

Chapitre 1

Statistiques descriptives

version élève

Sommaire du chapitre

I	Vocabulaire d'une étude statistique	2
II	Effectifs et fréquences	3
III	La moyenne	4
IV	La médiane	6
V	L'étendue	7
VI	Comparer deux séries statistiques	8
VII	Calculatrice et tableur	8
VIII	Bilan du chapitre	9



Objectifs du chapitre

Dans ce chapitre, on apprend à :

- identifier la population, les individus, le caractère et les modalités d'une étude statistique ;
- déterminer des effectifs, des fréquences et des effectifs cumulés croissants ;
- organiser les données dans un tableau ;
- calculer et interpréter une moyenne ;
- déterminer et interpréter une médiane ;
- calculer et interpréter l'étendue d'une série ;
- comparer deux séries statistiques à l'aide de leurs caractéristiques.

I Vocabulaire d'une étude statistique



Approche historique

Le mot « statistique » vient du latin status, qui désigne l'État. À l'origine, les statistiques servaient notamment à décrire la population et les ressources d'un pays. Aujourd'hui, elles sont utilisées dans les sciences, la médecine, l'économie, le sport et de nombreux domaines de la vie quotidienne.

Définition 1 (Vocabulaire fondamental)

Une étude statistique (statistical study) porte sur un ensemble appelé population (population).

- Chaque élément de cette population est un individu (individual).
- La propriété étudiée est appelée le caractère (variable).
- Les différentes valeurs prises par le caractère sont ses modalités (values / categories).

Un caractère peut être :

- qualitatif (qualitative) lorsqu'il n'est pas mesuré par un nombre ;
- quantitatif (quantitative) lorsqu'il est exprimé par un nombre.



Deux études statistiques

Exemple 1 : lancer d'un dé

- Population : les lancers effectués.
- Individu : un lancer.
- Caractère : le résultat obtenu.
- Modalités : 1, 2, 3, 4, 5 et 6.

Exemple 2 : enquête dans une classe

- Population : les élèves interrogés.
- Individu : un élève.
- Caractère : le nombre d'écrans dans son foyer.
- Modalités : 0, 1, 2, 3, etc.



Point de vigilance

Dans l'étude d'une série de lancers, les individus sont les lancers et non les faces du dé. La population doit toujours désigner l'ensemble des éléments effectivement observés.

II

Effectifs et fréquences

II.1

Définitions

Définition 2 (Effectif et fréquence)

- Dans une série statistique :
- l'effectif (frequency) d'une valeur x_i est le nombre de données égales à cette valeur ; on le note n_i ;
 - l'effectif total (total frequency) est le nombre total de données ; on le note N ;
 - la fréquence (relative frequency) d'une valeur x_i est le quotient de son effectif par l'effectif total ; on la note f_i .

$f_i = \frac{n_i}{N}$

avec

$N = n_1 + n_2 + \cdots + n_p.$

Propriété 1

Une fréquence est comprise entre 0 et 1. Elle peut être écrite sous forme décimale ou sous forme de pourcentage. La somme de toutes les fréquences est égale à 1, soit 100%.

 Résultats de 40 lancers

On lance 40 fois un dé cubique et on obtient les résultats suivants.

Valeurs x_i	1	2	3	4	5	6	Total
Effectifs n_i	5	10	5	7	3	10	40
Fréquences f_i	12,5%	25%	12,5%	17,5%	7,5%	25%	100%

Par exemple, la fréquence de la valeur 4 est :

$f_4 = \frac{\cdots}{\cdots}$

=

=

 Faux ami en anglais

En anglais, frequency désigne généralement l'effectif, tandis que relative frequency désigne la fréquence au sens français.

II.2 Effectifs cumulés croissants

Définition 3 (Effectif cumulé croissant)

Pour un caractère quantitatif dont les valeurs sont rangées dans l'ordre croissant, l'effectif cumulé croissant (cumulative frequency) d'une valeur x_i est la somme des effectifs correspondant aux valeurs inférieures ou égales à x_i .



Construction des effectifs cumulés

On reprend la série des 40 lancers.

Valeurs x_i	1	2	3	4	5	6
Effectifs n_i	5	10	5	7	3	10
Effectifs cumulés croissants	5	15	20	27	30	40

Par exemple :

$$15 = 5 + 10 \quad \text{et} \quad 27 = 5 + 10 + 5 + 7.$$

Il y a donc 27 résultats inférieurs ou égaux à 4.



À quoi servent les effectifs cumulés ?

Les effectifs cumulés permettent notamment de repérer rapidement le rang des données et de déterminer la médiane d'une série donnée sous forme de tableau.

III La moyenne

Définition 4 (Moyenne pondérée)

On considère une série dont les valeurs $x_1, x_2, \dots, x_p$ ont pour effectifs respectifs $n_1, n_2, \dots, n_p$. La moyenne (mean / average), notée $\bar{x}$, est donnée par :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{n_1 + n_2 + \dots + n_p}$$

soit encore :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N}$$



Calcul de la moyenne

Pour la série des 40 lancers :

$$\begin{aligned} \bar{x} &= \frac{5 \times 1 + 10 \times 2 + 5 \times 3 + 7 \times 4 + 3 \times 5 + 10 \times 6}{40} \\ &= \frac{\dots\dots\dots}{40} \\ &= \dots\dots\dots \end{aligned}$$

$$\bar{x} = \dots\dots\dots$$

Le résultat moyen d'un lancer est donc égal à 3,575.

**Interpréter une moyenne**

La moyenne n'est pas nécessairement une valeur observée dans la série. Ici, aucun lancer ne donne 3,575, mais cette valeur résume l'ensemble des résultats.

**Piège à éviter**

Lorsqu'une valeur apparaît plusieurs fois, il faut tenir compte de son effectif. On ne calcule donc pas ici $\frac{1 + 2 + 3 + 4 + 5 + 6}{6}$.

IV La médiane

Définition 5 (Définition de la médiane)

Une médiane (median) d'une série statistique ordonnée est une valeur qui partage la série en deux groupes :

- au moins 50% des valeurs lui sont inférieures ou égales ;
- au moins 50% des valeurs lui sont supérieures ou égales.

Méthode

On commence toujours par ranger les données dans l'ordre croissant.

- Si l'effectif total N est impair, la médiane est la valeur de rang $\frac{N+1}{2}$.
- Si l'effectif total N est pair, les deux valeurs centrales ont pour rangs $\frac{N}{2}$ et $\frac{N}{2} + 1$. Au collège, on choisit généralement pour médiane la moyenne de ces deux valeurs.

Effectif total pair

Dans la série des lancers, $N = 40$. Les deux valeurs centrales ont donc pour rangs :

$$\frac{40}{2} = 20 \quad \text{et} \quad \frac{40}{2} + 1 = 21.$$

Les effectifs cumulés montrent que :

- la valeur de rang 20 est 3 ;
- la valeur de rang 21 est 4.

On choisit donc :

$$M_e = \frac{\dots + \dots}{2}$$

$$= \dots$$

$$M_e = \dots$$

Effectif total impair

On considère la série ordonnée suivante :

4 ; 5 ; 5 ; 7 ; 8 ; 9 ; 12.

Elle contient $N = 7$ valeurs. La médiane est la valeur de rang :

$$\frac{7+1}{2} = \dots$$

La quatrième valeur est $\dots$, donc :

$$M_e = \dots$$



Moyenne ou médiane ?

La moyenne utilise toutes les valeurs et peut être fortement modifiée par une valeur extrême. La médiane dépend surtout de l'ordre des données et résiste mieux aux valeurs extrêmes.

v L'étendue

Définition 6 (Étendue)

L'étendue (range) d'une série statistique est la différence entre sa valeur maximale et sa valeur minimale :

$$\text{étendue} = \text{maximum} - \text{minimum}.$$



Calcul de l'étendue

Dans la série des lancers, la valeur minimale est 1 et la valeur maximale est 6.

$$e = \dots - \dots$$

$$= \dots$$

$$e = \dots$$



Interpréter l'étendue

L'étendue mesure la dispersion globale de la série. Plus elle est grande, plus les valeurs sont éloignées les unes des autres. Cependant, elle ne dépend que des deux valeurs extrêmes.

VI Comparer deux séries statistiques



Températures de deux villes

On relève les températures maximales, en degrés Celsius, pendant sept jours.

Ville A	15	16	16	17	17	18	20
Ville B	10	14	16	17	18	22	22

Les deux séries ont la même moyenne, égale à 17, et la même médiane, égale à 17. En revanche :

$$e_A = 20 - 15 = 5 \quad \text{et} \quad e_B = 22 - 10 = 12.$$

Les températures de la ville B sont donc plus dispersées que celles de la ville A.

Propriété 2

Pour comparer deux séries statistiques, une seule caractéristique ne suffit pas toujours. On peut comparer :

- leur position à l'aide de la moyenne et de la médiane ;
- leur dispersion à l'aide de l'étendue.

VII Calculatrice et tableur



Calculatrice

Pour calculer les caractéristiques d'une série à la calculatrice :

1. saisir les valeurs dans une liste ;
2. saisir les effectifs dans une seconde liste si les données sont regroupées ;
3. lancer le calcul statistique à une variable ;
4. repérer la moyenne $\bar{x}$, l'effectif total N , le minimum, la médiane et le maximum.

Les touches exactes dépendent du modèle utilisé. Le manuel de la TI-34 est disponible ici : [TI-34 – Statistiques](#).



Contrôle indispensable

La calculatrice fournit des résultats, mais elle ne remplace pas l'interprétation. Il faut toujours vérifier que l'effectif total et l'ordre de grandeur de la moyenne sont cohérents.

VIII

Bilan du chapitre

 Je dois savoir faire

- À la fin du chapitre, je dois savoir :
- utiliser correctement le vocabulaire d’une étude statistique ;
 - calculer un effectif total et une fréquence ;
 - compléter une ligne d’effectifs cumulés croissants ;
 - calculer une moyenne pondérée ;
 - déterminer une médiane lorsque l’effectif total est pair ou impair ;
 - calculer une étendue ;
 - interpréter et comparer les caractéristiques de deux séries ;
 - distinguer les faux amis anglais frequency et relative frequency.

 Lexique français–anglais

Français	English	Français	English
étude statistique	statistical study	effectif	frequency
population	population	effectif total	total frequency
individu	individual	fréquence	relative frequency
caractère	variable	effectif cumulé	cumulative frequency
modalité	value / category	moyenne	mean / average
caractère qualitatif	qualitative variable	médiane	median
caractère quantitatif	quantitative variable	étendue	range
valeur minimale	minimum value	valeur maximale	maximum value

↩️ Fin du cours ➡️