Premières notions de statistiques

I. Vocabulaire statistique.

1. Etude statistique.

Une étude statistique porte sur une population composée d'individus. (personnes, animaux, objets...)

On étudie souvent un aspect particulier de cette population, appelé caractère.

Lorsque la population est nombreuse, on travaille sur un ou plusieurs échantillons.

Pour étendre les résultats collectés à l'ensemble de la population, il est important que l'échantillon soit représentatif et non biaisé.

2. Caractère.

Le caractère étudié peut être :

► quantitatif : c'est un caractère mesurable, on parle alors de variable (poids, taille, pointure d'un individu).

Cette variable peut être discrète : elle prend des valeurs telles que 0, 1, 2, 3.

Ou alors elle est continue : le caractère étudié peut prendre toutes les valeurs d'un intervalle.

On classera par exemple les tailles d'individus dans des différents intervalles appelés classes.

► qualitatif : c'est un caractère non mesurable. On parle alors plutôt de modalités (couleur des yeux ; activités sportives pratiquées).

3. Représentations graphiques.

On appelle effectif d'une valeur (ou modalité) le nombre d'individus de la population étudiée ayant cette valeur.

On appelle fréquence d'une valeur, notée f , le rapport

$$\frac{\text{effectif de la valeur}}{\text{effectif total}}$$

La fréquence est un nombre entre 0 et 1. Elle peut être mise en pourcentage en multipliant par 100.

► Diagramme circulaire.

Il sert à représenter en général un caractère qualitatif ou un caractère quantitatif discret.

Chaque secteur angulaire représentant une valeur ou modalité est proportionnel à l'effectif (ou fréquence) correspondant.

► Diagramme en bâtons.

Il sert à représenter en général un caractère qualitatif ou un caractère quantitatif discret.

La hauteur de chaque bâton représentant une valeur est proportionnelle à l'effectif correspondant.

► Histogramme

Il sert à représenter un caractère quantitatif continu.

L'aire de chaque rectangle est proportionnelle à l'effectif de la valeur correspondante.

Exemple : Petite enquête sur les Seconde Thalès.

Parmi les élèves de Seconde Thalès,

1. Pratiquez-vous une activité extra-scolaire ?

OUI	NON

2. Combien avez-vous de frères et sœurs ?

0	1	2	3	4	5

3. Quelle est votre taille ?

[1,5; 1,6[	[1,6; 1,65[	[1,65; 1,7[	[1,7; 1,75[	[1,75; 1,8[	[1,8; 2]

Représenter chacun des tableaux d'effectifs par un diagramme.

II. Caractéristiques d'une série statistique.

1. Caractéristiques de position.

a. Moyenne.

Soit une série statistique donnée par son tableau d'effectif.

Valeur	x_1	x_2	...	x_p
Effectif	n_1	n_2	...	n_p

On note N l'effectif total = $n_1 + n_2 + \dots + n_p$.

La **moyenne** de cette série est le réel, noté $\bar{x}$, tel que $\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N}$.

Si au lieu des effectifs, on a les fréquences : $\bar{x} = f_1x_1 + f_2x_2 + \dots + f_px_p$

Démonstration :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N} = \frac{n_1}{N}x_1 + \frac{n_2}{N}x_2 + \dots + \frac{n_p}{N}x_p = f_1x_1 + f_2x_2 + \dots + f_px_p.$$

b. Médiane.

Les N valeurs d'une série statistique sont rangés dans l'ordre croissant.

► Si N est impair ($N = 2q + 1$) alors la médiane est le terme de rang $q + 1$.

► Si N est pair ($N = 2q$) alors la médiane est la demi-somme des termes de rang q et $q + 1$.

Exemple :

On donne les notes à un test dans un groupe de 13 personnes :

8	9	9	9	10	10	11	11	13	15	15	16	18
---	---	---	---	----	----	----	----	----	----	----	----	----

Calculer la moyenne et la médiane de la série.

Notes à un devoir dans une classe de 18 élèves.

8	9	10	11	15	18	10	13	12
9	13	17	18	9	15	11	13	10

Même question : moyenne et médiane.

c. Mode et classe modale.

Dans le cas d'un caractère qualitatif ou quantitatif discret, on appelle mode la valeur du caractère qui a le plus grand effectif.

Dans le cas d'une série regroupée en classes, la classe modale est la classe ayant le plus grand effectif uniquement lorsque les classes sont de même amplitude.

2. Caractéristiques de dispersion.

Pour un caractère quantitatif, on appelle étendue la différence entre la plus grande et la plus petite valeur du caractère.

III. Propriétés de la moyenne.