Probabilités et statistiqyes - Cours 3 : Tests d'hypothèses paramétriques

2021

Plan

Introduction

Les tests statistiques ont pour objet de décider sur la base d'un échantillon si une caractéristique de la loi mère répond ou non à une certaine spécification que l'on appelle *hypothèse*.

Exemple

La moyenne de la loi est supérieure à 10.

Ces spécifications peuvent avoir diverses provenances :

- normes imposées
- affirmations faites par un tiers (ex. le fabricant d'un produit)
- valeurs cruciales de paramètres de modèles

Introduction

Cadre paramétrique : les hypothèses portent sur un paramètre inconnu θ ou sur une fonction de ce paramètre $h(\theta)$, correspondant à une caractéristique d'intérêt de la loi.

Cas simple : L'hypothèse spécifie une valeur ou un intervalle de valeurs pour θ (ou $h(\theta)$).

Un test décide si un ensemble de valeurs spécifié est *plausible* ou non.

Introduction

Les tests statistiques permettent d'aborder une grande variété d'hypothèses bien au-delà du test d'une hypothèse portant sur un paramètre.

Exemple

Comparaison de plusieurs lois, existence de liens entre plusieurs v.a., adéquation d'un modèle, etc.

Ce sera l'objet d'un prochain cours.

Cadre paramétrique

La loi observée appartient à une famille de lois, décrites par la famille de densités de probabilités (resp. de fonctions de probabilités dans le cas discrêt)

$$\{f(x, \theta)/\theta \in \Theta\}.$$

La forme fonctionnelle de f est connue, θ est inconnue. La fonction de répartition est notée $F(x, \theta)$.

L'ensemble Θ est l'*espace paramétrique*.

Hypothèse nulle

Dans l'approche paramétrique, un test statistique consiste à décider d'accepter ou de rejeter une hypothèse spécifiant que θ appartient à un ensemble de valeurs Θ_0.

Cette hypothèse de référence est appelée **hypothèse nulle** et est notée H_0.

Hypothèse alternative

L'**hypothèse alternative**, notée H_1, est l'hypothèse pour laquelle $\theta \in \Theta \backslash \Theta_0$.

On teste donc : $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$.

Hypothèse simple ou multiple

Suivant la nature de Θ_0 et Θ_1, on distingue trois cas :

- Hypothèse nulle simple et alternative simple :
  $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$
- Hypothèse nulle simple et alternative multiple :
  $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$
- Hypothèse nulle multiple et alternative multiple :
  $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$

Plan

Statistique de test

Dans le cas d'hypothèses simples, l'espace paramétrique est $\Theta = \{\theta_0, \theta_1\}$, et le test veut décider : $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$.

Un test pour H_0 est une règle de décision fondée sur la valeur réalisée t sur un échantillon, d'une statistique T, appelée **statistique du test**, à valeurs dans $\mathbb{R}$.

Région d'acceptation

La règle de décision pour le test est la suivante :

- Si $t \in A$ (une partie de $\mathbb{R}$), on accepte H_0
- Si $t \in \overline{A}$ (le complémentaire de A), on rejette H_0

La partie A est appelée *région d'acceptation* du test, et la partie $\overline{A}$ la **région de rejet** du test. Lorsque A est un intervalle, il est appelé **intervalle de confiance** lié au test.

Risques

Une telle règle de décision recèle deux types d'erreurs possibles du fait que la vraie valeur du paramètre est inconnue.

Definition (Risque de première espèce)

On appelle *risque de première espèce* pour un test, la valeur α telle que : $P_{H_0}(T \in \overline{A}) = \alpha$, c'est-à-dire la probabilité de rejeter H_0 alors qu'elle est vraie.

Remarque : on note aussi $P_{H_0}(T \in \overline{A}) = P(T \in \overline{A} | H_0)$.

Risques

Definition (Risque de deuxième espèce)

On appelle *risque de deuxième espèce* pour un test, la valeur β telle que : $P_{H_1}(T \in A) = \beta$, c'est-à-dire la probabilité d'accepter H_0 alors qu'elle est fausse.

Remarque : on note aussi $P_{H_1}(T \in A) = P(T \in A | H_1)$.

Les deux risques sont interdépendants puisque l'un repose sur A et l'autre sur $\overline{A}$. Par le choix de A on peut donc vouloir contrôler l'un ou l'autre, mais pas les deux.

Risques

Dans un test statistique, on privilégie en fait le risque α (par ex. $\alpha = 0.05$) que l'on fixe *a priori*.

La valeur α est aussi appelée *niveau* ou *niveau de signification* du test.

Une fois ce niveau choisi, on détermine l'intervalle de confiance associé à ce niveau.

$\Rightarrow$ la loi de la statistique de test sour H_0 doit être parfaitement connue.

Construction d'un test

1. Déterminer les hypothèses H_0 et H_1
2. Rechercher une statistique pertinente dont on connaît la loi sous H_0
3. Fixer un niveau α
4. Déterminer l'intervalle de confiance associé à ce niveau, en utilisant la loi de la statistique de test
5. Prendre une décision en considérant la réalisation de la statistique sur l'échantillon

Remarques

Face à une spécification sur une caractéristique d'un produit, le fait de rejeter cette spécification est une preuve quasi-irréfutable qu'elle n'est pas correcte, alors que le fait de l'accepter ne signifie pas qu'elle soit correcte mais simplement que, sur la base des observations effectuées, rien ne permet de conclure qu'elle soit fausse.

Les rôles de H_0 et H_1 ne sont pas interchangeables. En particulier, il n'est pas nécessaire de connaître la loi sous H_1.

Puissance d'un test

Comment juger de la pertinence de la statistique choisie ?

Il est naturel de la prendre de telle sorte que la probabilité de rejeter H_0 soit plus élevée sous H_1 que sous H_0, et si possible, *nettement* plus élevée.

Definition (Puissance d'un test)

On appelle *puissance d'un test* la probabilité de rejeter H_0 alors qu'elle est effectivement fausse, soit :

$$P(T \in \overline{A} | H_1).$$

Puissance d'un test

Remarque : La puissance, qui est la capacité à détecter qu'une hypothèse nulle est fausse, est égale à $1 - \beta$.

Definition

On dit qu'un test est *sans biais* si sa puissance est supérieure ou égale à son niveau α :

$$P(T \in \overline{A} | H_1) \geq P(T \in \overline{A} | H_0).$$

Conclusion : une condition naturelle pour qu'une statistique soit éligible pour tester une hypothèse est qu'elle induise un test sans biais.

Puissance d'un test

Le choix entre plusieurs tests s'effectue sur la puissance.

Definition

On dit que le test $\mathcal{T}_1$ est plus puissant que le test $\mathcal{T}_2$ au niveau α, si :

- $\mathcal{T}_1$ est de niveau α
- $\mathcal{T}_2$ est de niveau $\leq \alpha$
- la puissance de $\mathcal{T}_1$ est supérieure à la puissance de $\mathcal{T}_2$

Objectif : Chercher le test le plus puissant parmi tous.

Remarques

Dans le cas où H_0 et H_1 sont des hypothèses simples il existe un test le plus puissant, mais cela n'est pas toujours le cas lorsque l'on a des hypothèses multiples.

Lorsqu'une statistique de test donne le test le plus puissant à un niveau donné, généralement elle reste optimale à tout autre niveau.

Remarques

Dans le cas d'une loi discrête, la statistique sera elle-même discrête et le niveau α choisi ne pourra être exactement atteint. On appelle cela un *test conservateur*.

Exemple

Si on souhaite un risque de première espèce $\alpha = 0.05$, on cherche une region de rejet $\overline{A}$ de probabilité sous H_0 la plus proche possible de α, mais inférieure ou égale à α.

Remarques

On peut ajouter une autre condition que celle "être sans biais" : lorsque la taille de l'échantillon devient grande, la suite de tests correspondant $\{\mathcal{T}_n\}$ doit avoir une puissance $1 - \beta_n$ qui croît en tendant vers 1.

Cela garantit que l'on gagne à observer de très grands échantillons, étant pratiquement sûr, à la limite, de détecter une hypothèse nulle qui serait fausse.

On dit alors que la procédure de test est *convergente*.

Exemple 1

Deux machines A et B produisent le même type de produit, mais la machine A fournit un produit plus cher et de qualité supérieure.

La qualité du produit se mesure de façon aléatoire suivant une loi $\mathcal{N}(5, 1)$ pour la machine A et $\mathcal{N}(4, 1)$ pour la machine B.

Un client achète le produit le plus cher par lot de 10 et désire développer un test pour contrôler qu'un lot donné provient bien de la machine A. La loi mère du lot est notée $\mathcal{N}(\mu, 1)$.

Exemple 1

Détermination des hypothèses

Comme accuser à tort le vendeur de fraude sur la marchandise peut avoir de graves conséquences, le client doit limiter le risque correspondant, et tester :

$$H_0 : \mu = 5 \quad \text{vs.} \quad H_1 : \mu = 4$$

à un niveau $\alpha = 0.05$ par exemple.

Exemple 1

Choix de la statistique de test

Il semble naturel d'employer comme statistique de test :

$$\overline{X} = \frac{1}{10}(X_1 + X_2 + \ldots + X_{10}).$$

Sous l'hypothèse H_0, sa loi est $\mathcal{N}(5, \frac{1}{10})$.

Exemple 1

Détermination de l'intervalle de confiance

On cherche un intervalle de probabilité 0.95 pour cette statistique : $P(-t_{1-\frac{\alpha}{2}} < \frac{\overline{X}-5}{\frac{1}{\sqrt{10}}} < t_{1-\frac{\alpha}{2}}) = 1-\alpha$, où $t_{1-\frac{\alpha}{2}}$ est un quantile de la loi normale centrée réduite.

On en déduit l'intervalle de confiance : $\left[5 - \frac{1.96}{\sqrt{10}}, 5 + \frac{1.96}{\sqrt{10}}\right]$, soit $[4.38, 5.62]$.

Exemple 1

Règle de décision

On applique la règle de décision suivante à la réalisation de la statistique :

- Si la réalisation $\overline{x}$ de $\overline{X}$ sur le lot considéré est dans $[4.38, 5.62]$, on accepte H_0
- Sinon on rejette H_0

Exemple 1

Il est possible de calculer la puissance de ce test puisque la loi de $\overline{X}$ sous H_1 est connue, c'est $\mathcal{N}(4, \frac{1}{\sqrt{10}})$.

$$\begin{aligned}
\beta &= P(4.38 < \overline{X} < 5.62 | H_1) \\
&= P\left(\frac{4.38-4}{\frac{1}{\sqrt{10}}} < Z < \frac{5.62-4}{\frac{1}{\sqrt{10}}}\right) \text{ où } Z \rightsquigarrow \mathcal{N}(0.1) \\
&= P(1.20 < Z < 5.12) \simeq 0.115
\end{aligned}$$

D'où une puissance d'approximativement 0.885.

Exemple 1

Remarque : On a choisi précédemment un test *bilateral*, mais il y a peu de sens à rejeter H_0 lorsque la moyenne de la réalisation est supérieure à 5.62.

On passe alors à un test *unilatéral*, c'est-à-dire qu'au lieu de prendre une région d'acceptation symétrique, on considère une région du type $x \geq t_\alpha$ où $P(X \in [t_\alpha, +\infty] | H_0) = 0.95$.

Exemple 1

Dans notre cas : $t_\alpha = 5 - \dfrac{1.645}{\sqrt{10}}$, où 1.645 est le quantile d'ordre 0.05 de la loi $\mathcal{N}(0,1)$.

On a alors : $A = [4.48, +\infty[$.

Le test est alors plus puissant :
$\beta = P(4.48 < \overline{X}|H_1) = P\left(\frac{4.48-4}{\frac{1}{\sqrt{10}}} < Z\right) = P(1.52 < Z) \simeq 0.064$, d'où une puissance 0.936.

Exemple 2

Le nombre de particules émises par une source radioactive par unité de temps suit une loi de Poisson.

On observe l'émission d'un corps durant 20 unités de temps et on doit décider s'il s'agit d'une source de type A ou d'une source de type B.

La source A émet en moyenne 0.6 particules par unité de temps et la source B en émet en moyenne 0.8.

Exemple 2

Détermination des hypothèses

La loi de la source observée est $\mathcal{P}(\lambda)$.

On teste donc :

$$H_0 : \lambda = 0.6 \quad \text{vs.} \quad H_1 : \lambda = 0.8$$

Exemple 2

Choix de la statistique de test

On choisit comme statistique de test la variable aléatoire X représentant le nombre total de particules émises au cours des 20 unités de temps.

Cette statistique suit une loi $\mathcal{P}(0.6 \times 20 = 12)$ sous H_0 et une loi $\mathcal{P}(0.8 \times 20 = 16)$ sous H_1.

Exemple 2

Détermination de l'intervalle de confiance

Intuitivement, on choisit une région de rejet de la forme $X \geq k$, puisqu'un nombre plutôt élevé de particules va à l'encontre de l'hypothèse nulle.

On choisit a priori $\alpha = 0.05$, et on lit dans la table que pour une v.a. $X \rightsquigarrow \mathcal{P}(12)$ on a :

$$P(X \geq 18) = 0.0630 \text{ et } P(X \geq 19) = 0.0374.$$

On opte pour un test conservateur en prenant comme zone de rejet $[19, +\infty[$.

Exemple 2

On peut calculer la puissance du test :

$$P(X \geq 19|H_1) = 0.258$$

Le test vérifie alors la condition sans biais.

Plan

Risques, puissance et optimalité

Les définitions précédentes changent légèrement.

En effet, $P(T \in \overline{A} | H_0)$ n'a pas de sens si H_0 est multiple.

On aura plutôt : $P_\theta(T \in \overline{A})$, pour $\theta \in \Theta_0$.

Risques, puissance et optimalité

Definition

Soit $H_0 : \theta \in \Theta_0$ une hypohèse nulle multiple et $\alpha(\theta)$ le risque de première espèce pour la valeur $\theta \in \Theta_0$. On appelle *niveau du test* (ou seuil du test) la valeur $\alpha = \sup_{\theta \in \Theta_0} \alpha(\theta)$.

Risques, puissance et optimalité

Definition

La fonction puissance du test est : $h(\theta) = 1 - \beta(\theta) = P_\theta(T \in \overline{A})$ définie pour tout $\theta \in \Theta_1$.

Definition

On dit qu'un test est *sans biais* si sa fonction puissance reste supérieurs ou égale à α, soit :

$$\forall \theta \in \Theta_1 : P_\theta(T \in \overline{A}) \geq \alpha.$$

Risques, puissance et optimalité

Autrement dit, la probabilité de rejeter H_0 si elle est fausse quelle que soit la valeur de θ dans Θ_1, est toujours plus élevée que la probabilité de la rejeter si elle vraie, quelle que soit alors la valeur de θ dans Θ_0.

Choix de H_0 dans un test unilatéral

Face à une situation pratique, le choix du sens de H_0 ($\theta \leq \theta_0$ ou $\theta \geq \theta_0$) n'est pas toujours évident.

Il se fait en considérant les deux erreurs possibles. On doit faire en sorte que celle qui est jugée la plus grave soit l'erreur de première espèce.

Choix de H_0 dans un test unilatéral

Supposons qu'un médicament soit considéré efficace lorsqu'un paramètre θ dépasse un seuil θ_0.

On peut déclarer le médicament efficace alors qu'il ne l'est pas ou le déclarer inefficace alors qu'il est efficace.

La première erreur est plus critique car elle a pour conséquence la mise sur le marché d'un médicament inefficace. La deuxième a pour conséquence de ne pas diffuser le médicament par mesure conservatoire.

H_1 doit donc correspondre au fait que le médicament est efficace, soit $\theta > \theta_0$, d'où $H_0 : \theta \leq \theta_0$.

Tests et p-valeur

La décision d'accepter ou de refuser une hypothèse est sujette au choix du risque de première espèce α.

Afin d'éviter ce choix arbitraire, on peut recourir à la notion de p-valeur pour simplement rendre compte du résultat d'un test.

Definition

La *p-valeur* est la probabilité que, sous H_0, la statistique du test prenne une valeur au moins aussi extrême que celle qui a été observée.

Tests et p-valeur

Si la région de rejet est unilatérale du type $t > c$ alors pour une valeur observée t_0 après expérience, la p-valeur est $P(T > t_0|H_0)$ si H_0 est simple, ou bien $\max_{\Theta_0} P_\theta(T > t_0)$ si elle est multiple.

Si la région de rejet est bilatérale, par exemple $\{t/t < c_1 \text{ ou } t > c_2\}$, alors la p-valeur est définie par $2P(T < t_0|H_0)$ si t_0 est plus petit que la médiane de la loi de T sous H_0 ou $2P(T > t_0|H_0)$ s'il est plus grand, afin de tenir compte du rejet sur les deux extrémités.

Tests et p-valeur

Exemple : $H_0 \; : \mu \leq 5$ vs. $H_1 \; : \mu > 5$ pour une loi mère $\mathcal{N}(\mu, 1)$. On dispose d'un échantillon de taille 10.

On est amené à rejeter H_0 si $\overline{x} > 5 + \frac{1.645}{\sqrt{10}}$ au niveau 0.05.

De manière plus générale, on rejette H_0 si $\overline{x} > 5 + z_{1-\alpha} \times \frac{1}{\sqrt{10}}$ au niveau α, où $z_{1-\alpha}$ est le quantile d'ordre $1-\alpha$ de la loi $\mathcal{N}(0, 1)$.

Tests et p-valeur

Supposons que l'échantillon observé ait une moyenne égale à 6. Comme le risque de première espèce maximal est atteint pour $\mu = 5$, la p-valeur est $P(\overline{X} > 6)$ pour $\overline{X} \rightsquigarrow \mathcal{N}\left(5, \frac{1}{10}\right)$, soit $P(Z > \sqrt{10}(6-5)) = P(Z > 3.16) = 0.008$ pour $Z \rightsquigarrow \mathcal{N}(0,1)$.

Cela indique que la valeur observée est au-delà de la valeur critique au niveau 0.05 et même au niveau 0.01.

Tests et p-valeur

Si le test avait été bilatéral avec $H_0 \ : \mu = 5$ vs. $H_1 \ : \mu \neq 5$, la p-valeur correspondant à la même observation 6 aurait été 0.016 impliquant un rejet au niveau 0.05 mais pas au niveau 0.01.

Remarque : plus la p-valeur est faible, plus l'hypothèse nulle est suspecte.

Plan

Cas où σ^2 est connu

C'est le cas le plus simple. On teste :

$$H_0 \ : \mu = \mu_0 \text{ vs. } H_1 \ : \mu \neq \mu_0$$

On prend la statistique $\overline{X}$ moyenne empirique. On a, sous H_0 :

$$\frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow \mathcal{N}(0, 1).$$

Cas où σ^2 est connu

Intervalle de confiance :

$$
\begin{aligned}
1-\alpha \quad & P_{\mu_0}\left(-z_{1-\frac{\alpha}{2}} < \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} < z_{1-\frac{\alpha}{2}}\right) \\
& P_{\mu_0}\left(\mu_0 - z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} < \overline{X} < \mu_0 + z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right)
\end{aligned}
$$

d'où $A = \{\overline{x}/\mu_0 - z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} < \overline{x} < \mu_0 + z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\}$ est l'intervalle de confiance pour ce test.

Ce test est le plus puissant, et sans biais, pour ce cas.

Cas où σ^2 est connu

Pour des hypothèses unilatérales : $H_0 \ : \mu \leq \mu_0$ vs. $H_1 \ : \mu > \mu_0$, il est naturel de rejeter H_0 lorsque $\overline{x}$ est trop grand car cela reflète une moyenne μ élevée.

Pour déterminer la valeur critique on se place en $\mu = \mu_0$ qui est la valeur la plus défavorable pour H_0 (le risque de première espèce est maximal pour $\mu = \mu_0$).

Cas où σ^2 est connu

Comme $\alpha = P_{\mu_0}\left(z_{1-\alpha} < \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$, on a pour région de rejet

$$\overline{A} = \left\{\overline{x}/\overline{x} > \mu_0 + z_{1-\alpha} \times \frac{\sigma}{\sqrt{n}}\right\}.$$

De même, pour $H_0 \; : \mu \geq \mu_0$ vs. $H_1 \; : \mu < \mu_0$, la région de rejet sera :

$$\overline{A} = \left\{\overline{x}/\overline{x} < \mu_0 - z_{1-\alpha} \times \frac{\sigma}{\sqrt{n}}\right\}.$$

Cas où σ^2 est inconnu

C'est un cas plus réaliste... On s'intéressera alors aux statistiques $\overline{X}$ et S^2.

On sait que $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \rightsquigarrow t(n-1)$ loi de Student.

Ainsi pour une hypothèse nulle $H_0 \ : \mu = \mu_0$ vs. $H_1 \ : \mu \neq \mu_0$, on a, sous H_0 :

$$\frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} \rightsquigarrow t(n-1)$$

Cas où σ^2 est inconnu

Intervalle de confiance :

$$1 - \alpha \quad = P_{\mu_0}\left(-t_{1-\frac{\alpha}{2}}^{(n-1)} < \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} < t_{1-\frac{\alpha}{2}}^{(n-1)}\right)$$

On définit l'intervalle de confiance pour un test de risque de première espèce α :

$$\frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}} \in \left[-t_{1-\frac{\alpha}{2}}^{(n-1)}, t_{1-\frac{\alpha}{2}}^{(n-1)}\right].$$

Cas où σ^2 est inconnu

Pour des hypothèses bilatérales on aura comme régions d'acceptation, pour :

$$H_0 \ : \mu \leq \mu_0 \quad \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}} \leq t_{1-\alpha}^{(n-1)}$$

$$H_0 \ : \mu \geq \mu_0 \quad \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}} \geq -t_{1-\alpha}^{(n-1)}$$

Cas où σ^2 est inconnu

Puissance de ce test : on calcule pour "μ dans H_1"

$$h(\mu) = P_\mu\left(-t_{1-\frac{\alpha}{2}}^{(n-1)} < \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} < t_{1-\frac{\alpha}{2}}^{(n-1)}\right)$$

Mais $\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}$ ne suit plus une loi de Student car μ_0 n'est plus la moyenne de $\overline{X}$.

Cas où σ^2 est inconnu

On écrit :

$$\frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{(\overline{X} - \mu) + (\mu - \mu_0)}{\frac{S}{\sqrt{n}}}$$

qui suit une loi de Student non centrale, de paramètre de non-centralité $\frac{(\mu - \mu_0)}{\sigma}$.

Les tables des lois de Student non centrales sont peu utilisées.

Tests sur la variance d'une loi $\mathcal{N}(\mu, \sigma^2)$

On suppose μ inconnu et on s'intéresse à une hypothèse sur σ^2, par exemple :

$$H_0 \ : \sigma^2 = \sigma_0^2 \text{ vs. } H_1 \ : \sigma^2 \neq \sigma_0^2.$$

On constate que :

$$\frac{(n-1)S^2}{\sigma^2} \rightsquigarrow \chi^2(n-1)$$

Tests sur la variance d'une loi $\mathcal{N}(\mu, \sigma^2)$

On a donc, sous H_0 :

$$P\left(\chi^{2\ (n-1)}_{\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma_0^2} < \chi^{2\ (n-1)}_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

On dispose donc d'une statistique de test $\frac{(n-1)S^2}{\sigma^2}$ et d'une région d'acceptation $\left[\chi^{2\ (n-1)}_{\frac{\alpha}{2}}, \chi^{2\ (n-1)}_{1-\frac{\alpha}{2}}\right]$.

Tests sur la variance d'une loi $\mathcal{N}(\mu, \sigma^2)$

Hypothèse unilatérale : $H_0 \ : \sigma^2 \leq \sigma_0^2$: on accepte si $\frac{(n-1)s^2}{\sigma_0^2} \leq {\chi^2_{1-\alpha}}^{(n-1)}$

$H_0 \ : \sigma^2 \geq \sigma_0^2$: on accepte si $\frac{(n-1)s^2}{\sigma_0^2} \geq {\chi^2_{\alpha}}^{(n-1)}$

Remarque : Si μ est connu, on utilise la statistique de test

$$\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \rightsquigarrow \chi^2(n).$$

Test de comparaison des moyennes de deux lois de Gauss

On est en présence de deux échantillons *indépendants* de taille n_1 (resp. n_2), de moyenne empirique $\overline{X_1}$ (resp. $\overline{X_2}$), de variance de l'échantillon S_1^2 (resp. S_2^2) issu d'une loi mère $\mathcal{N}(\mu_1, \sigma_1^2)$ (resp. $\mathcal{N}(\mu_2, \sigma_2^2)$).

On souhaite comparer μ_1 et μ_2 sur la base des échantillons.

Test de comparaison des moyennes de deux lois de Gauss

Question : Cas bilatéral ($H_0 \ : \mu_1 = \mu_2$ vs. $H_1 \ : \mu_1 \neq \mu_2$) ou unilatéral ($H_0 \ : \mu_1 \leq \mu_2$ vs. $H_1 \ : \mu_1 > \mu_2$) ?

En expérimentation clinique, on veut démontrer l'efficacité d'un traitement en comparant un échantillon traité et un échantillon témoin. On cherche à savoir si le traitement est efficace *en moyenne*.

Dans ce cas on prendra alors $H_0 \ : \mu_1 \leq \mu_2$ vs. $H_1 \ : \mu_1 > \mu_2$.

Test de comparaison des moyennes de deux lois de Gauss

Si les lois ont même variance théorique $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

On considère comme statistique de test :

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow t(n_1 + n_2 - 2)$$

où $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ est un estimateur sans biais de la variance commune σ^2.

Test de comparaison des moyennes de deux lois de Gauss

Si les lois n'ont pas même variance théorique

On utilise une approximation pour de grands échantillons (n_1 et $n_2 \geq 100$) :

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \rightsquigarrow_{approx} \mathcal{N}(0,1)$$

Test de comparaison des moyennes de deux lois de Gauss

Cas des échantillons appariés

On se ramène au cas d'un seul échantillon en étudiant la série des différences entre paires. Soit $\overline{D}$ et S_D respectivement la moyenne et l'écart-type de l'échantillon des paires :

$$\frac{\overline{D}}{\frac{S_D}{\sqrt{n}}} \rightsquigarrow t(n-1).$$

Test de comparaison des variances de deux lois de Gauss

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \rightsquigarrow F(n_1 - 1, n_2 - 1)$$

On teste : $H_0 \; : \frac{\sigma_1^2}{\sigma_2^2} = 1$ vs. $H_1 \; : \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$

On accepte H_0 si $\frac{s_1^2}{s_2^2} \in \left[F_{\frac{\alpha}{2}}^{(n_1-1,n_2-1)}, F_{1-\frac{\alpha}{2}}^{(n_1-1,n_2-1)} \right]$

Rappel pour l'usage des tables : $F_{\frac{\alpha}{2}}^{(n_1-1,n_2-1)} = \frac{1}{F_{\frac{1-\alpha}{2}}^{(n_2-1,n_1-1)}}$.

Ce résultat est peu robuste par rapport à la condition gaussienne.

Test sur la corrélation d'un couple gaussien

On considère un couple de v.a. gaussiennes (X, Y), où $X \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1^2)$, $Y \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2^2)$, et (X, Y) est de covariance théorique ρ.

On considère un $n-$échantillon $(X_1, Y_1), \ldots, (X_n, Y_n)$. On veut tester $H_0 \ : \rho = 0$ vs. $H_1 \ : \rho \neq 0$.

Statistique de test ? La corrélation empirique ?

$$R = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \overline{Y})^2}}$$

En pratique on considère : $T = \frac{\sqrt{n-2}R}{\sqrt{1-R^2}} \rightsquigarrow t(n-2)$ sous H_0.