

# Présentation du fonctionnement et du déroulement du cours de Mathématiques et Statistique

## ✕ Déroulement général du cours.

Cet enseignement est un TD mais aussi des éléments de cours à revoir et/ou à apprendre. Les séances se décomposeront alors généralement de la façon suivante :

- $\frac{3}{4}$  h à 1h de cours selon les séances ;
- 1 petite pause de 5 à 10 minutes ;
- Le reste de la séance consacré aux exercices d'application.

## **Évaluation.**

- 2 contrôles continus. Le premier aura lieu lors de la 5<sup>ème</sup> séance et le deuxième lors de la dernière séance de TD.

Attendus pour ces examens :

- des exercices d'application à partir de ce qu'on aura vu en cours.
- des interprétations et analyses des résultats statistiques obtenus.
- des questions de cours pour voir si vous avez bien compris et retenu.

# Introduction, vocabulaire et éléments de base

- **Introduction**

- L'importance de la quantification en sociologie : un outil d'objectivation du social.

La statistique : une discipline qui se développe à la fin du XIXe siècle, en parallèle de la sociologie et sur laquelle s'appuie la sociologie. La statistique est un outil très utilisé en sociologie et ce, depuis la naissance de la discipline.

Le chiffre est un outil de connaissance du monde social.

**Durkheim** s'en est servi dans *Le Suicide*. Dans cet ouvrage, il montre que le suicide, qui constitue un acte intime, n'obéit pas seulement à des causes individuelles / psychologiques.

Il ne se produit pas par hasard. Il relève certaines régularités : le taux de suicide varie géographiquement (selon les sociétés) et, dans chaque société, il est plutôt stable.

Les hommes se suicident plus que les femmes, les non croyants plus que les croyants, etc.

Outre les motifs personnels, les variations du taux de suicide obéissent donc à une loi sociale. De son point de vue, la statistique constitue un outil pour objectiver le social, pour se détacher des prénotions.

**La statistique est un outil d'objectivation du social, des phénomènes sociaux** (mobilité sociale, inégalités hommes-femmes).

**C'est un outil d'administration de la preuve** : énoncer un fait avec une statistique à l'appui cela paraît plus probant que si ce même fait est énoncé à partir d'observations ou à partir d'un corpus de quelques dizaines d'entretiens.

Les méthodes quantitatives ont leurs défauts et leurs inconvénients mais un résultat semble souvent moins discutable s'il est fondé sur des statistiques

- **Qu'entend-on par statistique ?**

Il faut distinguer « statistiques » (au pluriel) et statistique (au singulier).

L'expression « les statistiques » désigne un ou plusieurs nombres relatifs à une population donnée. Ces nombres peuvent également résumer ou synthétiser un ensemble de données sur une population, pour les rendre plus intelligibles, plus compréhensibles.

- **Exemple1** : il est difficile de déceler une cohérence dans les 1500 votes lors d'une élection sans résumer ceux-ci.

- **Exemple 2 de données quantitatives qui synthétisent des informations sur la population** : selon l'INSEE, en 2022, l'âge moyen des mères à l'accouchement (donc quel que soit le rang de naissance de l'enfant) est 31,0 ans. Trente ans plutôt, soit en 1992, l'âge moyen à l'accouchement, était de 25,5 ans.
- On a besoin de ces résumés quantitatifs qui condensent l'information. Des institutions statistiques le font : **lesquelles connaissez-vous ?**

- **Le terme statistique (au singulier)** renvoie à **la science** qui a pour objectif de collecter des données sur une population, de les traiter, les analyser et d'en donner des interprétations.

On pourra par exemple calculer les pourcentages (calculer la proportion de filles et de garçons dans votre groupe de TD), les moyennes (l'âge moyen des étudiants du groupe). Il est aussi possible de faire des représentations graphiques qui, sous la forme de courbes ou de diagrammes, permettent de présenter l'information autrement.

- **Statistique univariée.**

Cette année, en L1, nous ferons de la **statistique descriptive**, plus précisément **univariée**. Autrement dit, nous ne travaillerons que sur une seule variable:

Exemple étudier l'âge des étudiant.e.s, ou étudier les ressources des étudiant.e.s .

En L2, nous entamerons la **statistique bivariée**, c'est-à-dire que nous travaillerons sur 2 variables. Nous allons par exemple croiser deux variables : voir par exemple, si le sexe influence le niveau de revenu.

- **I- Le vocabulaire ou terminologie**

1 - **La population**: c'est l'ensemble sur lequel porte l'étude que l'on réalise. Elle fait référence à l'ensemble des personnes visées par l'étude qui est menée.

Exemple 1: Une enquête porte sur les pratiques du numérique chez les étudiant.e.s inscrit.e.s en L1 à l'université de Tours en 2024-2025.

La population est celle **des étudiant.e.s** de L1 2024-2025, de l'université de Tours

Exemple 2: La population des voitures immatriculées en France en 2024

**2-Individu ou unité statistique:** c'est un élément quelconque de la population étudiée.

Exemple 1: .....est un individu de la population des étudiants du groupe .....de L1 de sociologie en 2024-2025

Exemple 2 :soit la population des personnes mariées au mois d'Août 2024 d'une commune d'Indre-et-Loire.

Un individu : un.e marié.e quelconque

**3. La taille de la population:** c'est l'effectif total de la population. Elle se note N.

Exemple1: La taille de la population française au 1<sup>er</sup> Janvier 2024 est:

N= 68 437 000 d'habitants

- Exemple2 : Prenons la population de notre groupe de TD...;

sa taille N = ..... Unité statistique : un.e étudiant.e quelconque de notre groupe: .....

**4. Un échantillon** : c'est une partie ou un sous-ensemble de la population.

Exemple: Considérons la population des étudiant.e.s de notre groupe de TD.

- Un échantillon est par exemple: l'ensemble des personnes de sexe féminin de notre groupe
- Un autre échantillon : les étudiant.e.s de 18 ans de notre groupe

**5. Variable statistique ou caractère statistique** : c'est une caractéristique commune à l'ensemble des individus de la population. Autrement dit, les variables statistiques sont des propriétés qui permettent de décrire les individus de la population. On arrive à trouver quelles sont les variables en répondant à la question : « sur quoi porte l'étude ? » La valeur d'une variable peut différer d'un individu à l'autre.

- Exemples : en considérant la population es étudiant.e.s inscrit.e.ss en sociologie à l'université de Tours en 2024-2025, **on peut étudier les variables suivantes** : leur âge, leur sexe, les professions et catégories socioprofessionnelles (PCS) (précédemment appelées Catégories Socio-Professionnelles CSP) de leurs parents, leur nombre de frères et soeurs, le type de logements qu'elles/ils occupent, leurs lieux d'habitation, leurs pratiques du numérique etc

- Autres exemples de variables: le statut matrimonial, les revenus, la nationalité, le mode de déplacement, la religion, la filière d'inscription à l'université de Tours en 2024-2025, les notes obtenues par les étudiant.e.s , le nombre d'enfants par ménage, l'âge des habitants d'une commune, etc

**NB:** Une variable est désignée ou notée par une lettre majuscule quelconque : par exemple X : le sexe ou Y : le revenu

- **6. Modalité d'une variable** : c'est l'ensemble des expressions possibles de la variable au sein de la population considérée.

**Exemples** : la variable statut matrimonial peut avoir pour modalités : marié, en couple, pacsé, célibataire, veuf, divorcé.

- La variable sexe peut avoir pour modalités : féminin, masculin, non binaire, autres.
- Le nombre de frères et sœurs peut avoir pour modalités : 0, 1, 2, 3 ....

Donnez vous-mêmes d'autres exemples de modalités pour les variables : le revenu, la nationalité, le mode de déplacement, la religion, la filière d'inscription à l'université des étudiant.es de Tours en 2024-2025

**7- Effectif relatif** de modalité d'une variable: c'est le nombre d'individus de la population correspondant cette modalité

Cet effectif relatif se note  $n_i$  et on lit n indice i ou  $n_i$

Exemple: Si on étudie la variable sexe, on peut avoir

Féminin 25 personnes =>  $n_1=25$

Masculin 15 personnes=>  $n_2 =15$

Remarque: La somme des effectifs relatifs ( $n_1$ ) donne l'effectif total ou taille de la population. L'effectif total se note **N**

$$N = n_1 + n_2 + n_3 + \dots + n_k$$

Sous forme de formule on écrit

$$N =$$

## **8. Les types de variables**

On distingue **deux types de variables: les variables qualitatives** et les **variables quantitatives**

### **8.1 Les variables qualitatives**

Ce sont des variables dont les modalités ne peuvent pas prendre une valeur numérique.

Autrement dit, leurs modalités ne sont que des **mots ou des expressions qui ne correspondent pas à des quantités numériques**

Exemple : La variable **statut matrimonial**.

**Ses modalités** peuvent être : célibataire, marié.e, en couple, pacsé.e, veuf ou veuve, divorcé.

Autres exemples: le sexe, la nationalité, la commune de naissance

## **8.2 Les variables quantitatives**

Les variables quantitatives sont celles dont les modalités sont mesurables, dénombrables, quantifiables.

Elles ne prennent donc que des valeurs numériques

**Exemple de variables quantitatives:** le revenu des employés d'une entreprise 2000€ ou 1587,5€ etc

,le nombre d'enfants des familles: 2 ou 0 ou 5,

la taille des élèves d'une classe: 1,56m ou 1,70 m

la distance parcourue entre le domicile et l'université: 2km, ou 3,5km  
etc

- On regroupe les valeurs prises par les variables quantitatives continues dans des intervalles ou classes du type  $[a; b[$

Exemple pour l'âge on a

$[0;5[$  ce qui regroupe tous les âges de 0 à 4,99ans, mais 5 ans est exclu.

**On distingue 2 types de variables quantitatives:** les variables quantitatives discontinues ou discrètes et les variables quantitatives continues

### **8.2.1 Les variables quantitatives discontinues ou discrètes**

Les variables quantitatives discontinues ou discrètes sont des variables dont les modalités ne prennent que des valeurs entières.

- Exemple1: La taille d'une famille est un exemple de variable discrète :  
une famille peut avoir 2, 3 ou 4 membres, mais ne peut pas avoir 1,7  
membres
- Exemple 2: le nombre d'appartements d'une agglomération
- A compléter...

## 8.2.2 Les variables quantitatives continues

Ce sont des variables dont les modalités peuvent **prendre aussi bien** des valeurs entières que des valeurs décimales

Exemple:

- les moyennes obtenues au semestre 1 en L1 Sociologie
- le salaire 3200 € 1526,78 €
- L'âge: 2 ans 18,5 ans
- La taille 1,78m 1,55m 2m

NB: regroupes sous formes d'intervalles [15; 20[

- Variable

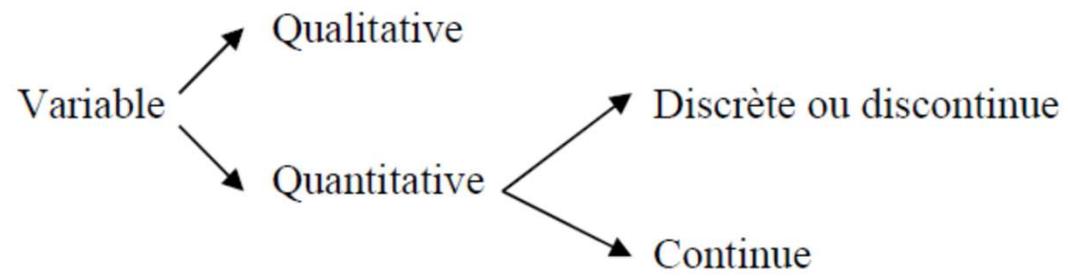
Qualitative

Discrète ou discontinue

Quantitative

Continue

Types de variables statistiques



## **9- Les échelles de mesures**

Les variables peuvent être classées sous une autre forme, celle de la façon dont elles se mesurent.

- Connaître à quelle échelle de mesure renvoie telle variable est là aussi nécessaire pour savoir quels sont les traitements statistiques qui sont autorisés avec elle, c'est-à-dire s'ils ont du sens.

- On distingue ainsi les variables **nominales, ordinales, métriques**.

### **9.1 L'échelle nominale.**

Les variables nominales sont simplement des classifications et leurs modalités ne peuvent être ordonnées.

Les échelles nominales correspondent à des variables qualitatives

**Exemple:** La répartition d'une population en Professions par Catégories Socioprofessionnelles (PCS) constitue un cas typique d'échelle nominale.

- Les différentes modalités des PCS (Agriculteurs exploitants, Artisans, commerçants et chefs d'entreprise, Cadres et professions intellectuelles supérieures, Employés etc), sont considérées comme égales les unes entre elles, et il **n'y a pas d'ordre entre elles. Les modalités n'ont pas de relation d'ordre entre elles.**
- Autres exemples, le sexe: il n'y a pas d'ordre entre Féminin et Masculin
- Donner des exemples....

## 9.2 L'échelle ordinale

Une variable ordinale est une variable dont **les modalités peuvent être ordonnées**. Les modalités ont une relation d'ordre entre elles.

-Exemple pour la variable « **la mention obtenue au bac** »

Mention Passable, Mention Assez Bien, Mention Bien, Mention Très Bien (il y a un ordre entre ces mentions)

- Un autre exemple de variable ordinale: **les échelles d'opinion.**

Quand on demande aux personnes si elles sont « pas du tout d'accord », « plutôt pas d'accord », « plutôt d'accord », « entièrement d'accord » avec une quelconque opinion. **Il y a un ordre sous-jacent.**

**Remarque:** on ne peut opérer aucune opération arithmétique ni avec les variables nominales, ni avec les variables ordinales

### 9.3 L'échelle métrique.

- L'échelle métrique correspond aux variables quantitatives. Les modalités sont mesurées numériquement (par des chiffres). La plupart des opérations statistiques sont autorisées à leur niveau.
- Exemple: le poids, le revenu, l'âge

## II. Les données construites et les tableaux de distribution

- Une fois qu'on a recueilli les données auprès de la population (à partir d'un questionnaire), on les réorganise et on va procéder à des comptages. Pour les organiser, on fait un **tri à plat**, qui correspond au comptage du nombre d'individus présentant chaque modalité de la variable. Ce faisant, on est en fait en train de calculer **les effectifs relatifs  $n_i$**

- **1. Tableau statistique.**

Le tableau statistique permet de rendre compte du dépouillement des données.

Il comprend évidemment des colonnes et des lignes.

- La première colonne, est celle de la variable et de ses modalités.

Chaque modalité

occupe une ligne du tableau.

- La deuxième colonne est celle des effectifs  $n_i$ .

- On peut ajouter une 3<sup>e</sup> colonne , celle des pourcentages et d'autres colonnes encore, selon les besoins

- On peut ajouter d'autres colonnes au tableau selon les besoins (fréquences, pourcentages, effectifs cumulés etc)

NB : Il faut toujours donner un titre au tableau et indiquer en bas la source des données.

Titre du tableau.

Variable	Effectifs	Pourcentages
Modalité 1 de la variable	nb de cas	
Modalité 2 de la variable	nb de cas	
.....	.....	
Total	N	100

Source du tableau.

## Exemple

- L'étude de la situation matrimoniale d'un groupe d'individus a permis, après le tri à plat, de construire le tableau statistique suivant :

Tableau : Répartition d'un groupe d'individus selon le sexe

<b>Sexe</b>	<b>Effectifs <math>n_i</math></b>	<b>Pourcentages</b>
Féminin	32	
Masculin	7	
Autres	1	
<b>Total</b>	<b>40</b>	<b>100</b>

*Source : L'auteur*

- **10. Les pourcentages**

Les effectifs relatifs  $n_i$  sont utiles pour résumer l'information, en revanche, **ils sont difficiles à interpréter**. La comparaison d'effectifs se complique s'ils ne sont pas basés sur le même nombre de cas.

- Le pourcentage indique, sur une base de 100, quelle partie de la population correspond à la modalité étudiée.
- On obtient le pourcentage par un calcul simple : une division et une multiplication. On divise l'effectif  $n_i$  par l'effectif total  $N$  puis on multiplie le résultat par 100.

$$\text{Pourcentage} = \frac{\text{Effectif relatif}}{\text{Effectif total}} \times 1000$$

- Sous forme de formule on écrit

$$\text{Pourcentage} = \frac{n_i}{N} \times 100$$

NB: le total des pourcentage est égale à 100