

# Chapitre 1

## Généralités sur la statistique

La statistique est l'étude de la collecte de données, leur analyse, leur traitement, l'interprétation des résultats et leur présentation afin de rendre les données compréhensibles par tous. C'est à la fois une science, une méthode et un ensemble de techniques.

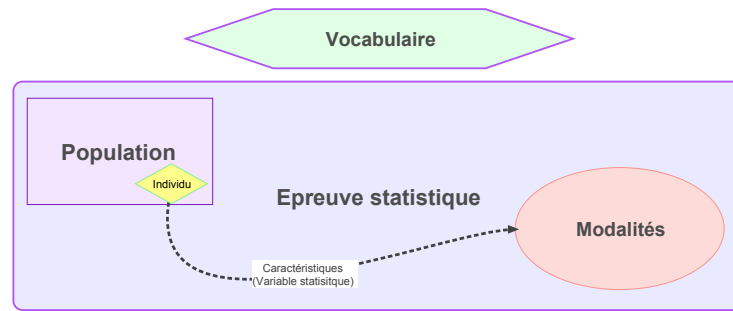
L'analyse des données est utilisée pour d'écrire les phénomènes étudiés, faire des prévisions et prendre des décisions à leur sujet. En cela, la statistique est un outil essentiel pour la compréhension et la gestion des phénomènes complexes.

Les données étudiées peuvent être de toute nature, ce qui rend la statistique utile dans tous les champs disciplinaires et explique pourquoi elle est enseignée dans toutes les filières universitaires, de l'économie à la biologie en passant par la psychologie et bien sûr les sciences de l'ingénieur. La statistique consiste à :

- Recueillir des données.
- Présenter et résumer ces données.
- Tirer des conclusions sur la population étudiée et d'aider à la prise de décision.
- En présence de données dépendant du temps, nous essayons de faire de la prévision.

### 1.1 Vocabulaire

Les statistiques consistent en diverses méthodes de classement des données tels que les tableaux, les histogrammes et les graphiques, permettant d'organiser un grand nombre de données. Les statistiques se sont développées dans la deuxième moitié du XIXe siècle dans le domaine des sciences humaines (sociologie, économie, anthropologie, ...). Elles se sont dotées d'un vocabulaire particulier.



### 1.1.1 Épreuve statistique

Les statistiques descriptives visent à étudier les caractéristiques d'un ensemble d'observations comme les mesures obtenues lors d'une expérience. L'expérience est l'étape préliminaire à toute étude statistique. Il s'agit de prendre "contact" avec les observations. De manière générale, la méthode statistique est basée sur le concept suivant.

#### Définition 1

*L'épreuve statistique est une expérience que l'on provoque.*

#### Exemple 1 (La durée de vie des lampes)

*Imaginons le cas suivant : un fabricant d'ampoules électriques ayant le choix entre 4 types de filaments se propose d'étudier l'influence de la nature du filament sur la durée de vie des ampoules fabriquées. Pour ce faire, il va faire fabriquer 4 échantillons d'ampoules identiques, sauf en ce qui concerne le filament, faire brûler les ampoules jusqu'à extinction, puis comparer les résultats obtenus.*

### 1.1.2 Population

En statistique, on travaille sur des populations. Ce terme vient du fait que la démographie, étude des populations humaines, a occupé une place centrale aux débuts de la statistique, notamment au travers des recensements de population. Mais, en statistique, le terme de population s'applique à tout objet statistique étudié, qu'il s'agisse d'étudiants (d'une université ou d'un pays), de ménages ou de n'importe quel autre ensemble sur lequel on fait des observations statistiques. Nous définissons la notion de population.

**Définition 2**

*On appelle population l'ensemble sur lequel porte notre étude statistique. Cet ensemble est noté  $\Omega$ .*

**Exemple 2**

- *On considère l'ensemble des étudiants de la section A. On s'intéresse aux nombre de frères et sœurs de chaque étudiant. Dans ce cas*

$\Omega =$  *ensemble des étudiants.*

- *Si l'on s'intéresse maintenant à la circulation automobile dans une ville, la population est alors constituée de l'ensemble des véhicules susceptibles de circuler dans cette ville à une date donnée. Dans ce cas*

$\Omega =$  *ensemble des véhicules.*

**1.1.3 Individu (unité statistique)**

Une population est composée d'individus. Les individus qui composent une population statistique sont appelés unités statistiques.

**Définition 3**

*On appelle individu tout élément de la population  $\Omega$ , il est noté  $\omega$  ( $\omega$  dans  $\Omega$ ).*

**Remarque 1**

*L'ensemble  $\Omega$  peut être un ensemble de personnes, de choses ou d'animaux...*

*L'unité statistique est un objet pour lequel nous sommes intéressés à recueillir de l'information.*

**Exemple 3**

- *Dans l'exemple indiqué ci-dessus, un individu est tout étudiant de la section.*
- *Si on étudie la production annuelle d'une usine de boîtes de boisson en métal (canettes). La population est l'ensemble des boîtes produites durant l'année et*

*une boîte constitue un individu.*

### 1.1.4 Caractère (variable statistique)

La statistique « descriptive », comme son nom l'indique cherche à décrire une population donnée. Nous nous intéressons aux caractéristiques des unités qui peuvent prendre différentes valeurs.

#### Définition 4

*On appelle caractère (ou variable statistique, dénotée  $V.S$ ) toute application*

$$X : \Omega \rightarrow C.$$

*L'ensemble  $C$  est dit : ensemble des valeurs du caractère  $X$  (c'est ce qui est mesuré ou observé sur les individus)*

#### Exemple 4

*Taille, température, nationalité, couleur des yeux, catégorie socioprofessionnelle ...*

#### Remarque 2

*Soit  $\Omega$  un ensemble. On appelle et on note  $\text{Card}(\Omega)$ , le nombre d'éléments de  $\Omega$ .*

$$\text{Card}(\Omega) := \text{nombre d'éléments de } \Omega = N.$$

### 1.1.5 Modalités

Les modalités d'une variable statistique sont les différentes valeurs que peut prendre celle-ci.

#### Exemple 5

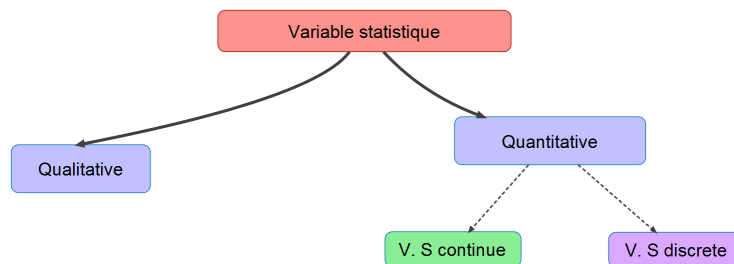
- Variable est " situation familiale "*
- Modalités sont " célibataire, marié, divorcé "*

- Variable est " statut d'interrupteur "  
Modalités sont " 0 et 1 ".
- Variable est " catégories socio-professionnelles "  
Modalités sont " Employés, ouvriers, retraités,... "

Les modalités sont les différentes situations dans lesquelles les individus peuvent se trouver à l'égard du caractère considéré.

## 1.2 Types des caractères

Nous distinguons deux catégories de caractères : les caractères qualitatifs et les caractères quantitatifs.



### 1.2.1 Caractère qualitatif

Les caractères qualitatifs sont ceux dont les modalités ne peuvent pas être ordonnées, c'est-à-dire que si l'on considère deux caractères pris au hasard, on ne peut pas dire de l'un des caractères qu'il est inférieur ou égal à l'autre. Plus précisément, nous avons la définition suivante.

#### Définition 5

*Les éléments de  $C$  sont représentés par autre chose que des chiffres.*

#### Exemple 6

*L'état d'une maison : on peut considérer les modalités suivantes*

- Ancienne.
- Dégradée.
- Nouvelle.

– *Rénovée.*

### 1.2.2 Caractère quantitatif

Les caractères quantitatifs sont des caractères dont les modalités peuvent être ordonnées. Ainsi, l'âge, la taille de vie ou le salaire d'un individu sont des caractères quantitatifs. Donc, nous avons la définition suivante.

#### Définition 6

*L'ensemble des valeurs est représenté par des chiffres. De même, il est partagé en deux sortes de caractères, discret et continu (voir l'exemple).*

#### Exemple 7

– *Le salaire d'employés d'une usine.*

*Modalités : 10000da , 20000da...*

*Type : Discret.*

– *La rigidité des ressorts.*

*Modalités : [10, 20] N/m*

*Type : continu.*

En général, la variable quantitative discrète est une variable ne prenant que des valeurs entières (plus rarement décimales). Le nombre de valeurs distinctes d'une telle variable est habituellement assez faible. Citons, par exemple, le nombre de maisons par quartier d'une ville. Une variable quantitative est dite continue lorsque les observations qui lui sont associées ne sont pas des valeurs précises, mais des intervalles. C'est le cas lorsque nous avons un grand nombre d'observations distinctes.

La statistique descriptive a pour objectif de synthétiser l'information contenue dans les jeux de données au moyen de tableaux, figures ou résumés numériques. Les variables statistiques sont analysées différemment selon leur nature (quantitative, qualitative).

## 1.3 Exercices corrigés

### Exercice 1

- La variable statistique "couleur de maisons d'un quartier" est-elle :

qualitative  quantitative

discrète  continue

La variable statistique "revenu brut" est-elle :

qualitative  quantitative

discrète  continue

La variable statistique "nombre de maisons vendues par ville" est-elle :

qualitative  quantitative

discrète  continue

**Solution :** Pour le premier cas, la variable statistique est qualitative. Pour le deuxième cas, la variable statistique est quantitative continue. Pour le troisième cas, la variable statistique est quantitative discrète.

### Exercice 2

- Parmi ces assertions, préciser celles qui sont vraies, celles qui sont fausses.

1. On appelle variable, une caractéristique que l'on étudie.
2. La tâche de la statistique descriptive est de recueillir des données.
3. La tâche de la statistique descriptive est de présenter les données sous forme de tableaux, de graphiques et d'indicateurs statistiques.
4. En Statistique, on classe les variables selon différents types.
5. Les valeurs des variables sont aussi appelées modalités.
6. Pour une variable qualitative, chaque individu statistique ne peut avoir qu'une seule modalité.
7. Pour faire des traitements statistiques, il arrive qu'on transforme une variable quantitative en variable qualitative.
8. La variable quantitative poids d'automobile peut être reclassée en compacte, intermédiaire et grosse.

9. *En pratique, lorsqu'une variable quantitative discrète prend un grand nombre de valeurs distinctes, on la traite comme continue.*

**Solution** : le corrigé en ordre est donné par

1. VRAI
2. FAUX
3. VRAI
4. VRAI
5. VRAI
6. VRAI
7. VRAI
8. VRAI
9. VRAI

### Exercice 3

- *Proposer des exemples de variable quantitative transformée en variable qualitative. Préciser les modalités de cette dernière.*

**Solution** : Les variables quantitatives dans le tableau ci-dessous peuvent être transformées en variable qualitative. Les modalités de cette dernière sont précisées dans la seconde colonne.

Variable quantitative	Modalités envisageables
Hauteur	Petit, Moyen, Grand
Poids	Très léger, Léger, Moyen, Lourd, Très lourd
Rendement	Faible, Moyen, Elevé
Chiffre d'affaire	Modéré, Moyen, Important, Très important
Cylindrée	Petite, Moyenne, Grosse

## 1.4 Exercices supplémentaires

### Exercice 4

- *Pour chacune des variables suivantes, préciser si elle est qualitative, quantitative discrète ou quantitative continue,*

(a) *Revenu annuel.*

(b) *Citoyenneté.*

(c) *Distance.*

(d) *Taille.*



(e) *Lieu de résidence.*

(g) *Couleur des yeux.*

(f) *Âge.*

(h) *Nombre de langues parlées.*

### Exercice 5

- *Pour les sujets d'étude qui suivent, spécifier : l'unité statistique, la variable statistique et son type,*

1. *Étude du temps de validité des lampes électriques.*
2. *Étude de l'absentéisme des ouvriers, en jours, dans une usine.*
3. *Répartition des étudiants d'une promotion selon la mention obtenue sur le diplôme du Bac.*
4. *On cherche à modéliser<sup>1</sup> le nombre de collisions impliquant deux voitures sur un ensemble de 100 intersections routières choisies au hasard dans une ville. Les données sont collectées sur une période d'un an et le nombre d'accidents pour chaque intersection est ainsi mesuré.*

### Exercice 6

- *Quelles formes de présentation de données correspondent ces propriétés ?*

1. *Il donne une bonne idée des données, mais on lui préfère en général les graphiques.*
2. *Il n'est pas nécessaire de lire des nombres. D'un simple coup d'œil, on a une vision d'ensemble des données.*

---

1. Fabriquer un modèle à partir des informations disponibles, généralement dans un but précis comme mieux appréhender un projet de construction ou se rendre compte de l'ampleur d'un sujet à partir d'une échelle réduite.