Haute école de gestion de Genève
Geneva School of Business Administration

Test de la moyenne

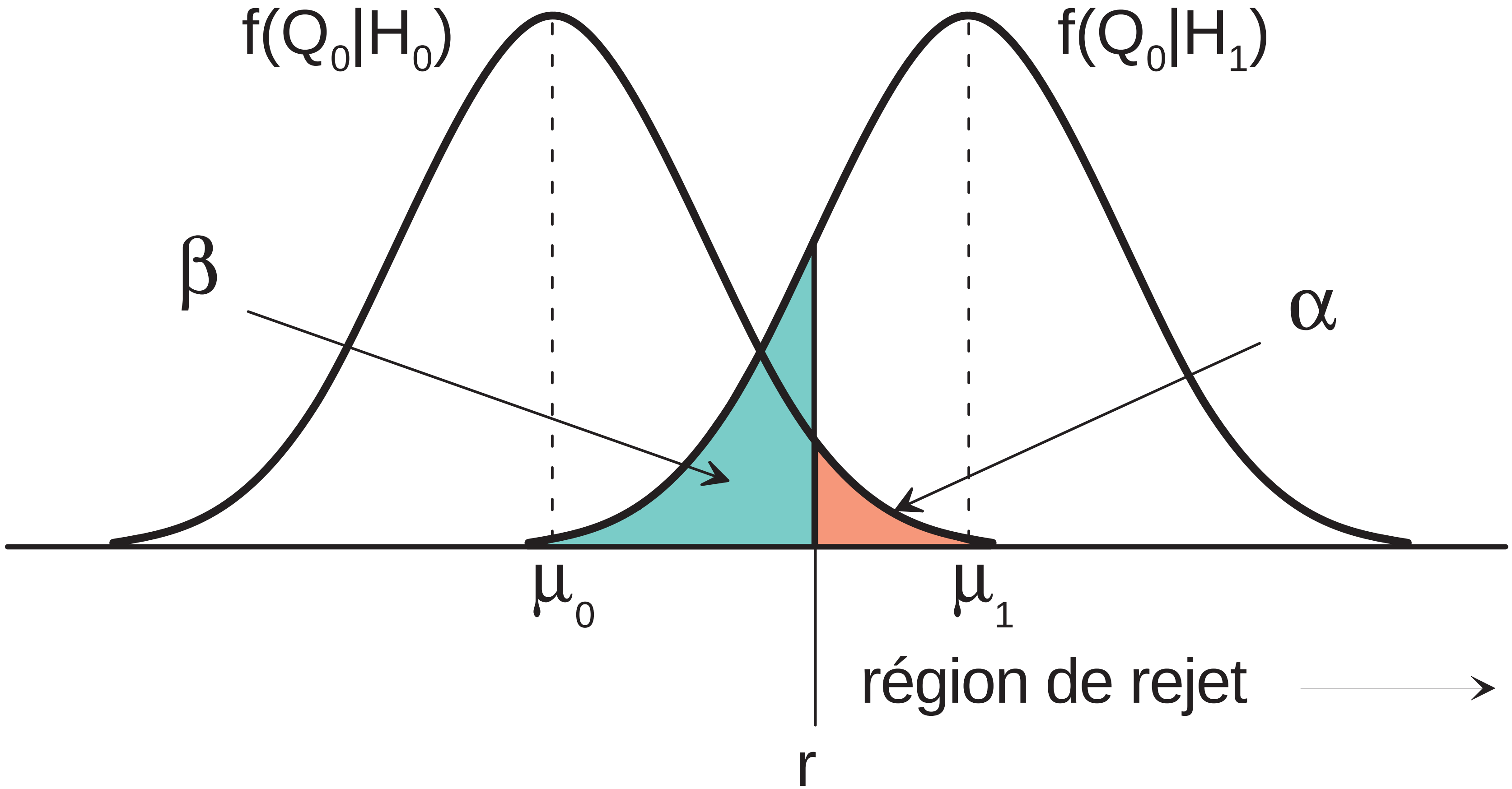
Dr Sacha Varone

Objectif

Savoir effectuer un test sur la moyenne

Rappels

Illustration d'un test



Étapes d'un test

1. Spécifier la valeur de la population d'intérêt.
2. Formuler l'hypothèse nulle H_0 et l'hypothèse alternative H_1
3. Choisir le niveau de signification α
4. Déterminer la région critique.
5. Calculer la statistique associée à l'échantillon.
6. Rejeter H_0 si la statistique appartient à la région critique. Ne pas rejeter H_0 dans le cas contraire.
7. Énoncer une conclusion.

h e g

Test de la moyenne

Test de μ, σ^2 connu

Population qui suit une loi normale de variance σ^2 connue
Par le TCL, $\bar{x} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$.

Test de μ, σ^2 connu

Population qui suit une loi normale de variance σ^2 connue
Par le TCL, $\bar{x} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$.

Par conséquent, pour tester une hypothèse sur μ, on utilisera la statistique

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

où :

$\bar{x}$	$=$	moyenne de l'échantillon
$z_{\alpha/2}$	$=$	valeur critique de la distribution normale standard pour un degré de confiance de $1-\alpha$
σ	$=$	écart type de la population
n	$=$	taille de l'échantillon

Remarque

- Lors d'un test unilatéral, la valeur critique est déterminée par
  - z_α s'il s'agit d'un test unilatéral à droite
  - $-z_\alpha$ s'il s'agit d'un test unilatéral à gauche
- Lors d'un test bilatéral, les valeurs critiques sont déterminées par $\pm z_{\alpha/2}$

Exemple

Une étude sur la mobilité des employés de l'entreprise "Jeux Mendors" affirme que la moyenne des temps de trajet des employés pour venir à leur travail excède 40 minutes. Vous devez tester cette affirmation avec une niveau de signification de 0.05. Pour cela, 100 employés vous ont indiqué leur temps de trajet actuel, dont la moyenne est 43.5 minutes. Basé sur une étude précédente, vous pouvez supposer que l'écart type de la population est de 8 minutes.

1. La valeur de la population d'intérêt est

Exemple

Une étude sur la mobilité des employés de l'entreprise "Jeux Mendors" affirme que la moyenne des temps de trajet des employés pour venir à leur travail excède 40 minutes. Vous devez tester cette affirmation avec une niveau de signification de 0.05. Pour cela, 100 employés vous ont indiqué leur temps de trajet actuel, dont la moyenne est 43.5 minutes. Basé sur une étude précédente, vous pouvez supposer que l'écart type de la population est de 8 minutes.

1. La valeur de la population d'intérêt est le temps de trajet.
2. Les hypothèses nulle et alternative sont :

Exemple

Une étude sur la mobilité des employés de l'entreprise "Jeux Mendors" affirme que la moyenne des temps de trajet des employés pour venir à leur travail excède 40 minutes. Vous devez tester cette affirmation avec une niveau de signification de 0.05. Pour cela, 100 employés vous ont indiqué leur temps de trajet actuel, dont la moyenne est 43.5 minutes. Basé sur une étude précédente, vous pouvez supposer que l'écart type de la population est de 8 minutes.

1. La valeur de la population d'intérêt est le temps de trajet.
2. Les hypothèses nulle et alternative sont :
   $H_0 \quad \mu \leq 40$ minutes
   $H_1 \quad \mu > 40$ minutes

Exemple (suite)

3. Le niveau de signification est

Exemple (suite)

3. Le niveau de signification est 0.05

4. La région critique est

Exemple (suite)

3. Le niveau de signification est 0.05
4. La région critique est $[z_{0.05} = 1.645; \infty[$
5. La statistique est

Exemple (suite)

3. Le niveau de signification est 0.05
4. La région critique est $[z_{0.05} = 1.645; \infty[$
5. La statistique est $z = \frac{43.5-40}{\frac{8}{\sqrt{100}}} = 4.38$
6. Comme 4.38 appartient à la région critique, l'hypothèse H_0 est rejetée.
7. La conclusion est

Exemple (suite)

3. Le niveau de signification est 0.05
4. La région critique est $[z_{0.05} = 1.645; \infty[$
5. La statistique est $z = \frac{43.5-40}{\frac{8}{\sqrt{100}}} = 4.38$
6. Comme 4.38 appartient à la région critique, l'hypothèse H_0 est rejetée.
7. La conclusion est que la moyenne des temps de trajet excède 40 minutes.

Rappel du Théorème Central Limite

Soit une suite $(X_1, X_2, \ldots, X_n)$ de n variables aléatoires identiquement et indépendamment distribuées (μ, σ^2).
Lorsque $n \to \infty$,

$$\overline{X} =\sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

où

$$\begin{array}{rcl} \mu & = & \text{moyenne dans la population} \\ \sigma^2 & = & \text{variance dans la population} \\ n & = & \text{taille de l'échantillon} \end{array}$$

h e g

Test de μ, σ^2 inconnu, n grand

Par le TCL, si $n \geq 30$,

$$z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

où

$\bar{x}$ = moyenne de l'échantillon

μ = moyenne supposée de la population étudiée

s = écart type de l'échantillon

n = taille de l'échantillon

Test de μ, σ^2 inconnu, n petit

Utiliser la distribution de Student, à $n-1$ degrés de liberté

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

où

$\bar{x}$ = moyenne de l'échantillon
μ = moyenne supposée de la population étudiée
s = écart type de l'échantillon
n = taille de l'échantillon

Supposition :POPULATION NORMALEMENT DISTRIBUÉE

Exemple

Tester la moyenne d'un échantillon du nombre d'employés par PME suisse contre la moyenne de la population, $\mu = 35$, en ne connaissant pas la valeur de la variance de la population.
Par des études précédentes, nous savons que le nombre d'employés est distribué selon une loi normale.
20 PME suisses ont été interrogées sur le nombre d'employés qu'elles occupent.
La moyenne calculée vaut $\overline{x} = 27$ et la variance calculée vaut $s^2 = 334.7$.
Choix d'un niveau de signification de 0.05.

1. La valeur de la population d'intérêt est

Exemple

Tester la moyenne d'un échantillon du nombre d'employés par PME suisse contre la moyenne de la population, $\mu = 35$, en ne connaissant pas la valeur de la variance de la population.
Par des études précédentes, nous savons que le nombre d'employés est distribué selon une loi normale.
20 PME suisses ont été interrogées sur le nombre d'employés qu'elles occupent.
La moyenne calculée vaut $\overline{x} = 27$ et la variance calculée vaut $s^2 = 334.7$.
Choix d'un niveau de signification de 0.05.

1. La valeur de la population d'intérêt est
   le nombre moyen d'employés des PME suisses.

2. Les hypothèses nulle et alternative sont :

Exemple

Tester la moyenne d'un échantillon du nombre d'employés par PME suisse contre la moyenne de la population, $\mu = 35$, en ne connaissant pas la valeur de la variance de la population.
Par des études précédentes, nous savons que le nombre d'employés est distribué selon une loi normale.
20 PME suisses ont été interrogées sur le nombre d'employés qu'elles occupent.
La moyenne calculée vaut $\overline{x} = 27$ et la variance calculée vaut $s^2 = 334.7$.
Choix d'un niveau de signification de 0.05.

1. La valeur de la population d'intérêt est
   le nombre moyen d'employés des PME suisses.

2. Les hypothèses nulle et alternative sont :
   $H_0 \quad \mu = 35$ personnes
   $H_1 \quad \mu \neq 35$ personnes

h e g

Exemple (suite)

3. Le niveau de signification est

Exemple (suite)

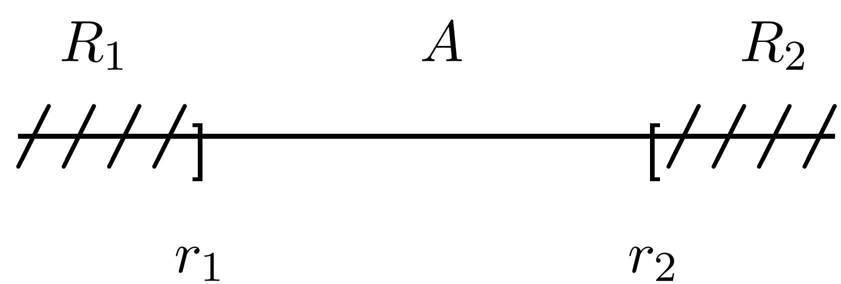
3. Le niveau de signification est 0.05

4. La région critique est composée de

Exemple (suite)

3. Le niveau de signification est 0.05

4. La région critique est composée de 2 zones distinctes car nous avons un test bilatéral à effectuer.
   $R = R_1 \cup R_2 = \{t_0 \mid t_0 \leq r_1\} \cup \{t_0 \mid t_0 \geq r_2\}$



$$P(T_0 \leq r_1) = \frac{\alpha}{2} = 0.025 \Rightarrow r_1 = t^{(19)}_{\alpha/2} = -2.093$$
$$P(T_0 \geq r_2) = 1 - \frac{\alpha}{2} = 0.975 \Rightarrow r_2 = t^{(19)}_{1-\alpha/2} = 2.093$$

Exemple (fin)

5. La statistique est

Exemple (fin)

5. La statistique est

$$t = \frac{27 - 35}{\sqrt{\frac{334.7}{20}}} = -1.955$$

6. Comme $t = -1.955$ n'appartient pas à la région critique, l'hypothèse H_0 n'est pas rejetée.

7. La conclusion est que

Exemple (fin)

5. La statistique est

$$t = \frac{27 - 35}{\sqrt{\frac{334.7}{20}}} = -1.955$$

6. Comme $t = -1.955$ n'appartient pas à la région critique, l'hypothèse H_0 n'est pas rejetée.

7. La conclusion est que la vraie moyenne de la population peut être égale à 35 employés par petite entreprise suisse.

Value at Risk

La *Value at Risk* (VaR) est définie comme la perte maximale sur un horizon donné T, avec un niveau de confiance $1-\alpha$.

Considérons une variable aléatoire $X \sim \mathcal{N}(\mu, \sigma^2)$

$$P\left(X \leq \mathsf{VaR}(\alpha, T)\right) = \alpha \quad \Leftrightarrow \quad P\left(X \geq \mathsf{VaR}(\alpha, T)\right) = 1-\alpha$$

VaR en Finance

La notion de risque financier peut être estimée à l'aide de l'écart type annuel des performances financières, appelé dans le jargon bancaire la *volatilité*. L'horizon est quant à lui souvent donné en jours (1, 10, ...), qui est à mettre en regard du nombre de jours ouvrables considéré dans la branche (250 ou 252 généralement). Considérons X une variable aléatoire indiquant les rendements, suivant une même loi normale $\mathcal{N}(\mu, \sigma^2)$ sur une seule période. Alors, sur la période T, les rendements sont également gausssiens de moyenne μT et de variance $\sigma^2 T$. La variable centrée réduite s'écrit :

$$z_\alpha = \frac{\text{VaR}(\alpha, T) - \mu T}{\sigma\sqrt{T}}$$

Et donc

$$\text{VaR}(\alpha, T) = \mu T + z_\alpha \sigma \sqrt{T}$$

Exemple

Source : Rapport annuel 2007 de l'UBS
version anglaise, "Risk, Treasury and Capital Management", p. 39
La position du secteur Banque d'investissement de l'UBS, sur un horizon de 1 jour, à un niveau de confiance de 99%, en utilisant des données sur 5 ans, est de

- 149 millions en 2007
- 160 millions en 2006