Probabilités pour ingénieurs (STT-2920)

Chapitre 13 : Tests d'hyptohèse

Ce document, inspiré des notes de cours de Claude Bélisle,
a été construit par Maxime Genest et Line Baribeau

Département de mathématiques et de statistique

Automne 2024

UNIVERSITÉ LAVAL

Faculté des sciences et de génie
Département de mathématiques et
de statistique

Table des matières

Test d’hypothèse (Présentation)

- Un test d’hypothèse est un procédé de l’inférence statistique permettant de tester la validité d’une hypothèse de recherche. Dans un test d’hypothèse, deux hypothèses se confrontent, l’hypothèse H_0 nommée l’hypothèse nulle et l’hypothèse H_1 nommée la contre-hypothèse (ou hypothèse alternative)

- Dans ce chapitre, nous nous attarderons aux tests d’hypothèse paramétriques, i.e. dont les hypothèses portent sur un paramètre, disons θ.

- L’hypothèse H_0 est l’hypothèse qui est acceptée jusqu’à présent, par norme établie ou étude précédente. Elle sera de la forme $H_0 : \theta = \theta_0$. Cette hypothèse est est considérée vraie jusqu’à preuve du contraire dans la mécannique du test.

- L’hypothèse H_1 suggère une modification possible de θ. Cette hypothèse peut être de une des 3 formes suivantes :

$H_1 : \theta < \theta_0$	$H_1 : \theta \neq \theta_0$	$H_1 : \theta > \theta_0$
test unilatéral à gauche	test bilatéral	test unilatéral à droite

Dans un test d’hypothèse, on se sert d’un échantillon aléatoire (plus précisément d’une statistique de l’échantillon) pour arriver à une des deux décisions suivantes : (on rejette H_0 i.e. on confirme l’hypothèse H_1) ou bien (on ne rejette pas l’hypothèse H_0, i.e. on ne peut pas confirmer l’hypothèse H_1)

Test d'hypothèse sur la moyenne μ d'une population

Les hypothèses et la statistique utilisée

Dans le cas des tests d'hypothèse sur la moyenne μ d'une population,

l'hypothèse H_0 sera de la forme $H_0 : \mu = \mu_0$

l'hypothèse H_1 peut être d'une des 3 formes :

$H_1 : \mu < \mu_0$ ou bien $H_1 : \mu \neq \mu_0$ ou bien $H_1 : \mu > \mu_0$.

Pour prendre la décision, on construira une règle de décision qui s'appuiera sur la statistique $\overline{X}$, ou plus précisément sur la satistique $Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ et sur le fait que

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} = Z \sim \mathcal{N}(0,1) \qquad \text{sous l'hypothèse } H_0$$

Vous aurez compris qu'on se place ici dans le cas d'une population même normale et σ connu, on traitera les autres cas de façon analogue plus tard.

Test d'hypothèse sur la moyenne μ d'une population
Le seuil de signification et la zone de rejet

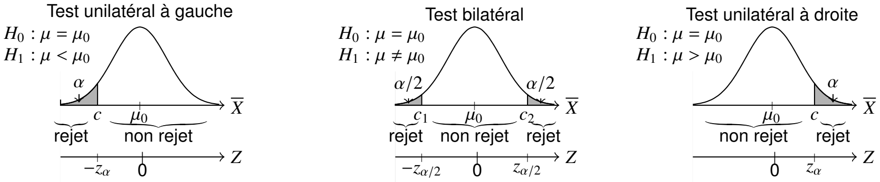
Le seuil de signification, noté α, est une probabilité (généralement faible : 0,1, 0,05, 0,01) permettant de déterminer si une observation est **une évidence statistique** contre H_0. On considère qu'une observation (ici $\overline{x}$) est **une évidence statistique** contre H_0 si sa valeur se retrouve dans une zone dont la probabilité (α) est faible si H_0 est vraie. α est la probabilité de rejeter H_0 si H_0 est vraie, c'est-à-dire de rejeter à tort H_0.

$$\alpha = \mathbb{P}\,[\text{Rejeter } H_0 \mid H_0 \text{ est vraie}]$$

α est la probabilité associée à la zone de rejet de H_0.

La zone de rejet de H_0 (ou si vous préférez la zone de validation de H_1) sera une zone significativement éloignée de μ_0. Son emplacement, à gauche, bilatéral ou bien à droite, dépendra de la forme de l'hypothèse alternative H_1.

Représentation graphique de la zone de rejet selon le type de test



Test d'hypothèse sur la moyenne μ d'une population
Le seuil de signification et la zone de rejet

La règle de décision sera donc d'une des 3 formes suivantes selon la forme de l'hypothèse H_1.

Hypothèse H_1	Règle de décision
$\mu < \mu_0$	On rejette H_0 si $Z < -z_\alpha$
$\mu \neq \mu_0$	On rejette H_0 si $\lvert Z \rvert > z_{\alpha/2}$
$\mu > \mu_0$	On rejette H_0 si $Z > z_\alpha$

On prend **la décision** de rejeter H_0 ou non en calculant $Z_{\text{obs}} = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ et en vérifiant si l'inégalité approprié du tableau précédent est satisfaite ou non.

La conclusion s'énoncera ainsi dépendamment de la décision rendue.
Si rejet de H_0 : Au seuil α, on a une évidence statistique nous permettant de croire l'hypothèse H_1 est vraie.
Si non-rejet de H_0 : Au seuil α, on n'a pas d'évidence statistique nous permettant de croire que l'hypothèse H_1 est vraie.

Bien entendu, on prend soin de contextualisé la conclusion par rapport au contexte.
Également, souvent on fait l'abus de langage en disant que l'on accepte H_0 lorsque la décision rendue est de ne pas rejeter H_0. Il faut toutefois toujours se souvenir que lorsque l'on ne rejette pas H_0, cela ne veut pas dire que l'on a démontré que H_0 est vrai, cela veut simplement dire que nous n'avons pas de preuve pour dire que H_0 est faux.

Exemple 1

Une industrie produit des piles. Un ingénieur propose des modifications au processus de fabrication dans le but d'augmenter la durée de vie des piles produites. Toutefois, il ne faudrait pas que ces modifications affectent le voltage moyen des piles produites. Avec le procédé actuel, on sait que le voltage des piles suit une loi normale de moyenne 12.5 V, avec un écart-type 0.25 V. Avant de faire les modifications à grande échelle, il faudrait vérifier qu'effectivement le voltage moyen n'est pas affecté. On supposera ici que la distribution du voltage des piles produites avec le nouveau procédé est normale avec même écart-type qu'avant. On produit un échantillon de 100 piles avec le nouveau procédé et on obtient un voltage moyen de $12,53$ V. Faites un test d'hypothèse au seuil de signification $\alpha = 0.05$ permettant de vérifier si le voltage moyen du nouveau procédé diffère de 12.5 V.

Le seuil de signification empirique (la valeur p ou p-value)

Dans un test d'hypothèse, le seuil de signification α est subjectif. Si on change ce seuil, la conclusion du test peut changer. Un chercheur mal intentionné tentant de valider une hypothèse de recherche H_1 pourrait prendre un seuil α très grand. C'est donc important comme lecteur de vérifier le seuil utilisé dans la conclusion et pour le chercheur honnête de mentionné le seuil utilisé dans la conclusion.

Également, de façon complémentaire (ou de façon alternative), il est très commun de fournir le seuil de signification empirique, aussi nommé p-value, d'un test d'hypothèse. La p-value constitue, étant donné les observations obtenues (données échantillonnales), le plus petit seuil de signification possible faisant en sorte que le test mène au rejet de H_0 (i.e. à la validation de H_1). Autrement dit, la p-value est la probabilité d'avoir une statistique du test aussi extrême ou pire que ce que l'on a observé si on refaisait l'expérience dans une situation où H_0 est vraie.

Plus formellement, dans un test utilisant une statistique Z, la p-value sera donnée par

$\mathbb{P}\left[|Z| > |Z_{\text{obs}}| \;|H_0 \text{ vraie}\right]$ dans un test bilatéral

$\mathbb{P}\left[Z < Z_{\text{obs}} \;|H_0 \text{ vraie}\right]$ dans un test unilatéral à gauche

$\mathbb{P}\left[Z > Z_{\text{obs}} \;|H_0 \text{ vraie}\right]$ dans un test unilatéral à droite

Exemple de calcul d'une p-value dans le cas bilatéral

Calculez la p-value obtenue au test d'hypothèse de l'exemple 1.

Rappel des distributions d'échantillonnage

On peut faire un test d'hypothèse sur différents paramètres d'une population (ou différence de paramètres de deux populations). Par exemple, μ, σ, $\mu_1 - \mu_2$, ...

Les différentes formules de règles de décision et de p-value sont données dans les notes de cours aux pages 367, 368, 369. Voici ici un rappel des différents résultats portant sur les distributions d'échantillonnage sur lesquels s'appuient ces formules.

Conditions	Distribution de $\overline{X}$
Pop. mère normale et σ connu	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = Z \sim \mathcal{N}(0,1)$
Pop. mère normale et σ inconnu	$\frac{\overline{X}-\mu}{S/\sqrt{n}} = T \sim t_{n-1}$
n grand ($n \geq 30$)	$\frac{\overline{X}-\mu}{\widehat{\sigma}\sqrt{n}} = Z \sim \mathcal{N}(0,1)$

où l'on prend généralement $\widehat{\sigma} = s$.

Conditions	Distribution de S^2
Pop. mère normale	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Conditions	Distribution de $\widehat{P}$
$n \geq 30, np \geq 5, n(1-p) \geq 5$	$\frac{\widehat{P}-p}{\sqrt{\frac{p(1-p)}{n}}} = Z \sim \mathcal{N}(0,1)$

Conditions	Distribution de $\overline{X}_1 - \overline{X}_2$
Pop. mères normales σ_1 et σ_2 connus	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} = Z \sim \mathcal{N}(0,1)$
Pop. mères normales σ_1 et σ_2 inconnus mais $\sigma_1 = \sigma_2$	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{s_c\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = T \sim t_{n_1+n_2-2}$
n_1 et n_2 grand	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\widehat{\sigma}_1^2}{n_1}+\frac{\widehat{\sigma}_2^2}{n_2}}} = Z \sim \mathcal{N}(0,1)$

où $s_c = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

Exemple 2

Un certain type de pneu possède une durée de vie 30000 km. On tente une amélioration au procédé de fabrication. Au seuil de signification de 5%, peut-on conclure à une amélioration au procédé de fabrication si un échantillon de taille 16 donne une durée de vie moyenne de 32000 km avec un écart-type 3500 km. Donnez également une estimation de la p-value du test. On suppose que le procédé génère des pneus avec une durée de vie normalement distribué.

Solution

Soit μ la réelle durée de vie moyenne des pneus avec le nouveau procédé.

Hypothèses

$H_0 : \mu = 30000$

$H_1 : \mu > 30000$

On a donc un test unilatéral à droite au seuil $\alpha = 5\%$

Collecte des données échantillonnales

On a $n = 16$, $\overline{x} = 32000$ km et $s = 3500$ km.

Distribution d'échantillonnage à utiliser

Puisque que l'on a une population mère normale et σ est inconnu, on a que $T = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$

Règle de décision

La valeur critique est $t_{15;0.05} = 1.753$.

Règle de décision : Rejet de H_0 si $T > 1.753$, non rejet de H_0 sinon.

Valeur observée et décision

Valeur observée : $T_{\text{obs}} = \frac{\overline{x}-\mu_0}{s/\sqrt{n}} = \frac{32000-30000}{3500/\sqrt{16}} = 2.29.$

Décison : Puisque $T_{\text{obs}} = 2.29 > 1.753$, on rejette H_0 et on valide donc H_1

Conclusion

Au seuil de signification 5%, on a une évidence statistique permettant de croire que le nouveau procédé améliore la durée de vie moyenne des pneus.

Estimation de la p-value

p-value $= \mathbb{P}[T > T_{\text{obs}}] = \mathbb{P}[T_{15} > 2.29] =?$

Tout ce que l'on peut dire avec la table de la loi de Student est que la p-value est entre 0.010 et 0.025.

Exemple 3

On prend 9 mesures du PH d'une solution avec une méthode d'analyse produisant des mesures normalement distribuées, on obtient un écart-type échantillonnal de 0.045. Cet échantillon permet-il de conclure, avec un seuil de signification de 1% que le réel écart-type de cette méthode d'analyse est inférieur à 0.10. Donnez également une estimation de la p-value du test.

Solution

Soit σ^2 la réelle variance de cette méthode d'analyse.

Hypothèses

$H_0 : \sigma^2 = 0.10^2$
$H_1 : \sigma^2 < 0.10^2$

On a donc un test unilatéral à gauche au seuil $\alpha = 1\%$

Collecte des données échantillonnales

On a $n = 9$, $s^2 = 0.045^2$

Suite de la solution

Distribution d'échantillonnage à utiliser

Puisque que l'on a une population mère normale, on a que $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$

Règle de décision

La valeur critique est $\chi^2_{8;0.99} = 1.65$.
Règle de décision : Rejet de H_0 si $\chi^2 < 1.65$, non rejet de H_0 sinon.

Valeur observée et décision

Valeur observée : $\chi^2_{\text{obs}} = \frac{(n-1)S^2}{\sigma_0^2} = \frac{8(0.045)^2}{(0.10)^2} = 1.62$.

Décison : Puisque $\chi^2_{\text{obs}} = 1.62 < 1.65$, on rejette H_0 et on donc on peut valider H_1.

Conclusion

Au seuil de signification 1%, on a une évidence statistique permettant de croire que la méthode d'analyse possède un réel écart-type inférieur à 0.10.

Estimation de la p-value

p-value $= \mathbb{P}\left[\chi^2 < \chi^2_{\text{obs}}\right] = \mathbb{P}\left[\chi^2_8 < 1.62\right] = ?$

Tout ce que l'on peut dire avec la table de la loi du khi-deux est que la p-value est entre 0.005 et 0.01.

Exemple 4

Un premier échantillon de taille 15 issu d'une première population mère normale donne une moyenne de 0.8 avec un écart-type de 0.12.
Un deuxième échantillon de taille 16 issu d'une deuxième population mère normale donne une moyenne de 0.5 avec un écart-type de 0.11.
On suppose que les deux populations mères sont de même écart-type.
Au seuil de signification 5%, déterminez si les moyennes des deux populations mères sont égales ou non ?
Déterminez également la p-value approximative du test.

Solution

Soit μ_1 la réelle moyenne de la première population.

Soit μ_2 la réelle moyenne de la deuxième population.

Hypothèses

$H_0 : \mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$
$H_1 : \mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$

On a donc un test bilatéral au seuil $\alpha = 5\%$

Collecte des données échantillonnales

Pour l'échantillon 1 : $n_1 = 15$, $\overline{x}_1 = 0.8$ km et $s_1 = 0.12$ km

Pour l'échantillon 2 : $n_2 = 16$, $\overline{x}_2 = 0.5$ km et $s_2 = 0.11$ km

Distribution d'échantillonnage à utiliser

Puisque que les populations mères sont normales avec et σ_1 et σ_2 inconnu (mais $\sigma_1 = \sigma_2$), on a que
$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{s_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$

Règle de décision

La valeur critique est $t_{29;0.025} = 2.045$.

Règle de décision : Rejet de H_0 si $|T| > 2.045$, non rejet de H_0 sinon.

Valeur observée et décision

Valeur observée :

$$T_{\text{obs}} = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{s_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(0.8 - 0.5) - (0)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{(0.8 - 0.5) - (0)}{\sqrt{\frac{(14)0.12^2 + (15)0.11^2}{29}} \sqrt{\frac{1}{15} + \frac{1}{16}}} = \frac{0.3}{0.041307822} = 7.26$$

Solution (suite)

Décison : Puisque $|T_{\text{obs}}| = 7.26 > 2.045$, on rejette H_0 et on valide donc H_1

<u>Conclusion</u>
Au seuil de signification 5%, Une évidence statistique permet de croire que la population 1 et la population 2 n'ont pas la même moyenne.

<u>Estimation de la p-value</u>
p-value $= \mathbb{P}\left[|T| > |T_{\text{obs}}|\right] = \mathbb{P}\left[|T_{29}| > 7.26\right] = ?$

Tout ce que l'on peut dire avec la table de la loi de Student est que la p-value est plus petite que $2 \cdot (0.0005) = 0.001$. Bref, l'évidence statistique contre H_0 est ici très grande puisque la p-value est très petite.

Types d'erreur dans un test d'hypotèse et fonction puissance

Dans un test d'hypothèse, on peut commettre deux types d'erreur, rejeter H_0 alors que H_0 est en réalité vraie ou bien ne pas rejeter H_0 alors que H_0 était en réalité fausse. Dans le premier cas, on dit que l'on commet une erreur de première espèce, dans le deuxième cas, on dit que l'on commet une erreur de deuxième espèce. Le tableau suivant résume la situation.

État de H_0 (réalité)	Décision : ne pas rejeter H_0	Décision : rejeter H_0
H_0 est vraie	Bonne décision $(1-\alpha)$	Mauvaise décision (α)
H_0 est fausse	Mauvaise décision (β)	Bonne décision $(1-\beta)$

α est l'erreur de première espèce et β l'erreur de deuxième espèce.

Plus formellement, on a que

$$\alpha = \mathbb{P}\left[\text{Rejeter } H_0 | H_0 \text{ est vraie}\right] = \mathbb{P}_{H_0}\left[\text{Rejeter } H_0\right]$$

$$\beta = \mathbb{P}\left[\text{Ne pas rejeter } H_0 | H_0 \text{ est fausse}\right]$$

Remarques :

- α est ce que l'on a défini comme le seuil de signification précédemment.
- Pour calculer β, il faut savoir de quelle façon l'hypothèse H_0 $(\theta = \theta_0)$ est fausse, c'est-à-dire savoir la vraie valeur de θ.

La **fonction puissance** d'un test d'hypothèse est notée $\Pi(\theta)$ et est définie par $\Pi(\theta) = \mathbb{P}_\theta\left[\text{on rejette } H_0\right]$. En particulier, on a $\Pi(\theta_0) = \alpha$. Pour $\theta \neq \theta_0$, on a que $\Pi(\theta) = 1 - \beta$.

Principe de dualité entre un test bilatéral et un intervalle de confiance

Disons que l'on fait un IC sur θ au niveau de confiance $1-\alpha$ et que l'on obtient (u, v) comme intervalle de confiance.
Alors la règle de décision «Rejet de H_0 si $\theta_0 \notin (u, v)$» est une règle de décision pour le test bilatéral $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ au seuil de signification α.

Par exemple, dans le cas d'une population mère normale et σ connu, l'I.C. pour μ au niveau de confiance $1-\alpha$ est $\left[\overline{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} ; \overline{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$.
On a donc

$$\mathbb{P}\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha \Longleftrightarrow \mathbb{P}\left[-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu - \overline{X} \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$$

$$\Longleftrightarrow \mathbb{P}\left[-z_{\alpha/2} \leq \frac{\mu - \overline{X}}{\frac{\sigma}{\sqrt{n}}} \leq z_{\alpha/2}\right] = 1-\alpha \Longleftrightarrow \mathbb{P}\left[z_{\alpha/2} \geq \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \geq -z_{\alpha/2}\right] = 1-\alpha \Longleftrightarrow \mathbb{P}\left[-z_{\alpha/2} \leq \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\alpha/2}\right] = 1-\alpha$$

$$\Longleftrightarrow \mathbb{P}\left[\left|\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right| \leq z_{\alpha/2}\right] = 1-\alpha \Longleftrightarrow \mathbb{P}\left[\left|\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right| > z_{\alpha/2}\right] = 1-(1-\alpha) \Longleftrightarrow \mathbb{P}\left[|Z| > z_{\alpha/2}\right] = \alpha$$

On reconnaît la règle $|Z| > z_{\alpha/2}$ qui est la règle de décision au seuil α pour le rejet de $H_0 : \mu = \mu_0$ lorsque l'hypothèse alternative est $H_1 : \mu \neq \mu_0$.

Par exemple, si l'on mène un test bilatéral avec les hypothèses :

$H_0 : \mu = 30$
$H_1 : \mu \neq 30$

On mène le test, au seuil α, à l'aide d'un certain échantillon.

Si l'on construit un I.C. pour μ et que l'on obtient : $[30{,}21; 30{,}87]$
On peut baser notre règle de décision sur cet I.C. : Rejet de H_0 si $\mu_0 = 30 \notin [30{,}21; 30{,}87]$.
Ici, on a effectivement que $30 \notin [30{,}21; 30{,}87]$, ainsi on rejette H_0 et on valide H_1.

Bibliographie

- Bélisle, Claude. *STT-2920 Probabilités pour ingénieurs*, notes de cours (chapitre 13), Université Laval, 2024.
- ROSS, Sheldon M., *Initiation aux probabilités, traduction de la 4e édition*. Presses polytechniques et universitaires romandes, 1994.