

Chapitre 0 : Introduction à la Statistique descriptive

2024/2025

Sommaire

- 1 Généralités
- 2 Définitions fondamentales
- 3 Les types de variables

1. Généralités

La **Statistique** est une science qui s'intéresse à l'étude de phénomènes complexes selon la démarche suivante :

- (i) observer et collecter des données ;
- (ii) traiter les données collectées pour en faire une synthèse (**Statistique Descriptive**) ;
- (iii) interpréter les résultats obtenus en (ii) et proposer des modèles théoriques (**Statistique Inférentielle**, en lien avec la théorie des probabilités).

L'objectif est de mieux comprendre les phénomènes observés, de les contrôler, de les influencer, de les prédire... La statistique apparaît ainsi comme un outil d'aide à la décision.

Des exemples ?

1. Généralités

La Statistique Descriptive regroupe des méthodes dont l'objectif est de **synthétiser** ou **décrire** les données observées.

Par synthétiser ou décrire les données, on entend **résumer** ces données :

- en les structurant (**tableaux**, **tri**),
- en produisant des graphiques (**résumés graphiques**),
- en calculant quelques indicateurs significatifs (**résumés numériques**).

Le but est de faire ressortir de ces données les informations les plus pertinentes (a priori cachées ou noyées dans la masse de données).

Des exemples ?

1. Généralités

Dans ce cours, nous allons présenter des *techniques* de la Statistique Descriptive qui permettent de déterminer les *caractéristiques* d'une *population*.

On parlera :

- de la notion de population et de la notion de variable,
- de résumés numériques (caractéristiques de tendance centrale, caractéristiques de dispersion),
- de résumés graphiques (histogrammes, boîtes à moustaches, nuages de points).

On ne parlera pas de la Statistique Inférentielle (extrapoler à toute la population les résultats obtenus sur un échantillon).

2. Définitions fondamentales

Définition 2.1

La **population** est l'ensemble sur lequel porte l'étude statistique. On note généralement Ω cet ensemble et N le cardinal de cet ensemble. Le nombre N est aussi appelé taille de la population.

Définition 2.2

Les éléments de la population sont appelés **unités statistiques** ou **individus**.

Remarque

Les individus sont de natures très diverses. La population peut être un ensemble de personnes, un ensemble de pièces produites par une usine, un ensemble de dates, etc.

2. Définitions fondamentales

Définition 2.3

On appelle **caractère** ou **variable** toute caractéristique observée sur les individus d'une population. Plus précisément, un caractère ou variable est une application $X : \Omega \rightarrow E$ qui, à tout individu ω de Ω , associe la valeur $X(\omega) \in E$.

$X(\omega)$ est la valeur de X observée sur l'individu ω . La valeur observée $X(\omega)$ est aussi appelée **observation**.

L'ensemble E est l'ensemble des valeurs possibles du caractère.

Exemples de caractères ou variables

- poids, taille, glycémie d'une personne
- nombre d'enfants, nombre de pièces dans un logement
- couleur des yeux, couleur des cheveux
- puissance, consommation d'un véhicule

2. Définitions fondamentales

Définition 2.4

Un **recensement** est l'opération qui consiste à recueillir des données sur tous les individus d'une population. Les données sont les valeurs des variables observées sur les individus.

Définition 2.5

Un **échantillon** est un sous-ensemble de la population sur lequel des données sont recueillies.

Définition 2.6

Un **sondage** est l'opération qui consiste à recueillir des données sur tous les individus d'un échantillon de la population.

3. Les types de variables

Définition 3.1

Le type d'une variable dépend de la nature des valeurs que peut prendre cette variable. Il existe deux grandes familles de variables :

- Les **variables quantitatives** dont les valeurs sont **numériques** ou **mesurables**.
- Les **variables qualitatives** dont les valeurs sont **non mesurables** mais décrites par des mots ou des phrases. Les valeurs de ce type de variable sont appelées **modalités**.

Exemples

Le poids ou la taille d'une personne, le nombre d'enfants, le nombre de pièces d'un logement, sont des variables quantitatives.

La couleur des yeux d'une personne, sa profession ou encore sa tranche d'âge, sont des variables qualitatives.

3. Les types de variables

Définition 3.2

Une **variable quantitative** est dite **discrète** si elle ne peut prendre que quelques valeurs numériques distinctes généralement entières.

Exemples

Le nombre d'enfants dans une famille, le nombre de pièces dans un logement, la puissance fiscale d'un véhicule.

Définition 3.3

Une **variable quantitative** est dite **continue** si elle peut prendre des valeurs appartenant à un intervalle de nombres réels.

3. Les types de variables

Exemples

Le poids, la taille ou encore le taux de glycémie sont des variables quantitatives continues.

Définition 3.4

Une **variable qualitative** est dite **nominale** si ses modalités ne peuvent pas être ordonnées (l'ordre n'a pas de sens).

Exemples

La couleur des yeux d'une personne, sa profession ou encore son signe astrologique sont des variables qualitatives nominales.

3. Les types de variables

Définition 3.5

Une **variable qualitative** est dite **ordinaire** si ses modalités peuvent être ordonnées (l'ordre a un sens).

Exemples

La tranche d'âge d'une personne (enfant, adolescent, adulte, aîné), la note d'un élève sur une échelle de A à E .