# Tests statistiques
# Comprendre simplement avec des exemples et formules

Linda MATSING

## Introduction

Les **tests statistiques** servent à prendre des décisions à partir de données, souvent pour vérifier si une hypothèse sur une population est plausible ou non. Par exemple : est-ce que la moyenne d'un groupe est différente de celle d'un autre ?

## Concepts clés

— **Hypothèse nulle** $H_0$ : hypothèse de départ, souvent l'absence d'effet ou de différence.
— **Hypothèse alternative** $H_1$ : hypothèse contraire, celle qu'on cherche à prouver.
— **Statistique de test** : une valeur calculée à partir des données qui permet de décider.
— **Valeur-p (p-value)** : probabilité d'observer une statistique aussi extrême que celle obtenue si $H_0$ est vraie.
— **Seuil de signification** $\alpha$ : seuil choisi (souvent $5\% = 0.05$) sous lequel on rejette $H_0$.

## Comment ça marche ?

1. On formule $H_0$ et $H_1$. 2. On choisit le test statistique adapté. 3. On calcule la statistique de test sur les données. 4. On calcule la **valeur-p**. 5. Si $p \leq \alpha$, on rejette $H_0$ au profit de $H_1$. Sinon, on ne rejette pas $H_0$.

## Tests les plus courants

### 1. Test de la moyenne (Test t de Student)

On compare la moyenne d'un échantillon à une valeur connue ou entre deux groupes.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- $\bar{x}$ : moyenne observée - $\mu_0$ : moyenne hypothétique sous $H_0$ - $s$ : écart-type de l'échantillon - $n$ : taille de l'échantillon

Si $t$ est très grand ou très petit, on rejette $H_0$.

### 2. Test du khi-deux ($\chi^2$)

Utilisé pour tester l'indépendance entre deux variables catégorielles (ex : sexe et préférence).

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

- $O_i$ : fréquence observée - $E_i$ : fréquence attendue sous $H_0$
Une grande valeur de $\chi^2$ indique que $H_0$ (indépendance) est peu probable.

### 3. Test de corrélation

Vérifie s'il existe une relation linéaire entre deux variables continues. Coefficient de corrélation de Pearson $r$ :

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

- $r$ proche de 1 ou -1 : forte corrélation - $r$ proche de 0 : pas de corrélation
Un test peut évaluer si $r$ est significatif.

## Attention aux erreurs

— **Erreur de type I** : rejeter $H_0$ alors qu'il est vrai (faux positif). Probabilité $= \alpha$.
— **Erreur de type II** : ne pas rejeter $H_0$ alors qu'il est faux (faux négatif). Probabilité $= \beta$.

## Résumé rapide

| **Test** | **But** | **Quand l'utiliser ?** |
|---|---|---|
| Test t | Comparer une moyenne | Échantillon, données continues |
| Test $\chi^2$ | Tester indépendance | Variables catégorielles |
| Test corrélation | Relation entre variables | Deux variables continues |

## Exemple simple

Supposons qu'un professeur affirme que la moyenne d'une classe est de 15/20. On a un échantillon de 30 élèves avec une moyenne observée de 14 et un écart-type de 3.

$$t = \frac{14 - 15}{3/\sqrt{30}} = \frac{-1}{0.5477} \approx -1.825$$

On regarde la table de Student avec 29 degrés de liberté. Si la valeur-p est plus grande que 0.05, on ne rejette pas l'hypothèse que la moyenne est bien 15.

---

*Fiche conçue par Linda MATSING – pour comprendre les bases des tests statistiques simplement.*