Les tests d'hypothèses

Introduction

On veut souvent tester une hypothèse en fonction des données disponibles.
Exemples

- Rendements d'échelles d'une fonction de production;
- Élasticité des prix;
- Rendement d'une année d'études supplémentaire

Une hypothèse peut être simple ou composée.

La plupart du temps, on a un ensemble paramétrique Θ, divisé en Θ_0 et Θ_1, et on veut tester

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1.$$

On construit donc une règle de decision basée sur un test statistique .
L' hypothèse H_0 est l'hypothèse nulle et l' hypothèse H_1 est l'alternative.
H_0, tient en général par défaut pour vrai, et l'on cherchera à l'infirmer. H_1 est contradictoire a H_0.

En se fondant sur les données on va soit rejeter H_0 et donc admettre H_1 soit échouer à la rejeter.

Exemple

On cherche à tester si une pièce est équilibrée ou non, soit θ la probabilité de pile, $\theta = P(H)$. On a deux hypothèses:

- H_0 la pièce est équilibrée, i.e. $\theta = \theta_0 = 1/2$.
- H^1 la pièce est pipée : $\theta = \theta_1$ différent de 1/2 ‾.

Solution

On fabrique un test pour savoir si la pièce suit plutôt H_0 ou H_1. Comme expérience on va lancer 100 fois la pièce compter les piles. soit X le nombre de pile on obtient :

$$X \sim Bin(100, \theta).$$

Donc, si H_0 est vrai, then $\theta = \theta_0 = 1/2$, on attend que le nombre de pile soit proche de 50, on acceptera donc H0 si il est suffisamment proche et le rejettera sinon.

Exemple: les critères

On fixera plus précisément une valeur limite:

si $|X - 50|$ est inférieur ou égal à la limite on ne rejette pas H_0.

si $|X - 50|$ est supérieur à la limite on rejette H_0 et accepte H_1.

Soit t.

- si $|X - 50| \leq t$, accepte H_0.
- si $|X - 50| > t$, accepte H_1.

Mais comment doit on déterminer t?

Exemple: Erreur de type 1

Pour choisir t de manière adéquate, nous avons quelques critères :

Dont le plus important est la probabilité d'erreur.

On fait une erreur lorsqu'on rejette H_0 alors qu'il est vrai. On appellera cette erreur erreur de type I.

Plus spécifiquement, $|X - 50| > t$ quand H_0 est vrai.

$$P(\text{type I}) = P(|X - 50| > t \mid H_0).$$

Il s'agit de la probabilité que le test indique $|X - 50| > t$ quand H_0 est vrai.

Exemple: la significativité

Pour choisir t needs on peut choisir une valeur cible pour :
$P($type I$)$.

Par exemple :

$$P\,(\text{type I}) \leq \alpha \;=0.05$$

On souhaite avoir 5 % de chance de rejeter H_0 quand H_0 est vrai.

Ici, α est appelé niveau de significativité. On prendra par

convention :

$$P\,(|X-50|>t\;|H_0)=\alpha=0.05$$

Comme on connait la distribution de X pour H_0, on pourra choisir t
De sorte a vérifier cette équation.

Exemple: les valeurs critiques

Pour simplifier on écrira c =t/5

- si $|Y| \leq c$, on accepte H_0.
- si $|Y| > c$, on accepte H_1.

ou Y = (X-50) / 5 et c est appelée la valeur critique. On doit donc décider la valeur de C en fonction de la distribution :

Par CLT $Y = \frac{X - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}} = \frac{X-50}{5}$ suit une loi normale N(0;1) donc on veut :

$$\begin{aligned}\alpha &= P(|Y| > c) \\ &= 1 - P(-c \leq Y \leq c) \\ &\approx 1 - (\Phi(c) - \Phi(-c)) = 2 - 2\Phi(c) \quad \left(\text{using } \Phi(x) = 1 - \Phi(-x)\right).\end{aligned}$$

Et :

$$2 - 2\Phi(c) = 0.05 \Rightarrow \Phi(c) = \frac{2 - 0.05}{2} = 0.975$$

$$c = \Phi^{-1}(0.975) = 1.96$$

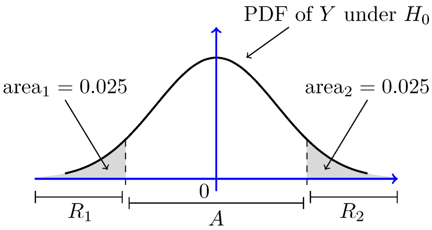
Exemple: Région acceptée :

En conclusion du test

- si $|Y| \leq 1.96$, on accepte H_0.
- si $|Y| > 1.96$, on accepte H_1.

Donc $A = [-1.96, 1.96]$ est appelée la zone acceptée .

Et $R = (-\infty, -1.96) \cup (1.96, \infty)$ est appelée la zone de rejet.



A = Acceptance Region

$R = R_1 \cup R_2$ = Rejection Region

α = P(type I error) = $\text{area}_1 + \text{area}_2 = 0.05$

Exemple: En chiffres :

On rappelle que $Y = \frac{X-50}{5}$, soit :

- si $|X - 50| \leq 9.8$, admet H_0.
- si $|X - 50| > 9.8$, admet H_1.

donc,

- si X est inclus dans $\{41, 42, \cdots, 59\}$, on accepte H_0
- si X est dans $\{0, 1, \cdots, 40\} \cup \{60, 61, \cdots, 100\}$, on rejette H_0 (accepte H_1).

En conclusion si il y a plus de 9 piles de plus ou moins que 50, on rejette H_0.

Généralisation

Considérons un échantillon aléatoire $\{X\}_n \sim f_X(\cdot;\theta)$.

Et un test statistique $T_n = T_n(X_1, \ldots, X_n)$

On postule une hypothèse, une règle de décision :

rejeter H_0 si $T_n \in R$ Ne pas rejeter H_0 si $T_n \in A$,

R est la zone de rejet et A la zone d'acceptation.

Ici, rejeter H_0 implique d'accepter H_1.

Types d’erreurs

Cependant comme T_n est aléatoire, toute decision basée dessus est prone à deux types d’erreur.

	Vrai	
Decision	H_0	H_1
rejette H_0	OK	Type II
rejette H_0	Type I	OK

- Type-I (faux negatif) quand on rejette H_0 à tort.
- Type-II (faux positif) quand on échoue à rejeter H_0 à tort.

Il faut alors chercher à minimiser ces erreurs

Niveau de significativité

Si la probabilité d'erreur de type I satisfait :

$$P(\text{type I erreur}) \leq \alpha, \quad \text{pour tout } \theta \in \Theta_0, \tag{1}$$

on dira que le test a un niveau de significativité α.

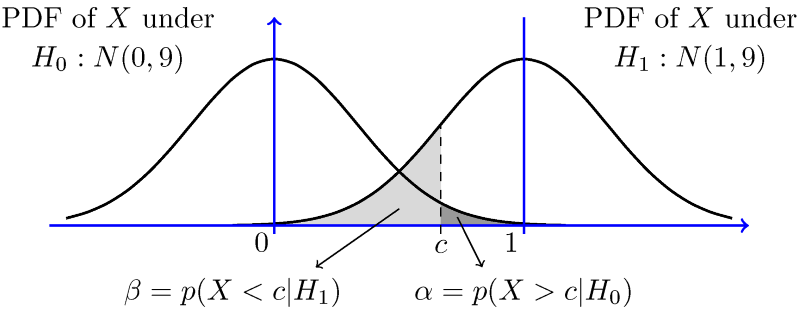
Souvent on choisira H_0 comme une hypothèse simple de sorte que Θ_0 ait un seul élément.

La deuxième erreur possible est qu'on accepte H_0 quand H_0 est fausse. Comme l'alternative, H_1, est le plus souvent une hypothèse composite (plusieurs valeurs possibles de θ), la probabilité d' erreur de type II est une fonction de θ. Notée β:

$$\beta(\theta) = P(\text{accepte } H_0 / \theta), \quad \text{pour } \theta \in \Theta_1. \tag{2}$$

Trade-off entre α et β

Comme α et β sont des erreurs en probabilités, on les voudrait le plus petit possible,
Cependant il s'agit en fait d'un échange entre α et β.
En fait , si on veut minimiser la probabilité d' erreur type I (α), alors la probabilité d' erreur type II (β) augmente, et vise versa.
En image :



Les principaux points :

Pour tester une hypothèse nulle et son alternative on doit choisir un test et une valeur critique.

Un test statistique, noté T, est une fonction de l'échantillon. Quand on calcule sa réalisation , on obtient une valeur notée t.

Avec un test on détermine une zone de rejet pour laquelle on pourra rejeter H_0 en faveur de H_1. cette zone s'obtient en comparant le résultat du test , t, à une valeur critique, c.

Pour choisir la valeur critique, on doit choisir un niveau de test α. Puis la valeur de c associée.

Généralités

Notre règle de rejet dépendra de la forme de l'hypothèse à tester .

les trois alternatives quand on veut tester $H_0 : \theta = 0$

$H_1 : \theta > 0$

$H_1 : \theta < 0$

$H_1 : \theta$ *différent de* 0

Une statistique est dite significative si t se trouve dans une région critique . On rejette alors H0.

À l'inverse lorsque t se trouve en dehors des régions critique le test n'es pas significatif et on ne peut pas rejeter H0.

Exemple

Supposons $\{X\}_n \sim N(\mu, \sigma^2)$; on cherche à déterminer la moyenne

Si l'on connait la variance on sait qu'est standard :

$$T_n = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

si σ^2 est inconnue, notre test est basé sur

$$\frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

Qui admet pour distribution t_{n-1}.

Modèle Normal: σ^2 connu

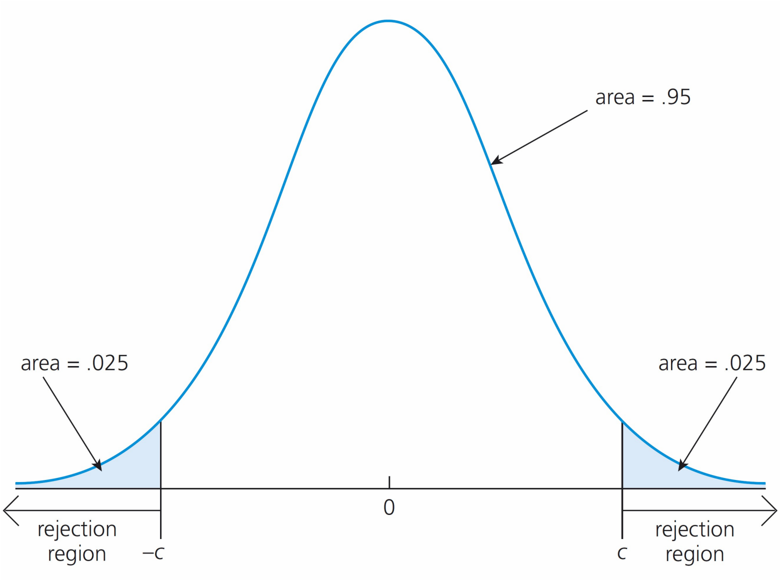
Pour un niveau de significativité α, la règle de décision sera ($H_0 : \mu = \mu_0$; $H_1 : \mu \neq \mu_0$) :

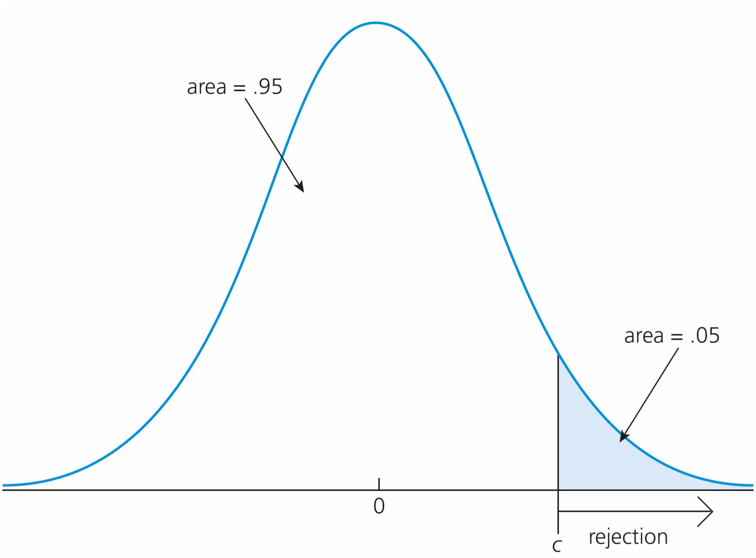
rejeter H_0 si : $$\left| \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \right| > z(\alpha/2)$$

Test unilatéral à droite : ($H_0 : \mu \leq \mu_0$; $H_1 : \mu > \mu_0$):

rejette H_0 $$\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z(\alpha)$$

Et inversement à gauche.


area = .95
area = .025
area = .025
0
rejection region
−c
c
rejection region


area = .95
area = .05
0
c
rejection

Choisir α ou indiquer p-value?

Notre règle décision repose sur α, que nous choisissons arbitrairement .
Les plus courants sont $\alpha \in \{.01, .05, .10\}$.

Définition

La p-value est :

$$p = \sup_{\theta \in \Theta_0} P\,(T \geq t),$$

Où t est la réalisation de T calculée depuis l'échantillon.

La p-value est définie comme le niveau le plus élevé pour lequel un test basé sur T échoue à rejeter l'hypothèse nulle. De manière équivalente la p-value donne le plus petit alpha qui permet de rejeter H_0.

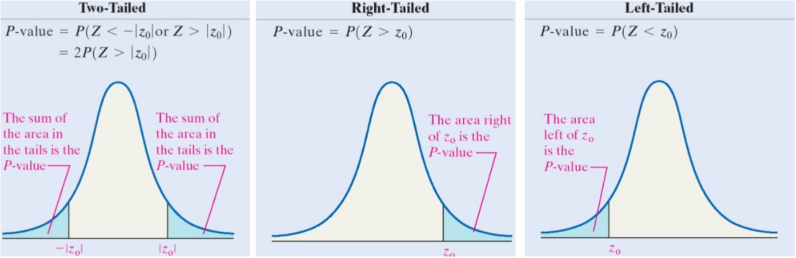
Donc le test rejette H0 pour alpha supérieur à la p-value et échoue à rejeter pour les niveaux plus faibles que la p-value.

Calculer la p-value

Assumons que l' hypothèse nulle est vraie.

On calcule :

$$z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$



Modèle Normal: σ^2 Inconnue

On remplace σ par son estimateur S, et $z(\alpha)$ par $t(n-1;$
$\alpha)$.

Test bilatéral :

rejeter H_0 si :
$$\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right| > t(n-1;\alpha/2)$$

Unilatéral à droite :

rejeter H_0 si :
$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}} > t(n-1;\alpha)$$

Et inversement pour le test unilatéral à gauche.