



UNIVERSITÉ DE MONTPELLIER
Département de Mathématiques Appliquées

Outstanding : Tests d'hypothèses

Atelier Projet — HAX916X

Réalisé par :
DIALLO Ousmane
ATTOUMANI Ibrahim



STATISTIQUE
SCIENCE DES DONNÉES
UNIVERSITÉ DE MONTPELLIER

L'INSTITUT
agro Montpellier

Année Universitaire 2025 – 2026

Table des matières

1. Introduction	2
2. Test d'hypothèse : cadre général	3
2.1. Cas d'un test pur	3
2.2. Cas d'un test randomisé	4
3. Cas Gaussien	4
3.1. Test concernant m avec variance supposée connue	4
3.1.1. Test bilatéral	4
3.1.2. Test unilatéral	5
3.2. Test concernant m avec variance supposée inconnue	5
3.2.1. Test bilatéral	6
3.2.2. Test unilatéral	6
3.3. Test concernant σ^2 avec moyenne m supposée connue	6
3.3.1. Test bilatéral	7
3.3.2. Test unilatéral	7
3.4. Test concernant σ^2 avec moyenne m supposée inconnue	7
3.4.1. Test bilatéral	8
3.4.2. Test unilatéral	8
4. Test asymptotique dans le cas quelconque	9
4.1. Test bilatéral	9
4.2. Test unilatéral	9

1. Introduction

Nous allons illustrer un problème de décision consistant à trancher entre deux hypothèses contradictoires.

Dans une telle situation, il est impossible de choisir une hypothèse sans prendre un certain risque : on peut en effet commettre des erreurs.

Ces erreurs sont appelées **erreur de première espèce** et **erreur de seconde espèce**.

En général, on privilégie l'hypothèse H_0 (appelée *hypothèse nulle*) par rapport à H_1 (appelée *hypothèse alternative*).

En effet, il est souvent préférable de ne pas rejeter H_0 tant qu'on ne dispose pas d'assez de preuves contraires.

Par analogie avec le domaine judiciaire, il est préférable **d'innocenter un coupable que d'enfermer un innocent**.

Exemple : illustration du concept

On juge une personne et on formule les hypothèses suivantes :

H_0 : personne innocente et H_1 : personne coupable.

	Personne innocente (Θ_0)	Personne coupable (Θ_1)
Innocenter la personne	$1 - \alpha_\varphi$	(Erreur type 2) β_φ
Condamner la personne	(Erreur type 1) α_φ	$1 - \beta_\varphi$ (la puissance)

Cet exemple illustre que :

- une erreur de type 1 consiste à condamner une personne innocente (rejeter H_0 alors qu'il est vrai) ;
- une erreur de type 2 consiste à innocenter une personne coupable (accepter H_0 alors qu'il est faux).

Ainsi, dans la plupart des situations, on cherche à **contrôler le risque d'erreur de première espèce** (noté α), car il correspond à une décision trop audacieuse, potentiellement injuste.

2. Test d'hypothèse : cadre général

Soit $(\mathcal{X}, \mathcal{B}, (P_\theta)_{\theta \in \Theta})$ un modèle statistique.

où :

- \mathcal{X} est l'**espace d'observation** ;
- \mathcal{B} est la **tribu** (ou σ -algèbre) sur \mathcal{X} ;
- $(P_\theta)_{\theta \in \Theta}$ est une **famille de probabilités** sur $(\mathcal{X}, \mathcal{B})$.

On cherche à trancher entre les deux hypothèses suivantes :

$$(H_0) : \theta \in \Theta_0 \quad \text{et} \quad (H_1) : \theta \in \Theta_1$$

avec $\Theta_0 \cap \Theta_1 = \emptyset$.

2.1. Cas d'un test pur

On appelle **statistique de test pur** toute fonction $\varphi : \mathcal{X} \rightarrow \{0, 1\}$.

La règle de décision associée à φ est :

$$\text{décider } (H_1) \text{ si } \varphi(x) = 1, \quad \text{et } (H_0) \text{ si } \varphi(x) = 0.$$

On appelle **région critique du test** la zone qui conduit à décider (H_1) :

$$R_\varphi = \{\varphi = 1\}.$$

Deux types d'erreurs sont possibles :

- **Erreur de première espèce (type 1)** : décider (H_1) alors que (H_0) est vraie.

$$\alpha_\varphi(\theta) = P_\theta(R_\varphi) = E_\theta(\varphi), \quad \theta \in \Theta_0.$$

- **Erreur de deuxième espèce (type 2)** : décider (H_0) alors que (H_1) est vraie.

$$\beta_\varphi(\theta) = P_\theta(R_\varphi^c) = E_\theta(1 - \varphi), \quad \theta \in \Theta_1.$$

On définit aussi la **puissance du test** :

$$\eta_\varphi(\theta) = 1 - \beta_\varphi(\theta) = P_\theta(R_\varphi) = E_\theta(\varphi),$$

qui représente la probabilité de décider (H_1) à raison.

2.2. Cas d'un test randomisé

On appelle **statistique de test randomisée** toute fonction $\Phi : \mathcal{X} \rightarrow [0, 1]$.

La règle de décision associée à Φ est la probabilité d'accepter (H_1) quand H_0 est vraie.

Les erreurs sont alors :

$$\alpha_\Phi(\theta) = E_\theta(\Phi), \quad \theta \in \Theta_0 \quad \text{et} \quad \beta_\Phi(\theta) = E_\theta(1 - \Phi), \quad \theta \in \Theta_1.$$

La puissance du test est alors :

$$\eta_\Phi(\theta) = 1 - \beta_\Phi(\theta) = E_\theta(\Phi) = P_\theta(R_\Phi),$$

qui représente la probabilité de décider (H_1) à raison.

3. Cas Gaussien

On considère un échantillon i.i.d. Y_1, Y_2, \dots, Y_n de loi

$$Y_i \sim \mathcal{N}(m, \sigma^2),$$

avec $m \in \mathbb{R}$ et $\sigma > 0$.

L'objectif est de tester une hypothèse portant sur la moyenne m .

Deux cas sont possibles selon que la variance σ^2 est **connue** ou **inconnue**.

3.1. Test concernant m avec variance supposée connue

On suppose que $\sigma^2 = \sigma_0^2$ est connue.

Les tests seront alors basés sur la fonction pivotale suivante :

$$Z_n = \frac{\sqrt{n}(\bar{Y}_n - m_0)}{\sigma_0} = \frac{\sqrt{n}(\bar{Y}_n - m)}{\sigma_0} + \frac{\sqrt{n}(m - m_0)}{\sigma_0}$$

où $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$ et m_0 est la valeur de m sous (H_0) .

Sous (H_0) :

$$\bar{Y}_n \sim \mathcal{N}\left(m, \frac{\sigma_0^2}{n}\right) \quad \Rightarrow \quad Z_n \sim \mathcal{N}\left(\frac{\sqrt{n}(m - m_0)}{\sigma_0}, 1\right)$$

3.1.1. Test bilatéral

$$(H_0) : m = m_0 \quad \text{vs} \quad (H_1) : m \neq m_0$$

Sous (H_0) , la statistique Z_n suit une loi $\mathcal{N}(0, 1)$.

Sous (H_1) , la valeur absolue $|Z_n|$ tend à être plus grande que celle d'une variable normale centrée réduite.

Ainsi, la région critique du test (au risque $\alpha \in]0, 1[$) est :

$$R_\alpha = \left\{ (y_1, \dots, y_n) : |z_n| > u_{1-\frac{\alpha}{2}} \right\}$$

où $u_{1-\frac{\alpha}{2}}$ désigne le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0, 1)$.

3.1.2. Test unilatéral

- Cas à droite :

$$(H_0) : m \leq m_0 \quad \text{vs} \quad (H_1) : m > m_0$$

- Cas à gauche :

$$(H_0) : m \geq m_0 \quad \text{vs} \quad (H_1) : m < m_0$$

Sous (H_0) , on a toujours $Z_n \sim \mathcal{N}(0, 1)$.

Sous (H_1) , Z_n aura tendance à être plus grande (ou plus petite) qu'une variable normale centrée réduite.

La région critique (au risque α) est donnée par :

$$R_\alpha = \begin{cases} \{(y_1, \dots, y_n) : z_n > u_{1-\alpha}\}, & \text{si } (H_1) : m > m_0, \\ \{(y_1, \dots, y_n) : z_n < -u_{1-\alpha}\}, & \text{si } (H_1) : m < m_0. \end{cases}$$

3.2. Test concernant m avec variance supposée inconnue

Lorsque σ^2 est **inconnue**, on l'estime à partir de l'échantillon à l'aide de la variance empirique corrigée :

$$S_n'^2 = \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y}_n)^2$$

La statistique de test devient alors :

$$T_n = \frac{\sqrt{n}(\bar{Y}_n - m_0)}{S'_n} = \frac{\sqrt{n}(\bar{Y}_n - m)}{S'_n} + \frac{\sqrt{n}(m - m_0)}{S'_n}$$

Sous (H_0) , cette statistique suit une loi de Student à $(n-1)$ degrés de liberté :

$$T_n \sim \mathcal{T}(n-1)$$

où m_0 est la valeur de m sous (H_0) .

3.2.1. Test bilatéral

$$(H_0) : m = m_0 \quad \text{vs} \quad (H_1) : m \neq m_0$$

Sous (H_0) , la statistique T_n suit une loi $\mathcal{T}(n-1)$ Student à $n-1$ degrés de libertés.

Sous (H_1) , la valeur absolue $|T_n|$ tend à être plus grande que celle d'une variable de loi $\mathcal{T}(n-1)$.

Ainsi, la région critique du test (au risque $\alpha \in]0, 1[$) est :

$$R_\alpha = \left\{ (y_1, \dots, y_n) : |t_n| > t_{1-\frac{\alpha}{2}}(n-1) \right\}$$

3.2.2. Test unilatéral

- Cas à droite :

$$(H_0) : m \leq m_0 \quad \text{vs} \quad (H_1) : m > m_0$$

- Cas à gauche :

$$(H_0) : m \geq m_0 \quad \text{vs} \quad (H_1) : m < m_0$$

Sous (H_0) , on a toujours $T_n \sim \mathcal{T}(n-1)$.

Sous (H_1) , Z_n aura tendance à être plus grande (ou plus petite) qu'une variable de loi $\mathcal{T}(n-1)$.

La région critique (au risque α) est donnée par :

$$R_\alpha = \begin{cases} \{(y_1, \dots, y_n) : t_n > t_{1-\alpha}(n-1)\}, & \text{si } (H_1) : m > m_0, \\ \{(y_1, \dots, y_n) : t_n < -t_{1-\alpha}(n-1)\}, & \text{si } (H_1) : m < m_0. \end{cases}$$

3.3. Test concernant σ^2 avec moyenne m supposée connue

Lorsque $m = m_0$ est **connue**, on peut estimer la variance à partir de l'échantillon par :

$$S_n^2 = \frac{1}{n} \sum_{j=1}^n (Y_j - m_0)^2$$

La statistique de test pivotale est alors :

$$Z_n = \frac{nS_n^2}{\sigma_0^2} = \frac{\sigma^2}{\sigma_0^2} \frac{nS_n^2}{\sigma^2}$$

avec $\frac{nS_n^2}{\sigma^2} \sim \chi^2(n)$.

Ainsi, sous (H_0) , cette statistique suit une loi du χ^2 à n degrés de liberté :

$$Z_n \sim \chi^2(n)$$

où σ_0^2 est la valeur de σ^2 sous (H_0) .

3.3.1. Test bilatéral

$$(H_0) : \sigma^2 = \sigma_0^2 \quad \text{vs} \quad (H_1) : \sigma^2 \neq \sigma_0^2$$

Sous (H_0) , $Z_n \sim \chi^2(n)$ et sous (H_1) , Z_n aura tendance à être plus grande ou plus petite qu'une variable de loi $\chi^2(n)$.

La région critique du test bilatéral (au risque $\alpha \in]0, 1[$) est donnée par :

$$R_\alpha = \left\{ (y_1, \dots, y_n) : Z_n < \chi_{\frac{\alpha}{2}}^2(n) \quad \text{ou} \quad Z_n > \chi_{1-\frac{\alpha}{2}}^2(n) \right\}$$

3.3.2. Test unilatéral

- Cas à droite :

$$(H_0) : \sigma^2 \leq \sigma_0^2 \quad \text{vs} \quad (H_1) : \sigma^2 > \sigma_0^2$$

- Cas à gauche :

$$(H_0) : \sigma^2 \geq \sigma_0^2 \quad \text{vs} \quad (H_1) : \sigma^2 < \sigma_0^2$$

Sous (H_0) , $Z_n \sim \chi^2(n)$.

Sous (H_1) , la valeur de Z_n tend à être plus grande (cas à droite) ou plus petite (cas à gauche) que celle d'une variable de loi $\chi^2(n)$.

La région critique du test (au risque $\alpha \in]0, 1[$) est :

$$R_\alpha = \begin{cases} \{(y_1, \dots, y_n) : Z_n > \chi_{1-\alpha}^2(n)\}, & \text{si } H_1 : \sigma^2 > \sigma_0^2, \\ \{(y_1, \dots, y_n) : Z_n < \chi_\alpha^2(n)\}, & \text{si } H_1 : \sigma^2 < \sigma_0^2. \end{cases}$$

3.4. Test concernant σ^2 avec moyenne m supposée inconnue

Lorsque m est **inconnue**, on peut estimer la variance à partir de l'échantillon par :

$$S_n'^2 = \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y}_n)^2$$

La statistique de test pivotale est alors :

$$Z_n = \frac{(n-1)S_n'^2}{\sigma_0^2} = \frac{\sigma^2}{\sigma_0^2} \frac{(n-1)S_n'^2}{\sigma^2}$$

avec $\frac{(n-1)S_n'^2}{\sigma^2} \sim \chi^2(n-1)$.

Ainsi, sous (H_0) , cette statistique suit une loi du χ^2 à n degrés de liberté :

$$Z_n \sim \chi^2(n)$$

où σ_0^2 est la valeur de σ^2 sous (H_0) .

3.4.1. Test bilatéral

$$(H_0) : \sigma^2 = \sigma_0^2 \quad \text{vs} \quad (H_1) : \sigma^2 \neq \sigma_0^2$$

Sous (H_0) , $Z_n \sim \chi^2(n-1)$ et sous (H_1) , Z_n aura tendance à être plus grande ou plus petite qu'une variable de loi $\chi^2(n-1)$.

La région critique du test bilatéral (au risque $\alpha \in]0, 1[$) est donnée par :

$$R_\alpha = \left\{ (y_1, \dots, y_n) : Z_n < \chi_{\frac{\alpha}{2}}^2(n-1) \quad \text{ou} \quad Z_n > \chi_{1-\frac{\alpha}{2}}^2(n-1) \right\}$$

3.4.2. Test unilatéral

- Cas à droite :

$$(H_0) : \sigma^2 \leq \sigma_0^2 \quad \text{vs} \quad (H_1) : \sigma^2 > \sigma_0^2$$

- Cas à gauche :

$$(H_0) : \sigma^2 \geq \sigma_0^2 \quad \text{vs} \quad (H_1) : \sigma^2 < \sigma_0^2$$

Sous (H_0) , $Z_n \sim \chi^2(n-1)$.

Sous (H_1) , la valeur de Z_n tend à être plus grande (cas à droite) ou plus petite (cas à gauche) que celle d'une variable de loi $\chi^2(n-1)$.

La région critique du test (au risque $\alpha \in]0, 1[$) est :

$$R_\alpha = \begin{cases} \{(y_1, \dots, y_n) : Z_n > \chi_{1-\alpha}^2(n-1)\}, & \text{si } H_1 : \sigma^2 > \sigma_0^2, \\ \{(y_1, \dots, y_n) : Z_n < \chi_\alpha^2(n-1)\}, & \text{si } H_1 : \sigma^2 < \sigma_0^2. \end{cases}$$

4. Test asymptotique dans le cas quelconque

Considérons un échantillon i.i.d. Y_1, \dots, Y_n d'espérance m et de variance σ^2 .
Les tests seront basés sur le Théorème Central Limite (TCL) :

$$\frac{\sqrt{n}(\bar{Y}_n - m)}{\sigma} \approx \mathcal{N}(0, 1)$$

Ces tests sont donc **asymptotiques**, c'est-à-dire appropriés uniquement pour des tailles d'échantillon suffisamment grandes ($n \geq 30$).

Soit $\hat{\sigma}_n^2$ un estimateur tel que $\hat{\sigma}_n^2 \xrightarrow{\mathbb{P}} \sigma^2$. Alors, par le TCL et Slutsky :

$$Z_n := \frac{\sqrt{n}(\bar{Y}_n - m_0)}{\sqrt{\hat{\sigma}_n^2}} \xrightarrow{L} \mathcal{N}(0, 1)$$

où m_0 est la valeur de m sous (H_0) .

4.1. Test bilatéral

$$(H_0) : m = m_0 \quad \text{vs} \quad (H_1) : m \neq m_0$$

Sous (H_0) et pour n grand : $Z_n \sim \mathcal{N}(0, 1)$.

Sous (H_1) , $|Z_n|$ aura tendance à être plus grande que celle d'une variable de loi $\mathcal{N}(0, 1)$.

La région critique du test bilatéral (au risque $\alpha \in]0, 1[$) est donnée par :

$$R_\alpha = \left\{ (Y_1, \dots, Y_n) : |Z_n| > u_{1-\frac{\alpha}{2}} \right\}$$

4.2. Test unilatéral

- Cas à droite :

$$(H_0) : m \leq m_0 \quad \text{vs} \quad (H_1) : m > m_0$$

- Cas à gauche :

$$(H_0) : m \geq m_0 \quad \text{vs} \quad (H_1) : m < m_0$$

Sous (H_0) et pour n grand : $Z_n \sim \mathcal{N}(0, 1)$.

Sous (H_1) , Z_n aura tendance à être plus grande (cas à droite) ou plus petite (cas à gauche) qu'une variable de loi $\mathcal{N}(0, 1)$.

La région critique du test unilatéral (au risque $\alpha \in]0, 1[$) est :

$$R_\alpha = \begin{cases} \{(Y_1, \dots, Y_n) : Z_n > u_{1-\alpha}\}, & \text{si } H_1 : m > m_0 \text{ (cas à droite),} \\ \{(Y_1, \dots, Y_n) : Z_n < -u_{1-\alpha}\}, & \text{si } H_1 : m < m_0 \text{ (cas à gauche).} \end{cases}$$