

# Comprendre et Appliquer le Test T de Student

Analyse Statistique et Interprétation

February 21, 2025

## 1 Introduction

Le test T de Student est un **test statistique** permettant de comparer les **moyennes** de deux échantillons pour déterminer si leur différence est **statistiquement significative**. Il est couramment utilisé en analyse de données et en sciences expérimentales.

## 2 Types de Test T

### 2.1 Test T à un échantillon

Ce test compare la **moyenne d'un échantillon** à une **valeur de référence** (moyenne théorique  $\mu$ ).

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

où :

- $\bar{X}$  est la **moyenne de l'échantillon**,
- $\mu$  est la **moyenne théorique** (valeur de référence),
- $s$  est l'**écart-type de l'échantillon**,
- $n$  est la **taille de l'échantillon**.

**Interprétation :** Si la **p-valeur** est inférieure à 0.05, alors la moyenne de l'échantillon est significativement différente de la moyenne théorique.

## 2.2 Test T pour échantillons indépendants

Ce test compare les moyennes de **deux groupes distincts**. Il est utilisé pour voir si une intervention (ex : vaccination) a eu un effet significatif.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

où :

- $\bar{X}_1, \bar{X}_2$  sont les **moyennes des deux groupes**,
- $s_1^2, s_2^2$  sont les **variances des deux groupes**,
- $n_1, n_2$  sont les **tailles des échantillons**.

**Interprétation :** Si  $p < 0.05$ , il y a une **différence significative** entre les groupes.

## 2.3 Test T pour échantillons appariés

Ce test est utilisé lorsque nous comparons **les mêmes individus avant et après une intervention** (ex : taux d'infection avant et après la vaccination).

$$t = \frac{\bar{D}}{\frac{s_D}{\sqrt{n}}}$$

où :

- $\bar{D}$  est la **moyenne des différences entre les paires de valeurs**,
- $s_D$  est l'**écart-type des différences**,
- $n$  est le **nombre d'observations**.

## 3 Exemple d'Application en Python

```
1 import pandas as pd
2 import numpy as np
3 import scipy.stats as stats
4
5 # Génération de données aléatoires
6 data = {
```

```

7      'date': pd.date_range(start='2020-06-01', periods=400,
8                               freq='D'),
9      'confirmed_cases': np.random.randint(100, 5000, 400)
10 }
11 df = pd.DataFrame(data)
12 df['date'] = pd.to_datetime(df['date'])
13
14 # Sparation des donn es avant et apr s 2021
15 df_pre_vaccine = df[df['date'] < "2021-01-01"]['confirmed_cases']
16 df_post_vaccine = df[df['date'] >= "2021-01-01"]['confirmed_cases']
17
18 # Test T pour comparer les p riodes
19 t_stat, p_value = stats.ttest_ind(df_pre_vaccine,
20                                   df_post_vaccine, equal_var=False)
21 print(f"T-statistique: {t_stat:.2f}, p-valeur: {p_value:.5f}")
22
23 if p_value < 0.05:
24     print("Diff rence significative entre les p riodes.")
25 else:
26     print("Aucune difference significative detectee.")

```

## 4 Interprétation des Résultats

Type de test	Hypothèse nulle $H_0$	Hypothèse alternative $H_A$	
Test T à un échantillon	$\bar{X} = \mu$	$\bar{X} \neq \mu$	Différence sig
Test T indépendant	$\bar{X}_1 = \bar{X}_2$	$\bar{X}_1 \neq \bar{X}_2$	Les c
Test T apparié	$\bar{D} = 0$	$\bar{D} \neq 0$	

Table 1: Résumé des différents tests T

## 5 Conclusion

Le test T de Student est un outil puissant permettant d'évaluer si une **différence observée** entre deux groupes est **statistiquement significative**. Il est très utilisé en **data science**, en médecine et en recherche scientifique.