeurs**, sur lequel porte l'étude statistique.
- **effectif** d'une valeur le nombre d'individus ayant cette valeur.

Remarque: Un caractère qui prend des valeurs isolées, comme dans l'exemple de référence, est dit **caractère discret**.

Dans le cas contraire, on dit que le **caractère** est **continu** (voir paragraphe III).

Définition 1: On appelle **fréquence** d'une valeur le quotient de l'effectif n de cette valeur sur l'effectif total N .

$$f = \frac{n}{N}$$

Remarque: En général, les fréquences sont données en pourcentage.

Toutes les données numériques sont consignées dans un même tableau.

Dans l'exemple:

Note	1	2	3	4	5	6	7	8	9	10
Effectif	2	0	1	3	8	5	2	0	3	1
Effectif cumulé croissant										
Effectif cumulé décroissant										
Fréquence en %										
Fréquence cumulée croissante en %										
Fréquence cumulée décroissante en %										

II- Représentation graphique et premiers indicateurs

La représentation graphique d'une série statistique à caractère discret s'appelle un **diagramme en bâtons**.

Indicateur 1: La moyenne

On multiplie la valeur du caractère par son effectif puis on ajoute les valeurs obtenues.
Le résultat est ensuite divisé par l'effectif total.

Indicateur 2: Le mode

C'est la valeur ayant le plus grand effectif.

Indicateur 3: La médiane

C'est la valeur M qui partage la série statistique en 2 groupes de même effectif tels que:
- tous les éléments du premier groupe ont des valeurs inférieures ou égales à M
- tous les éléments du second groupe ont des valeurs supérieures ou égales à M

(la médiane est la valeur à partir de laquelle la fréquence cumulée croissante dépasse 50 %)

Indicateur 4: L'étendue de la série

C'est la différence entre les deux valeurs extrêmes de la série.

III- Cas d'une série à caractère quantitatif continu

Etude à partir d'un exemple.

On mesure le nombre de voitures à la sortie d'une ville un jour de départ massif.

Heure h	$0 \leq h < 4$	$4 \leq h < 8$	$8 \leq h < 12$	$12 \leq h < 16$	$16 \leq h < 20$	$20 \leq h < 24$
Nombre de voitures	1100	3800	2100	700	1800	1200

Les valeurs prises par le caractère étudié sont regroupées en intervalles (appelés **classes**);
c'est ce qui caractérise une série statistique à caractère quantitatif continu.

La représentation graphique d'une série statistique à caractère continu s'appelle un **histogramme**.

Dans l'exemple:

Tableau de la série:

Heure	$0 \leq h < 4$	$4 \leq h < 8$	$8 \leq h < 12$	$12 \leq h < 16$	$16 \leq h < 20$	$20 \leq h < 24$
Effectif	1100	3800	2100	700	1800	1200
Effectif cumulé croissant						
Effectif cumulé décroissant						
Fréquence en %						
Fréquence cumulée croissante en %						
Fréquence cumulée décroissante en %						

Indicateur 1: La moyenne

Pour calculer la moyenne, on prend le centre de classe comme valeur du caractère, c'est à dire la moyenne des extrémités de chaque intervalle.

Indicateur 2: La classe modale

Indicateur 3: La médiane

En observant la ligne des fréquences cumulées croissantes, on observe qu'on dépasse les 50% de véhicules entre 8h et 12h. La médiane se situe donc dans cette classe.

Pour avoir une estimation plus précise de la médiane, on peut tracer le graphe représentant la fréquence cumulée en fonction de l'heure.

indicateur 4: L'étendue de la série

C'est la différence entre les deux valeurs extrêmes de la série.

Fin du chapitre 11
