Tests

Frédéric Bertrand[1]

[1]IRMA, Université de Strasbourg
Strasbourg, France

École doctorale SVS 12-01-2010

« Lorsque vous avez éliminé l'impossible, ce qui reste, si improbable soit-il, est nécessairement la vérité ».

De Arthur Conan Doyle, D'après « Le signe des Quatre ».

Sommaire

1 Introduction

2 Test entre deux hypothèses simples

3 Test entre hypothèses composées

4 Test de comparaison

Définition

Un test est un mécanisme qui permet de trancher entre deux hypothèses à la vue des résultats d'un échantillon.

Exemples

Dans les cas qui nous intéressent, ces hypothèses porteront sur des estimations (comparaison d'une moyenne, égalité de variances, nature d'une loi de probabilité...).

Soient $\mathcal{H}_0$ et $\mathcal{H}_1$ deux hypothèses, dont une et une seule est vraie. La décision du test aboutira à choisir $\mathcal{H}_0$ ou $\mathcal{H}_1$. Il y a donc quatre cas possibles dont les probabilités sont résumées dans le tableau ci-dessous :

	$\mathcal{H}_0$ vraie	$\mathcal{H}_1$ vraie
$\mathcal{H}_0$ décidée	$1-\alpha$	β
$\mathcal{H}_1$ décidée	α	$1-\beta$

où α **et** β **sont les erreurs de première et deuxième espèce.**

- α est la probabilité de décider $\mathcal{H}_1$ alors que $\mathcal{H}_0$ est vraie.
- β est la probabilité de décider $\mathcal{H}_0$ alors que $\mathcal{H}_1$ est vraie.

Remarque

Ces deux erreurs sont antogonistes, plus α sera grand (respectivement petit), plus β sera petit (respectivement grand).

Remarque

Le fait d'imposer un α faible conduit à une règle de décision plus stricte qui aboutit le plus souvent à n'abandonner l'hypothèse $\mathcal{H}_0$ que dans des cas rarissimes et donc à conserver cette hypothèse quelque fois à tort.

Remarque

Le compromis entre les valeurs de α et β est donc souhaitable bien que difficile à réaliser.

Définition

Nous appelons puissance d'un test la quantité $1 - \beta$.

Remarque

Dans la pratique des tests statistiques, il est de règle de se fixer α comme donné (les valeurs les plus courantes sont 5%, 1% ou 10%) de préférence en fonction du risque de première espèce.

Remarque

L'hypothèse nulle $\mathcal{H}_0$ joue le plus souvent un rôle prédominant par rapport à l'hypothèse alternative $\mathcal{H}_1$. Cela est la conséquence du fait que $\mathcal{H}_0$ joue le rôle d'hypothèse de référence alors que $\mathcal{H}_1$ est souvent limitée à l'hypothèse contraire.

Exemple

Nous pouvons avoir :

$$\mathcal{H}_0 : \mu = \mu_0$$

ce qui est relativement facile à tester et dans ce cas :

$$\mathcal{H}_1 : \mu \neq \mu_0.$$

Remarque

Cette pratique est liée au fait que l'évaluation d'un test passe par l'évaluation de fonctions complexes qui ont été tabulées pour de nombreuses valeurs de α mais ne sont pas connues pour tout α. **Nous sommes donc amené à choisir a priori α.**

Remarque

Cependant, l'apparition de plus en plus fréquente de processus numériques d'approximation rapides et précis permet une autre approche consistant à rechercher la plus petite valeur de α pour laquelle l'hypothèse nulle $\mathcal{H}_0$ reste vraie.

Remarque

Avant d'appliquer tout test statistique, il s'agit de bien définir le problème posé. En effet, selon les hypothèses formulées, nous appliquons soit un test bilatéral, soit un test unilatéral.

Définition

Un test bilatéral s'applique quand on cherche une différence entre deux estimations, ou entre une estimation et une valeur donnée sans se préoccuper du signe ou du sens de la différence. Dans ce cas, la zone de rejet (cf section suivante) de l'hypothèse principale se fait de part et d'autre de la distribution de référence.

Définition

Un test unilatéral s'applique quand nous cherchons à savoir si une estimation est supérieure (ou inférieure) à une autre ou à une valeur donnée. La zone de rejet de l'hypothèse principale est située d'un seul côté de la distribution de probabilité de référence.

Exemples

Certains tests comme l'analyse de la variance ou le test du χ^2 sont pratiquement toujours unilatéraux.

Quelle est la démarche générale ?

L'erreur de première espèce α étant fixée, il faut choisir une **variable de décision**, variable qui doit apporter de l'information sur le problème posé, à savoir le choix entre les deux hypothèses. La loi de cette variable doit être parfaitement connue dans au moins une hypothèse (le plus souvent pour l'hypothèse nulle $\mathcal{H}_0$) afin de ne pas introduire de nouvelles inconnues dans le problème.

Définition

Nous appelons alors région critique, et nous notons W, l'ensemble des valeurs de la variable de décision qui conduisent à écarter l'hypothèse nulle $\mathcal{H}_0$ au profit de l'hypothèse alternative $\mathcal{H}_1$.

Remarque

Nous pouvons relier la région critique W à l'erreur de première espèce α par

$$\mathbb{P}_{\mathcal{H}_0}[W] = \alpha.$$

Définition

Nous appelons région d'acceptation, et nous notons $\overline{W}$ la région complémentaire de la région critique W.

Remarque

Nous avons également des relations avec les erreurs de première et deuxième espèce, que nous noterons (abusivement) :

$$\mathbb{P}_{\mathcal{H}_0}\left[\overline{W}\right] = 1 - \alpha$$

et

$$\mathbb{P}_{\mathcal{H}_1}\left[W\right] = 1 - \beta.$$

Définition

La région ou la zone d'acceptation correspond à l'intervalle dans lequel les différences observées entre les réalisations et la théorie sont attribuables aux fluctuations d'échantillonnage.

Définition

La région critique ou zone de rejet correspond donc aux intervalles dans lesquels les différences sont trop grandes pour être le fruit du hasard d'échantillonnage.

La construction d'un test est la détermination a priori de la région critique W sans connaitre le résultat de l'expérience.

La démarche à adopter

1. Choix des hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Allure de la région critique en fonction de $\mathcal{H}_1$.
4. Calcul de la région critique en fonction de α.
5. Calcul éventuel de la puissance du test $1 - \beta$.
6. Calcul expérimental de la variable de décision.
7. Conclusion du test : acceptation ou rejet de l'hypothèse nulle $\mathcal{H}_0$.

Plusieurs tests de conception très différente sont souvent disponibles pour soumettre à une épreuve de vérité une hypothèse principale.

Définition

Dans un tel cas, le test qui fournit l'erreur β la plus petite, pour une même valeur de α, est par définition le plus puissant (celui ayant la plus grande valeur de la puissance de test $1-\beta$).

En effet, il peut détecter les plus petites différences entre les populations sans pour autant augmenter l'erreur de première espèce.

La majorité des tests statistiques repose sur le respect d'un certain nombre de conditions. Selon le degré de respect de ces conditions d'application, la validité des résultats se trouve plus ou moins affectée et elle l'est d'autant plus que le test est moins robuste. Ainsi, la robustesse d'un test équivaut à sa tolérance vis-à-vis du respect des conditions.

Si le statisticien dispose de plusieurs tests pour vérifier une hypothèse, il choisira bien sûr le plus puissant et le plus robuste.

Remarque

Les tests peu puissants augmentent la probabilité de commettre une erreur de deuxième espèce. Or, cette erreur peut s'avérer particulièrement grave.

Exemple

En effet, en médecine par exemple, une analyse qui classerait comme malade un individu bien portant peut avoir des conséquences aussi graves qu'une analyse qui classerait comme bien portants des individus malades (erreur de première espèce).

Remarque

Dans de tels cas, il y a intérêt à tracer la courbe de puissance du test, aussi appelée courbe caractéristique d'efficacité qui indique la probabilité de prendre une bonne décision si l'hypothèse alternative $\mathcal{H}_1$ est vraie.

Remarque

La puissance est mesurée par la valeur de $1 - \beta$ pour une erreur de première espèce α donnée.

Pour comparer les moyennes, les variances ou les autres paramètres estimés de deux échantillons, il faut prendre en considération la technique conduisant à la constitution des deux échantillons.

Définition

Si la sélection des éléments est aléatoire, et si le choix des éléments du premier échantillon n'a aucune influence sur le choix des éléments du second, les deux échantillons sont alors appelés indépendants.

Définition

Si nous prélevons aléatoirement des paires d'éléments, et non les éléments eux-mêmes, nous constituons deux échantillons appariés.

Remarque

Dans ce cas, le premier élément de chaque paire appartient au premier échantillon et le deuxième est affecté au second. Parfois, la paire d'éléments peut se rapporter au même individu sur lequel nous mesurons la même variable à deux occasions différentes, par deux moyens différents par exemple.

La technique de l'échantillonnage apparié présente l'avantage d'éliminer un maximum de sources de variations non reliées au facteur que nous étudions.

Règle générale

Plus les critères d'appariement des données sont nombreux, plus grand sera cet avantage.

Dans ce qui suit, nous allons aborder quelques tests classiques. Cette liste ne se veut pas exhaustive. Reportez-vous à des ouvrages plus spécialisés pour une approche plus systématique des tests statistiques.

Sommaire

1. Introduction
2. Test entre deux hypothèses simples
3. Test entre hypothèses composées
4. Test de comparaison

Soit X une variable aléatoire de densité $f(\mathbf{x},\theta)$ où θ est un paramètre réel inconnu. $L(\mathbf{x},\theta)$ désigne la densité de l'échantillon $\mathbf{x}$.

Un test entre deux hypothèses simples se traduit par :

$$\begin{cases} \mathcal{H}_0 & : \quad \theta = \theta_0 \\ \mathcal{H}_1 & : \quad \theta = \theta_1. \end{cases}$$

Supposons l'erreur de première espèce α connue.

Rappel

Nous avons vu que nous pouvons relier α à une région de l'espace $\mathbb{R}^n$ par :

$$\mathbb{P}_{\mathcal{H}_0}[W] = \alpha = \int_W L(\mathbf{x},\theta_0)d\mathbf{x}.$$

Nous cherchons par ailleurs le test le plus puissant, donc celui qui maximise la puissance du test :

$$\mathbb{P}_{\mathcal{H}_1}[W] = 1 - \beta = \int_W L(\mathbf{x}, \theta_1) d\mathbf{x}.$$

La solution est donnée par le théorème de Neyman et Pearson.

Théorème de Neyman et Pearson

La région critique optimale est définie par l'ensemble des points $\mathbf{x}$ de $\mathbb{R}^n$ tels que :

$$T_{NP} = \frac{L(\mathbf{x}, \theta_1)}{L(\mathbf{x}, \theta_0)} \geqslant k_\alpha,$$

où k_α est le quantile associé à la statistique T_{NP} au seuil α.

En conséquence de ce théorème, nous pouvons établir deux définitions :

Définition

Le test est sans biais

$$\text{si} \quad 1 - \beta > \alpha.$$

Définition

Le test est convergent lorsque la condition suivante est vérifiée

$$\text{si} \quad \text{lorsque} \quad n \to 1 \quad \text{alors} \quad 1 - \beta \to 1.$$

Exemple

Soit X une variable aléatoire normale de moyenne μ et de variance σ^2 connue. À la vue d'un échantillon de n réalisations indépendantes X_i, nous désirons savoir si la moyenne μ est égale à μ_0 ou à μ_1, ce qui se résume par :

$$\left\{ \begin{array}{lcl} \mathcal{H}_0 & : & \mu = \mu_0 \\ \mathcal{H}_1 & : & \mu = \mu_1 \end{array} \right. .$$

Définition

Les fonctions de vraisemblance, ou densité, de l'échantillon sont :

$$L(\mathbf{x}, \mu_0) = \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i - \mu_0)^2\right)$$

et

$$L(\mathbf{x}, \mu_1) = \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i - \mu_1)^2\right).$$

Définition

La région critique W est définie par le ratio de ces deux fonctions.

Méthode

En passant par le logarithme, nous obtenons facilement :

$$\sum_{i=1}^{n}(X_i-\mu_1)^2-\sum_{i=1}^{n}(X_i-\mu_0)^2 \leqslant -2\sigma^2\ln(k_\alpha).$$

En posant $\widehat{\mu}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$, on obtient :

$$\left(\widehat{\mu}_n-\frac{\mu_0+\mu_1}{2}\right)(\mu_0-\mu_1)\leqslant -\frac{\sigma^2\ln(k_\alpha)}{n}.$$

Si $\mu_0<\mu_1$, alors nous avons

$$\widehat{\mu}_n \geqslant \frac{\mu_0+\mu_1}{2}-\frac{\sigma^2\ln(k_\alpha)}{n(\mu_0-\mu_1)}=\lambda.$$

Nous supposera par la suite que $\mu_1 > \mu_0$. Le cas contraire se traiterait de la même manière.

Définition

La région critique W est donc définie par l'inégalité

$$\widehat{\mu}_n \geqslant \lambda$$

qu'il faut maintenant déterminer.

Méthode

Pour cela, nous introduisons l'erreur de première espèce α qui a été définie préalablement par :

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[W\right].$$

Règle de décision

Nous décidons l'hypothèse alternative $\mathcal{H}_1$ si $\widehat{\mu}_n \geqslant \lambda$. Donc

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[\widehat{\mu}_n \geqslant \lambda\right]$$

où $\widehat{\mu}_n$ est la variable aléatoire dont $\widehat{\mu}_n(obs)$ est une réalisation. X étant une variable aléatoire gaussienne, la distribution de $\widehat{\mu}_n$ est également gaussienne de moyenne μ et de variance $\dfrac{\sigma^2}{n}$.

Détermination de α

Nous avons alors (en supposant que l'hypothèse nulle $\mathcal{H}_0$ est vraie, c'est-à-dire que $\mu = \mu_0$)

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[\widehat{\mu}_n \geqslant \lambda\right]$$

avec

$$\widehat{\mu}_n = \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

ou encore

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[\frac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}} \geqslant \frac{\lambda - \mu_0}{\sigma/\sqrt{n}}\right].$$

La quantité $Z = \dfrac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}}$ suit une $\mathcal{N}(0,1)$. D'où, nous obtenons :

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[Z \geqslant \frac{\lambda - \mu_0}{\sigma/\sqrt{n}}\right]$$

où

$$Z = \mathcal{N}(0,1).$$

Si la valeur de α est fixée, nous pouvons par lecture dans une table de la loi normale centrée et réduite, trouver la valeur de $\dfrac{\lambda - \mu_0}{\sigma/\sqrt{n}}$ et donc celle de λ.

Règle de décision du test

La règle de décision du test est donc :

- **Si $\widehat{\mu}_n(obs) \geqslant \lambda$, alors nous décidons de refuser l'hypothèse nulle $\mathcal{H}_0$ et par conséquent nous acceptons l'hypothèse alternative $\mathcal{H}_1$ au risque α.**
- **Si $\widehat{\mu}_n(obs) < \lambda$, alors nous décidons de ne pas refuser l'hypothèse nulle $\mathcal{H}_0$ et par conséquent nous acceptons l'hypothèse nulle $\mathcal{H}_0$.**

Puissance du test

Par un raisonnement équivalent, nous pouvons évaluer l'erreur de deuxième espèce et donc la puissance du test, à savoir calculer la quantité suivante :

$$\begin{aligned} \beta &= \mathbb{P}_{\mathcal{H}_1}\left[\widehat{\mu}_n < \lambda\right] \\ &= \mathbb{P}_{\mathcal{H}_1}\left[Z < \frac{\lambda - \mu_1}{\sigma/\sqrt{n}}\right], \end{aligned}$$

où $Z = \dfrac{\widehat{\mu}_n - \mu_1}{\sigma/\sqrt{n}}$ est une variable aléatoire normale centrée et réduite $\mathcal{N}(0, 1)$.

Le raisonnement précédent s'applique jusqu'à la détermination de λ.

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[\frac{\widehat{\mu}_n - \mu_0}{S_{n,c}/\sqrt{n}} \geqslant \frac{\lambda - \mu_0}{S_{n,c}/\sqrt{n}}\right]$$

où $S_{n,c}^2$ désigne l'estimateur de la variance inconnue σ^2.

Attention

La quantité $T_{n-1} = \dfrac{\widehat{\mu}_n - \mu_0}{S_{n,c}/\sqrt{n}}$ ne suit plus une loi normale centrée et réduite car le dénominateur n'est plus une constante mais une variable aléatoire puisque $S_{n,c}^2$ est l'estimateur de la variance inconnue σ^2 de la variable X.

Définition

$S_{n,c}$ est défini comme étant la racine carrée de :

$$S_{n,c}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \widehat{\mu}_n)^2 .$$

Définition

Par construction, $(n-1)S_{n,c}^2$ suit une loi du χ^2. T_{n-1} est donc une variable aléatoire suivant une loi de Student à $n-1$ degrés de liberté. Ce qui donne :

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[T_{n-1} \geqslant \frac{\lambda - \mu_0}{S_{n,c}/\sqrt{n}}\right].$$

Là encore, il est possible grâce à une table de la loi de Student de trouver la valeur du seuil et donc celle de λ. La règle de décision est toujours la même.

De même, par un raisonnement analogue, nous accédons à l'erreur de deuxième espèce β et à la puissance du test $1-\beta$.

$$\begin{aligned}\beta &= \mathbb{P}_{\mathcal{H}_1}\left[\widehat{\mu}_n < \lambda\right] \\ &= \mathbb{P}_{\mathcal{H}_1}\left[T_{n-1} < \frac{\lambda-\mu_1}{S_{n,c}/\sqrt{n}}\right],\end{aligned}$$

où $T_{n-1} = \dfrac{\widehat{\mu}_n-\mu_1}{S_{n,c}/\sqrt{n}}$ est une variable aléatoire de Student à $n-1$ degrés de liberté.

Test d'une variance de loi normale

Soit X une variable aléatoire normale de moyenne μ connue. Nous supposons que la variance σ^2 inconnue ne peut prendre que deux valeurs σ_0^2 et σ_1^2. À la vue d'un échantillon de n réalisations indépendantes X_i, nous désirons savoir si la variance σ^2 est égale à σ_0^2 ou à σ_1^2, ce qui se résume par :

$$\left\{\begin{array}{lcl} \mathcal{H}_0 & : & \sigma^2 = \sigma_0^2 \\ \mathcal{H}_1 & : & \sigma^2 = \sigma_1^2 \end{array}\right. .$$

Définition

L'estimateur de la variance est égal à

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

Remarque

Nous utilisons l'estimateur S_n^2 et non pas S_{n-1}^2 car la moyenne μ ici est connue.

Définition

Les fonctions de vraisemblance, ou densité, de l'échantillon sont égales à :

$$L(\mathbf{x}, \sigma_0) = \left(\frac{1}{\sqrt{2\pi\sigma_0}}\right)^n \exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(X_i - \mu)^2\right)$$

et

$$L(\mathbf{x}, \sigma_1) = \left(\frac{1}{\sqrt{2\pi\sigma_1}}\right)^n \exp\left(-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(X_i - \mu)^2\right).$$

Définition

La région critique W est définie par le ratio de ces deux fonctions.

Méthode

En passant par le logarithme, nous obtenons facilement :

$$n \ln\left(\frac{\sigma_0}{\sigma_1}\right) + \frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^2\left(\frac{1}{\sigma_0^2}-\frac{1}{\sigma_1^2}\right) \geqslant \ln(k_\alpha).$$

Si $\sigma_1 > \sigma_0$, nous aboutissons à

$$\frac{nS_n^2}{\sigma_0^2} \geqslant \frac{2\sigma_1^2}{\sigma_1^2-\sigma_0^2}\left(\ln(k_\alpha) - n\ln\left(\frac{\sigma_0}{\sigma_1}\right)\right).$$

Méthode

La valeur de k_α est déterminée à partir de l'erreur de première espèce α. La quantité $\dfrac{nS_n^2}{\sigma_0^2}$ suit une loi du χ^2 à n degrés de liberté. La valeur critique est donc lue dans une table de la loi du χ^2.

Remarque

Le cas $\sigma_1 < \sigma_0$ est laissé au soin du lecteur. Il est conseillé de le détailler.

La variable de décision

$$S_{n,c}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \widehat{\mu}_n)^2$$

qui est telle que $\dfrac{(n-1)S_{n,c}^2}{\sigma_0^2}$ suit une loi du χ^2 à $n-1$ degrés de liberté.

La région critique W

La région critique W est définie par $S_{n,c}^2 \geqslant k_\alpha$ et k_α est déterminée par

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[S_{n,c}^2 \geqslant k_\alpha\right] = \mathbb{P}_{\mathcal{H}_0}\left[\chi_{n-1}^2 \geqslant (n-1)\frac{k_\alpha}{\sigma_0^2}\right].$$

La règle de décision du test est donc :

- **Si $S^2_{n,c,obs} \geqslant k_\alpha$, alors nous décidons de refuser l'hypothèse nulle $\mathcal{H}_0$ et par conséquent d'accepter l'hypothèse alternative $\mathcal{H}_1$ au risque α.**
- **Si $S^2_{n,c,obs} < k_\alpha$, alors nous décidons de ne pas refuser l'hypothèse nulle $\mathcal{H}_0$ et par conséquent d'accepter l'hypothèse nulle $\mathcal{H}_0$.**

Sommaire

1. Introduction
2. Test entre deux hypothèses simples
3. Test entre hypothèses composées
4. Test de comparaison

Test unilatéral

Soit X une v.a. normale de moyenne μ et de variance σ^2 connue. À la vue d'un échantillon de n réalisations indépendantes X_i, on veut choisir entre les deux hypothèses :

$$\mathcal{H}_0 : \mu = \mu_0$$

contre

$$\mathcal{H}_1 : \mu > \mu_0.$$

Comme toujours, l'erreur de première espèce α est fixée. Par ailleurs, la moyenne μ sera estimée par l'estimateur $\widehat{\mu}_n$.

Construction du test

La construction du test est similaire à ce que nous avons vu pour le cas du test simple d'une moyenne. Nous aboutissons à :

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[\frac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}} \geqslant \frac{\lambda - \mu_0}{\sigma/\sqrt{n}}\right],$$

où $\widehat{\mu}_n$ suit une loi normale de moyenne μ et de variance σ^2/n.

Remarque

La valeur du seuil de décision λ est indépendante de la valeur de μ sous l'hypothèse alternative $\mathcal{H}_1$. Il s'en suit que le test est uniformément le plus puissant.

Décision et conclusion

La variable $Z = \dfrac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}}$ suit une loi normale centrée et réduite.
La valeur critique est donc lue dans une table de la loi normale centrée et réduite.

Remarque

Il en est de même pour l'erreur de deuxième espèce β et pour la puissance du test $1 - \beta$.

Test bilatéral

Soit X une v.a. normale de moyenne μ et de variance σ^2 connue. À la vue d'un échantillon de n réalisations indépendantes X_i, nous voulons choisir entre les deux hypothèses :

$$\mathcal{H}_0 : \mu = \mu_0$$

contre

$$\mathcal{H}_1 : \mu \neq \mu_0.$$

Comme toujours, l'erreur de première espèce α est fixée. Par ailleurs, la moyenne μ sera estimée par l'estimateur $\widehat{\mu}_n$.

Construction du test

La construction du test est obtenue en remarquant que l'hypothèse alternative $\mathcal{H}_1$ peut se décomposer en deux hypothèses simples :

$$\mathcal{H}_1^{'} : \mu < \mu_0$$

ou

$$\mathcal{H}_1^{''} : \mu > \mu_0.$$

Remarque

À chacune de ces deux hypothèses sera associé un seuil de décision $\lambda^{'}$ et $\lambda^{''}$. Il s'en suit que le test n'est plus uniformément le plus puissant (UMP) car le seuil de décision λ dépend du sens de l'inégalité.

Détermination des seuils de décision

La détemination des seuils est simple puisque les deux hypothèses $\mathcal{H}_1'$ et $\mathcal{H}_1''$ sont disjointes. Nous avons

$$\alpha = \mathbb{P}_{\mathcal{H}_0}\left[\widehat{\mu}_n \geqslant \lambda''\right] + \mathbb{P}_{\mathcal{H}_0}\left[\widehat{\mu}_n \leqslant \lambda'\right] = \alpha'' + \alpha'.$$

Remarque

Comme la loi de $\widehat{\mu}_n$ est symétrique (loi de Laplace-Gauss), nous prenons $\alpha'' = \alpha' = \alpha/2$. Ce qui conduit à des valeurs de λ symétriques par rapport à μ_0.

Construction du test

$$\frac{\alpha}{2} = \mathbb{P}_{\mathcal{H}_0}\left[\frac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}} \geqslant \frac{\lambda - \mu_0}{\sigma/\sqrt{n}}\right],$$

où $\widehat{\mu}_n$ suit une loi normale de moyenne μ et de variance σ^2/n

Décision et conclusion

La variable $Z = \frac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}}$ suit une loi normale centrée et réduite. La valeur critique est donc lue dans une table de la loi normale centrée et réduite.

Remarque

Il en est de même pour l'erreur de deuxième espèce β et pour la puissance du test $1 - \beta$.

Construction du test

Les deux tests, unilatéral et bilatéral, se construisent selon le même procédé que précédemment.

Décision et conclusion

Les valeurs critiques sont lues dans une table de la loi de Student à $n - 1$ degrés de liberté.

Construction du test

Les deux tests, unilatéral et bilatéral, se construisent selon le même procédé que précédemment.

Décision et conclusion

Les valeurs critiques sont lues dans une table de la loi du χ^2 à n degrés de liberté.

Construction du test

Les deux tests, unilatéral et bilatéral, se construisent selon le même procédé que précédemment.

Décision et conclusion

Les valeurs critiques sont lues dans une table de la loi du χ^2 à $n-1$ degrés de liberté.

Sommaire

1. Introduction
2. Test entre deux hypothèses simples
3. Test entre hypothèses composées
4. Test de comparaison

Test de comparaison

Soient X_1 et X_2 deux variables aléatoires définies sur deux populations mères comparables (éventuellement égales). La loi de X_1 (respectivement X_2) dépend d'un paramètre inconnu θ_1 (respectivement θ_2). Nous souhaitons tester l'hypothèse nulle

$$\mathcal{H}_0 : \text{ces deux paramètres sont égaux}$$

contre l'hypothèse alternative

$$\mathcal{H}_1 : \text{ces deux paramètres sont différents,}$$

soit

$$\begin{cases} \mathcal{H}_0 & : & \theta_1 = \theta_2 \\ \mathcal{H}_1 & : & \theta_1 \neq \theta_2 \end{cases}.$$

Méthode

Pour effectuer ce test, nous disposons d'un échantillon de taille n_1 (respectivement n_2) de X_1 (respectivement X_2) permettant une estimation ponctuelle T_{n_1} (respectivement T_{n_2}) de θ_1 (respectivement θ_2). Nous supposons de plus que les variables aléatoires X_1 et X_2 sont normales ou approximativement normales. En supposant $\mathcal{H}_0$ vraie, nous déterminons un risque de première espèce α, une zone de rejet associée à deux valeurs critiques c_1 et c_2 telles que

$$\mathbb{P}\left[Z < c_1\right] = \mathbb{P}\left[Z > c_2\right] = \frac{\alpha}{2}$$

où Z est une fonction de T_{n_1} et T_{n_2}.

La règle de décision du test est donc :

- **Si z appartient à la région de critique alors nous décidons de refuser l'hypothèse nulle $\mathcal{H}_0$ et par conséquent d'accepter l'hypothèse alternative $\mathcal{H}_1$ au risque α.**
- **Si z appartient à la région d'acceptation alors nous décidons de nepas refuser l'hypothèse nulle $\mathcal{H}_0$ et par conséquent d'accepter l'hypothèse nulle $\mathcal{H}_0$.**

Comparaison de deux moyennes

Soient X_1 et X_2 deux lois normales de moyennes μ_1 et μ_2, et de variances σ_1^2 et σ_2^2. Nous testons

$$\begin{cases} \mathcal{H}_0 & : \quad \mu_1 = \mu_2 \\ \mathcal{H}_1 & : \quad \mu_1 \neq \mu_2 \end{cases}.$$

Nous utilisons le test de Student (dans sa version la plus générale).

Nous disposons de deux échantillons de tailles n_1 et n_2 sur lesquels nous pouvons construire des estimateurs $\widehat{\mu}_{n_1}$ et $\widehat{\mu}_{n_2}$ des moyennes μ_1 et μ_2 et des estimateurs $S^2_{n_1}$ et $S^2_{n_2}$ des deux variances.

Premier cas : si les variances σ_1^2 et σ_2^2 sont connues.

Nous calculons la réalisation de la variable aléatoire

$$Z = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}.$$

Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ au risque α si $z \notin \left[-u_{1-\alpha/2}; +u_{1-\alpha/2}\right]$ où la valeur $u_{1-\alpha/2}$ est lue dans la table de la loi normale centrée et réduite.

Deuxième cas : si les variances σ_1^2 et σ_2^2 sont inconnues.

Il faut alors tenir compte de la taille des deux échantillons. Trois cas se distinguent.

Premier cas

Si les tailles n_1 et n_2 sont toutes les deux supérieures à 30, nous calculons la réalisation de la variable aléatoire

$$Z = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\dfrac{S_{n_1}^2}{n_1 - 1} + \dfrac{S_{n_2}^2}{n_2 - 1}}}.$$

Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ au risque α si $z \notin \left[-u_{1-\alpha/2}; +u_{1-\alpha/2}\right]$ où la valeur $u_{1-\alpha/2}$ est lue dans la table de la loi normale centrée et réduite.

Deuxième cas

Si la taille n_1 ou la taille n_2 est inférieure à 30 et si $\sigma_1^2 = \sigma_2^2$ nous calculons la réalisation de la variable aléatoire

$$T = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\widehat{\sigma}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}.$$

où

$$\widehat{\sigma} = \sqrt{\frac{n_1 S_{n_1}^2 + n_2 S_{n_2}^2}{n_1 + n_2 - 2}}.$$

Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ au risque α si $t \notin \left[-t_{1-\frac{\alpha}{2};n_1+n_2-2}; +t_{1-\frac{\alpha}{2};n_1+n_2-2}\right]$ où $t_{1-\frac{\alpha}{2};n_1+n_2-2}$ est lue dans la table de la loi de Student à $n_1 + n_2 - 2$ degrés de liberté.

Troisième cas

Si la taille n_1 ou la taille n_2 est inférieure à 30 et $\sigma_1^2 \neq \sigma_2^2$ nous calculons la réalisation de la variable aléatoire

$$T = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\dfrac{S_{n_1}^2}{n_1 - 1} + \dfrac{S_{n_2}^2}{n_2 - 1}}}.$$

Troisième cas - Suite

Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ au risque α si $t \notin \left[-t_{1-\frac{\alpha}{2};\nu}; +t_{1-\frac{\alpha}{2};\nu}\right]$ où $t_{1-\frac{\alpha}{2};\nu}$ est lue dans la table de la loi de Student à ν degrés de liberté. ν est l'entier le plus proche de

$$\frac{\left[\dfrac{S_{n_1}^2}{n_1-1}+\dfrac{S_{n_2}^2}{n_2-1}\right]^2}{\dfrac{S_{n_1}^4}{(n_1-1)n_1^2}+\dfrac{S_{n_2}^4}{(n_2-1)n_2^2}}$$

Remarque :

Le test de Student est assez robuste mais si nous nous éloignons trop de la condition de normalité qui est nécessaire pour procéder à ce test, il est préférable d'utiliser un test non paramétrique, comme par exemple le test de Mann-Whitney.

Comparaison de deux variances

Avec les mêmes notations que précédemment, nous testons

$$\left\{ \begin{array}{lll} \mathcal{H}_0 & : & \sigma_1^2 = \sigma_2^2 \\ \mathcal{H}_1 & : & \sigma_1^2 \neq \sigma_2^2 \end{array} \right. .$$

Nous calculons les trois quantités suivantes :

$$S^2_{n_1,c} = \frac{n_1 S^2_{n_1}}{n_1 - 1}, \quad S^2_{n_2,c} = \frac{n_2 S^2_{n_2}}{n_2 - 1} \quad \text{et} \quad \frac{S^2_{n_1,c}}{S^2_{n_2,c}} = F.$$

Décision

Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ au risque α si

$$F \notin \left[F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1); F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)\right]$$

où la valeur F_α est lue dans la table de la loi de Fisher-Snedecor à $n_1 - 1$ et $n_2 - 1$ degrés de liberté.

Remarque

$$F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) = \frac{1}{F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)}.$$