Chapitre IV : Tests d'hypothèses

MATH-F-207

Davy Paindaveine

Bloc 2, Bachelier en sciences mathématiques
Université libre de Bruxelles

2023–2024

Contenu du chapitre

Terminologie et concepts de base

Le lemme fondamental de Neyman–Pearson

Tests unilatéraux à puissance uniformément maximale

Tests bilatéraux

Tests de rapport de vraisemblance

Tests χ^2

Contenu du chapitre

Terminologie et concepts de base

Le lemme fondamental de Neyman–Pearson

Tests unilatéraux à puissance uniformément maximale

Tests bilatéraux

Tests de rapport de vraisemblance

Tests χ^2

Introduction

Soit le modèle statistique paramétrique

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)} = \left\{ \mathrm{P}_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^k \right\}\right),$$

où les mesures de probabilité $\mathrm{P}_\theta^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \ldots, X_n)$.

Le chapitre précédent visait à faire un pari sur la valeur exacte de $\theta \in \Theta$.
Mais certaines applications demandent moins que ceci.

Exemple électoral : parier sur le fait que le candidat A va remporter l'élection ne requiert pas de parier sur la valeur exacte de p.

Partition de l'espace paramétrique

Dans ce chapitre, on veut décider, sur la base de $X^{(n)}$, si $\theta \in \Theta_0$ ou si $\theta \in \Theta_1$, où Θ_0 et Θ_1 forment une partition de l'espace paramétrique Θ.

Exemple électoral : on souhaite décider si $p \in [0, \frac{1}{2}]$ ou si $p \in \,]\frac{1}{2}, 1]$.

Terminologie

La terminologie est très structurée : le problème de test d'hypothèses s'écrit

$$\left\{\begin{array}{l} \mathcal{H}_0 : \theta \in \Theta_0 \\ \mathcal{H}_1 : \theta \in \Theta_1. \end{array}\right.$$

$\mathcal{H}_0$ est l'hypothèse nulle et $\mathcal{H}_1$ est la contre-hypothèse.

Deux décisions seulement sont possibles :

$\not R\mathcal{H}_0$: ne pas rejeter l'hypothèse nulle $\mathcal{H}_0$

$R\mathcal{H}_0$: rejeter l'hypothèse nulle $\mathcal{H}_0$.

Asymétrie des hypothèses

- Attention : $\not R\mathcal{H}_0$ ne signifie pas accepter $\mathcal{H}_0$...
- ...Mais doit plutôt être interprété comme "il n'y a pas suffisamment d'évidence statistique pour conclure au rejet de $\mathcal{H}_0$".
- Donc une procédure de test ne sera vraiment conclusive que si on prend la décision $R\mathcal{H}_0$.
- C'est la raison pour laquelle il convient de prendre pour $\mathcal{H}_1$ la proposition qu'on veut "prouver" statistiquement.

Exemple du verre de bière

Le barman de la Jefke promet que la quantité de bière moyenne servie dans un verre est supérieure ou égale à 15 cl.

Dans le modèle d'échantillonnage paramétrique gaussien, une personne qui en douterait considérera le problème de test

$$\begin{cases} \mathcal{H}_0 : \theta = (\mu, \sigma^2) \in \Theta_0 = \{(\mu, \sigma^2) \in \Theta : \mu \geq 15\} \\ \mathcal{H}_1 : \theta = (\mu, \sigma^2) \in \Theta_1 = \{(\mu, \sigma^2) \in \Theta : \mu < 15\}. \end{cases}$$

De façon plus informelle, on écrira

$$\begin{cases} \mathcal{H}_0 : \mu \geq 15 \\ \mathcal{H}_1 : \mu < 15. \end{cases}$$

Test unilatéral

Considérons le cas où $\Theta \subset \mathbb{R}$.

Les problèmes de test

$$\begin{cases} \mathcal{H}_0 : \theta \leq \theta_0 \\ \mathcal{H}_1 : \theta > \theta_0 \end{cases} \quad \text{ou} \quad \begin{cases} \mathcal{H}_0 : \theta \geq \theta_0 \\ \mathcal{H}_1 : \theta < \theta_0, \end{cases}$$

où $\theta_0 \in \Theta$ est fixé, seront dits unilatéraux.

Exemple électoral : pour parier sur le gagnant de l'élection, on considérera

$$\begin{cases} \mathcal{H}_0 : p \leq \frac{1}{2} \\ \mathcal{H}_1 : p > \frac{1}{2} \end{cases} \quad \text{ou} \quad \begin{cases} \mathcal{H}_0 : p \geq \frac{1}{2} \\ \mathcal{H}_1 : p < \frac{1}{2}. \end{cases}$$

Test bilatéral

Le problème bilatéral correspondant est

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

Exemple : si on souhaite tester qu'une pièce de monnaie est équilibrée, on se place dans un modèle d'échantillonnage de Bernoulli, et on considère

$$\begin{cases} \mathcal{H}_0 : p = \frac{1}{2} \\ \mathcal{H}_1 : p \neq \frac{1}{2}. \end{cases}$$

Remarques

- Dans le cas d'un modèle statistique non paramétrique, c'est la collection de lois $\mathcal{P}^{(n)} = \{\mathrm{P}^{(n)}\}$ qu'on partitionnera pour définir $\mathcal{H}_0$ et $\mathcal{H}_1$.
- $\mathcal{H}_0$ est simple $\Leftrightarrow$ $\mathcal{H}_0$ est associée à une unique mesure de probabilité du modèle statistique considéré (dans le cas paramétrique, $\Theta_0 = \{\theta_0\}$).
- Dans le cas contraire, $\mathcal{H}_0$ est composée.
- De la même façon, on parle de $\mathcal{H}_1$ simple ou composée.

Tests purs

La procédure statistique qui, sur base de $x^{(n)} = (x_1, \ldots, x_n)$, prend la décision de rejeter ou non $\mathcal{H}_0$ est appelée un test.

Définition 1

Un test pur ϕ est une fonction mesurable de $\mathcal{X}^{(n)}$ dans $\{0, 1\}$, où $\mathcal{X}^{(n)}$ désigne l'ensemble des valeurs possibles pour $x^{(n)}$.

Lecture:
Si le $x^{(n)}$ observé est tel que $\phi(x^{(n)}) = 0$, alors le test prend la décision $\not{R}\mathcal{H}_0$.
Si le $x^{(n)}$ observé est tel que $\phi(x^{(n)}) = 1$, alors le test prend la décision $R\mathcal{H}_0$.

L'ensemble $\phi^{-1}(\{1\})$ est appelé la zone critique du test.

Exemple électoral

Pour le problème de test

$$\begin{cases} \mathcal{H}_0 : p \leq \frac{1}{2} \\ \mathcal{H}_1 : p > \frac{1}{2}, \end{cases}$$

des tests raisonnables sont, pour $c \in \mathbb{N}$,

$$\phi_c(x^{(n)}) = \mathbb{I}\left[\sum_{i=1}^n x_i \geq c\right].$$

Question : que font les tests extrêmes $\phi_0(\cdot)$ et $\phi_{n+1}(\cdot)$?

Tests randomisés

Définition 2

Un test randomisé ϕ est une fonction mesurable de $\mathcal{X}^{(n)}$ dans $[0, 1]$.

Lecture:
Pour $x^{(n)} \in \mathcal{X}^{(n)}$, le test ϕ prend la décision $R\mathcal{H}_0$ avec probabilité $\phi(x^{(n)})$, sur la base d'un aléa indépendant du modèle statistique considéré.

(Si $\phi(x^{(n)}) \in \{0, 1\}$, alors la décision est compatible avec celle d'un test pur).

L'ensemble $\phi^{-1}(\{1\})$ est encore appelé la zone critique
L'ensemble $\phi^{-1}(]0, 1[)$ est appelé la zone de randomisation.

Probabilité de rejet en termes d'espérance

La formule des probabilités totales livre

$$\mathrm{P}_\theta^{(n)}[\phi \text{ prend la décision } R\mathcal{H}_0]$$

$$= \int_{\mathcal{X}^{(n)}} \underbrace{\mathrm{P}_\theta^{(n)}[\phi \text{ prend la décision } R\mathcal{H}_0 \mid X^{(n)} = x^{(n)}]}_{\phi(x^{(n)})} \, L_\theta^{(n)}(x^{(n)}) \, dx^{(n)}$$

$$= \mathrm{E}_\theta^{(n)}[\phi(X^{(n)})] \stackrel{\text{not}}{=} \mathrm{E}_\theta^{(n)}[\phi],$$

où $L_\theta^{(n)}(x^{(n)})$ désigne la vraisemblance associée à $\mathrm{P}_\theta^{(n)}$.

Types d'erreurs

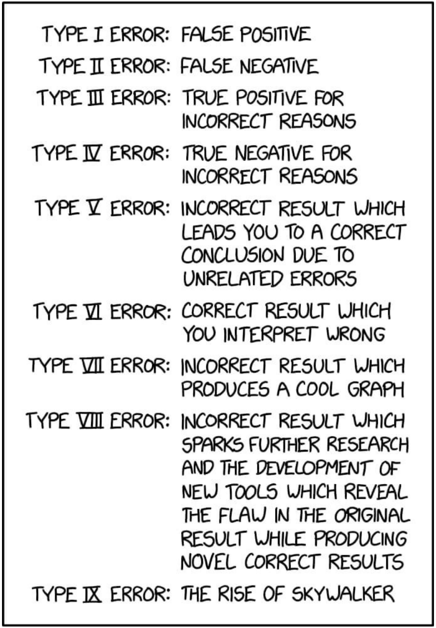
Deux erreurs sont possibles lorsqu'on conduit un test:

- l'erreur de 1ère espèce (prendre la décision $R\mathcal{H}_0$ quand $\mathcal{H}_0$ est vraie)
- l'erreur de 2nde espèce (prendre la décision $\not R\mathcal{H}_0$ quand $\mathcal{H}_0$ est fausse)

	$\not R\mathcal{H}_0$	$R\mathcal{H}_0$
$\mathcal{H}_0$ est vraie	OK!	Erreur de 1ère espèce
$\mathcal{H}_0$ est fausse	Erreur de 2nde espèce	OK!

Erreur de 1ère espèce $\neq$ Erreur de 2nde espèce

$\mathcal{H}_0$: pas enceinte vs $\mathcal{H}_1$: enceinte


TYPE I ERROR: FALSE POSITIVE
TYPE II ERROR: FALSE NEGATIVE
TYPE III ERROR: TRUE POSITIVE FOR INCORRECT REASONS
TYPE IV ERROR: TRUE NEGATIVE FOR INCORRECT REASONS
TYPE V ERROR: INCORRECT RESULT WHICH LEADS YOU TO A CORRECT CONCLUSION DUE TO UNRELATED ERRORS
TYPE VI ERROR: CORRECT RESULT WHICH YOU INTERPRET WRONG
TYPE VII ERROR: INCORRECT RESULT WHICH PRODUCES A COOL GRAPH
TYPE VIII ERROR: INCORRECT RESULT WHICH SPARKS FURTHER RESEARCH AND THE DEVELOPMENT OF NEW TOOLS WHICH REVEAL THE FLAW IN THE ORIGINAL RESULT WHILE PRODUCING NOVEL CORRECT RESULTS
TYPE IX ERROR: THE RISE OF SKYWALKER

Risques

La qualité d'un test ϕ sera mesurée par les probabilités respectives des erreurs (la décision prise par ϕ est fondée sur $X^{(n)}$, donc est aléatoire).

Définition 3

Soit ϕ un test pour le problème de test $\mathcal{H}_0 : \theta \in \Theta_0$ contre $\mathcal{H}_1 : \theta \in \Theta_1$. Alors,
(i) le risque de 1ère espèce de ϕ en $\theta(\in \Theta_0)$ est

$$\mathrm{P}_\theta^{(n)}[\phi \textit{ prend la décision } R\mathcal{H}_0] = \mathrm{E}_\theta^{(n)}[\phi].$$

(ii) Le risque de 2nde espèce de ϕ en $\theta(\in \Theta_1)$ est

$$\mathrm{P}_\theta^{(n)}[\phi \textit{ prend la décision } \not{R}\mathcal{H}_0] = 1 - \mathrm{E}_\theta^{(n)}[\phi].$$

Dimension et puissance d'un test

Définition 4

Soit ϕ un test pour le problème de test $\mathcal{H}_0 : \theta \in \Theta_0$ contre $\mathcal{H}_1 : \theta \in \Theta_1$. Alors la dimension de ϕ en $\theta (\in \Theta)$ est

$$\mathrm{P}_\theta^{(n)}[\phi \textit{ prend la décision } R\mathcal{H}_0] = \mathrm{E}_\theta^{(n)}[\phi].$$

Si $\theta \in \Theta_1$, on parlera de puissance de ϕ en θ.

Un bon test a de faibles risques de $1^{\text{ère}}$ et 2^{nde} espèces, ou encore, un faible risque de $1^{\text{ère}}$ espèce et une grande puissance.

Mais minimiser le risque de $1^{\text{ère}}$ espèce et maximiser la puissance sont des objectifs antagonistes…

Principe de Neyman

Le principe de Neyman consiste alors à maximiser la puissance[1] sous la contrainte que le risque de $1^{\text{ère}}$ espèce vaut au plus α, où $\alpha \in]0,1[$ est fixé.

Plus précisément, on veut maximiser la puissance (uniformément en $\theta(\in \Theta_1)$) dans la classe des tests *de niveau* α.

Définition 5

Soit ϕ un test pour le problème de test $\mathcal{H}_0 : \theta \in \Theta_0$ contre $\mathcal{H}_1 : \theta \in \Theta_1$. Alors ϕ est de niveau $\alpha(\in]0,1[)$ $\Leftrightarrow \mathrm{E}_\theta^{(n)}[\phi] \leq \alpha$ pour tout $\theta \in \Theta_0$.

Les *niveaux* les plus usuels sont $\alpha = 0.05$ et $\alpha = 0.01$.

[1] De manière équivalente, à minimiser le risque de 2^{nde} espèce

Tests à puissance uniformément maximale

Soit $\mathcal{C}_\alpha$ la classe des tests de niveau α.

Définition 6

ϕ_* *est à* *puissance uniformément maximale (PUM) dans la classe* $\mathcal{C}_\alpha$
$\Leftrightarrow$ *(i)* $\phi_* \in \mathcal{C}_\alpha$, *et*
(ii) pour tout $\phi \in \mathcal{C}_\alpha$, $\mathrm{E}_\theta^{(n)}[\phi_*] \geq \mathrm{E}_\theta^{(n)}[\phi]$ *pour tout* $\theta \in \Theta_1$.

De tels tests existent-ils?
Pour quels problèmes de test?

Contenu du chapitre

Cadre

Soit le modèle statistique paramétrique

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)} = \left\{ \mathrm{P}_\theta^{(n)} : \theta \in \{\theta_0, \theta_1\} \subset \mathbb{R}^k \right\}\right),$$

où les mesures de probabilité $\mathrm{P}_\theta^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \ldots, X_n)$.

Le problème de test considéré ici est

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

La classe $\mathcal{C}_\alpha$ des tests de niveau α est donc celle des tests ϕ tels que $\mathrm{E}_{\theta_0}^{(n)}[\phi] \leq \alpha$.

Le lemme fondamental

Théorème 7 (Lemme fondamental de Neyman–Pearson)

Considérons le problème de test associé à deux hypothèses simples et fixons $\alpha \in]0,1[$. *Désignons par* $L_\theta^{(n)}(x^{(n)})$ *la vraisemblance associée à* $\mathrm{P}_\theta^{(n)}$. *Alors,*

(i) il existe $k_\alpha \in \mathbb{R}^+$ *et* $\gamma_\alpha \in [0,1]$ *tels que le test* ϕ_* *défini par*

$$\phi_*(x^{(n)}) = \begin{cases} 1 & \textit{si} \quad L_{\theta_1}^{(n)}(x^{(n)}) > k_\alpha L_{\theta_0}^{(n)}(x^{(n)}) \\ \gamma_\alpha & \textit{si} \quad L_{\theta_1}^{(n)}(x^{(n)}) = k_\alpha L_{\theta_0}^{(n)}(x^{(n)}) \\ 0 & \textit{si} \quad L_{\theta_1}^{(n)}(x^{(n)}) < k_\alpha L_{\theta_0}^{(n)}(x^{(n)}) \end{cases}$$

vérifie $\mathrm{E}_{\theta_0}^{(n)}[\phi_*] = \alpha$;

(ii) ϕ_* *est à puissance (uniformément) maximale dans* $\mathcal{C}_\alpha$.

Exemple

Pour $\theta_1 > \theta_0 > 1$, considérons le problème de test

$$\begin{cases} \mathcal{H}_0 : \; X_1 \sim f_{\theta_0}(x) = \theta_0 x^{\theta_0 - 1} \mathbb{I}[0 < x < 1] \\ \mathcal{H}_1 : \; X_1 \sim f_{\theta_1}(x) = \theta_1 x^{\theta_1 - 1} \mathbb{I}[0 < x < 1] \end{cases}$$

au niveau $\alpha (\in]0, 1[)$. Alors le test ϕ_* donné par

$$\phi_*(x) = \begin{cases} 1 & \text{si} \quad x > (1-\alpha)^{1/\theta_0} \\ 0 & \text{si} \quad x \leq (1-\alpha)^{1/\theta_0} \end{cases}$$

est à puissance maximale pour ce problème.

Unicité du test de Neyman–Pearson

Théorème 8

Considérons le cadre du lemme de Neyman–Pearson.

(i) Si ϕ est de niveau α et est aussi puissant que ϕ_*, alors il existe une fonction $\gamma_\alpha(\cdot) : \mathcal{X}^{(n)} \to [0,1]$ telle que, pour presque tout $x^{(n)} \in \mathcal{X}^{(n)}$,

$$\phi(x^{(n)}) = \begin{cases} 1 & \textit{si} \quad L_{\theta_1}^{(n)}(x^{(n)}) > k_\alpha L_{\theta_0}^{(n)}(x^{(n)}) \\ \gamma_\alpha(x^{(n)}) & \textit{si} \quad L_{\theta_1}^{(n)}(x^{(n)}) = k_\alpha L_{\theta_0}^{(n)}(x^{(n)}) \\ 0 & \textit{si} \quad L_{\theta_1}^{(n)}(x^{(n)}) < k_\alpha L_{\theta_0}^{(n)}(x^{(n)}), \end{cases}$$

où k_α est la constante intervenant dans ϕ_*;

(ii) $\mathrm{E}_{\theta_1}^{(n)}[\phi_*] > \alpha$.

Exemple

Dans l'exemple précédent, on vérifiera directement que

$$\begin{aligned} E_{\theta_1}[\phi_*] &= 1 - (1-\alpha)^{\theta_1/\theta_0} \\ &> \alpha. \end{aligned}$$

Contenu du chapitre

Tests unilatéraux

Soit un modèle statistique paramétrique à paramètre scalaire de la forme

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)} = \left\{ \mathrm{P}_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R} \right\}\right),$$

où les mesures de probabilité $\mathrm{P}_\theta^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \ldots, X_n)$.

Soit un problème de test unilatéral de la forme

$$\begin{cases} \mathcal{H}_0 : \theta \leq \theta_0 \\ \mathcal{H}_1 : \theta > \theta_0, \end{cases}$$

où θ_0 est une valeur fixée dans l'intérieur de Θ.

Existe-t-il un test à PUM dans la classe des tests de niveau α?

Modèle à rapport de vraisemblance monotone

Définition 9

Le modèle statistique ci-dessus est à rapport de vraisemblance monotone en la statistique $T(x^{(n)}) \Leftrightarrow$ *pour tout* θ' *et* $\theta'' \in \Theta$ *avec* $\theta' < \theta''$,

$$\frac{L_{\theta''}^{(n)}(x^{(n)})}{L_{\theta'}^{(n)}(x^{(n)})} = h_{\theta',\theta''}(T(x^{(n)}))$$

pour une fonction $h_{\theta',\theta''}(\cdot)$ *strictement croissante.*

Exemple électoral

Soit $X^{(n)} = (X_1, \ldots, X_n)$, où les observations X_i sont i.i.d. Bern(p) et où le paramètre $\theta = p \in \Theta = [0, 1] \subset \mathbb{R}$. Alors, pour tout p', p'' avec $p' < p''$,

$$\begin{aligned} \frac{L^{(n)}_{p''}(x^{(n)})}{L^{(n)}_{p'}(x^{(n)})} &= \frac{p''^{\sum_{i=1}^n x_i}(1-p'')^{n-\sum_{i=1}^n x_i}}{p'^{\sum_{i=1}^n x_i}(1-p')^{n-\sum_{i=1}^n x_i}} \\ &= \left(\frac{p''}{p'}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p''}{1-p'}\right)^{n-\sum_{i=1}^n x_i} \\ &= \left(\frac{p''}{p'}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p'}{1-p''}\right)^{(\sum_{i=1}^n x_i)-n} \end{aligned}$$

est une fonction strictement croissante de $T(x^{(n)}) = \sum_{i=1}^n x_i$.

Exemple du verre de bière (variance connue)

Soit $X^{(n)} = (X_1, \ldots, X_n)$, où les X_i sont i.i.d. $\mathcal{N}(\mu, \sigma_0^2)$, avec $\sigma_0^2 > 0$ fixé, de sorte que le paramètre est $\theta = \mu \in \Theta = \mathbb{R}$. Alors, pour tout μ', μ'' avec $\mu' < \mu''$,

$$
\begin{aligned}
\frac{L^{(n)}_{\mu''}(x^{(n)})}{L^{(n)}_{\mu'}(x^{(n)})} &= \frac{\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i - \mu'')^2\right)}{\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i - \mu')^2\right)} \\
&= \frac{\exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2 + \frac{\mu''}{\sigma_0^2}\sum_{i=1}^n x_i - \frac{n(\mu'')^2}{2\sigma_0^2}\right)}{\exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2 + \frac{\mu'}{\sigma_0^2}\sum_{i=1}^n x_i - \frac{n(\mu')^2}{2\sigma_0^2}\right)} \\
&= \exp\left(\frac{n}{2\sigma_0^2}((\mu')^2 - (\mu'')^2)\right) \exp\left(\frac{\mu''-\mu'}{\sigma_0^2}\sum_{i=1}^n x_i\right)
\end{aligned}
$$

est une fonction strictement croissante de $T(x^{(n)}) = \sum_{i=1}^n x_i$ ou, de manière équivalente, de $T(x^{(n)}) = \frac{1}{n}\sum_{i=1}^n x_i = \bar{x}$.

Test à PUM dans un modèle à RV monotone

Théorème 10 (Théorème de Lehmann)

Soit un modèle statistique paramétrique à paramètre scalaire et à rapport de vraisemblance monotone en la statistique $T(x^{(n)})$*. Considérons le problème*

$$\begin{cases} \mathcal{H}_0 : \theta \leq \theta_0 \\ \mathcal{H}_1 : \theta > \theta_0 \end{cases}$$

au niveau $\alpha \in]0,1[$ *(où* $\theta_0 \in \mathrm{int}(\Theta)$ *est fixé). Alors,*

(i) il existe $t_\alpha^+ \in \mathbb{R}$ *et* $\gamma_\alpha \in [0,1]$ *tels que le test* ϕ_{*T}^+ *défini par*

$$\phi_{*T}^+(x^{(n)}) = \begin{cases} 1 & \textit{si} \quad T(x^{(n)}) > t_\alpha^+ \\ \gamma_\alpha & \textit{si} \quad T(x^{(n)}) = t_\alpha^+ \\ 0 & \textit{si} \quad T(x^{(n)}) < t_\alpha^+ \end{cases}$$

vérifie $\mathrm{E}_{\theta_0}^{(n)}[\phi_{*T}^+] = \alpha$*;*

(ii) $\theta \mapsto \mathrm{E}_\theta^{(n)}[\phi_{*T}^+]$ *est strictement croissante;*

(iii) ϕ_{*T}^+ *est à PUM dans la classe des tests de niveau* α.

Exemple du verre de bière (avec variance fixée)

Soit $X^{(n)} = (X_1, \dots, X_n)$, où les X_i sont i.i.d. $\mathcal{N}(\mu, \sigma_0^2)$, avec $\sigma_0^2 > 0$ fixé. Soit $\mu_0 \in \mathbb{R}$. (i) Le test à PUM au niveau $\alpha \in]0,1[$ pour $\mathcal{H}_0 : \mu \leq \mu_0$ contre $\mathcal{H}_1 : \mu > \mu_0$ est

$$\phi_{*T}^{+}(x^{(n)}) = \begin{cases} 1 & \text{si} \quad \bar{x} > \mu_0 + \frac{\sigma_0}{\sqrt{n}} \Phi^{-1}(1-\alpha) \\ 0 & \text{si} \quad \bar{x} \leq \mu_0 + \frac{\sigma_0}{\sqrt{n}} \Phi^{-1}(1-\alpha). \end{cases}$$

Puissance du test à PUM

(ii) La dimension du test sous $\mathrm{P}_\mu^{(n)}$ est

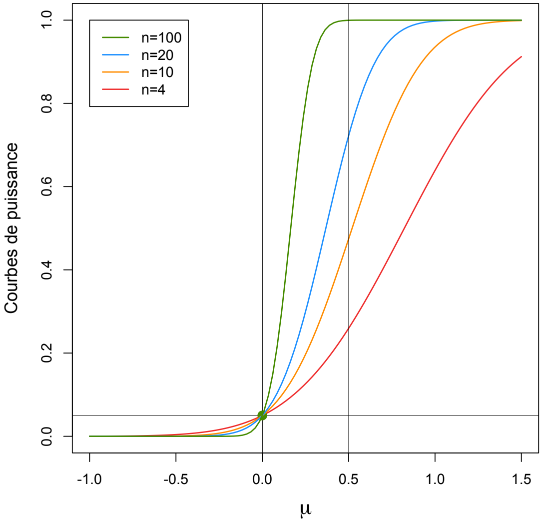
$$\mathrm{E}_\mu^{(n)}[\phi_{*T}^+] = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{\sqrt{n}(\mu-\mu_0)}{\sigma_0}\right),$$

où Φ désigne la fonction de répartition de la loi normale standard.

Remarques (voir aussi les illustrations dans R):

1. La "fonction de puissance" $\mu \mapsto \mathrm{E}_\mu^{(n)}[\phi_{*T}^+]$ est strictement croissante.
   ($\rightsquigarrow$ puisque $\mathrm{E}_{\mu_0}^{(n)}[\phi_{*T}^+] = \alpha$, le test est bien de niveau α).
2. Pour $\mu(>\mu_0)$ et α fixés, la puissance $\mathrm{E}_\mu^{(n)}[\phi_{*T}^+] \to 1$ quand $n \to \infty$.
   On parlera de test convergent.
3. La puissance est d'autant plus grande que σ_0 est petit.

Puissance du test à PUM ($\mu_0 = 0$)



Test à PUM pour l'autre problème unilatéral

Théorème 11 (exercice)

Soit un modèle statistique paramétrique à paramètre scalaire et à rapport de vraisemblance monotone en la statistique $T(x^{(n)})$. Considérons le problème

$$\left\{\begin{array}{l}\mathcal{H}_0 : \theta \geq \theta_0 \\ \mathcal{H}_1 : \theta < \theta_0\end{array}\right.$$

au niveau $\alpha \in]0,1[$ (où $\theta_0 \in \mathrm{int}(\Theta)$ est fixé). Alors,

(i) il existe $t_\alpha^- \in \mathbb{R}$ et $\gamma_\alpha \in [0,1]$ tels que le test ϕ_{*T}^- défini par

$$\phi_{*T}^-(x^{(n)}) = \left\{\begin{array}{ccc} 1 & \textit{si} & T(x^{(n)}) < t_\alpha^- \\ \gamma_\alpha & \textit{si} & T(x^{(n)}) = t_\alpha^- \\ 0 & \textit{si} & T(x^{(n)}) > t_\alpha^- \end{array}\right.$$

vérifie $\mathrm{E}_{\theta_0}^{(n)}[\phi_{*T}^-] = \alpha$;

(ii) $\theta \mapsto \mathrm{E}_\theta^{(n)}[\phi_{*T}^-]$ est strictement décroissante;

(iii) ϕ_{*T}^- est à PUM dans la classe des tests de niveau α.

Test à PUM pour l'autre problème unilatéral

Exercice : déterminer ϕ^{-}_{*T} dans le modèle d'échantillonnage gaussien à variance fixée et calculer sa fonction de puissance.

Contenu du chapitre

Introduction

Soit un modèle statistique paramétrique à paramètre scalaire

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)} = \left\{\mathrm{P}_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}\right\}\right)$$

où les mesures de probabilité $\mathrm{P}_\theta^{(n)}$ représentent les distributions possibles pour $X^{(n)} = (X_1, \ldots, X_n)$.

Le problème de test bilatéral

$$\begin{cases} \mathcal{H}_0 : \theta = \theta_0 \\ \mathcal{H}_1 : \theta \neq \theta_0 \end{cases}$$

admet-il un test à PUM au niveau α?

Un résultat négatif

Théorème 12

Soit un modèle statistique paramétrique à paramètre scalaire et à rapport de vraisemblance monotone en la statistique $T(x^{(n)})$. *Quel que soit* $\alpha \in]0,1[$, *il n'existe pas de test à PUM au niveau* α *pour le problème de test bilatéral.*

Tests sans biais

Pour obtenir un résultat d'existence, il est nécessaire de se restreindre à une plus petite classe de tests.

Définition 13

Le test ϕ de niveau α pour le problème $\mathcal{H}_0 : \theta \in \Theta_0$ contre $\mathcal{H}_1 : \theta \in \Theta_1$ est sans biais $\Leftrightarrow \mathrm{E}_\theta^{(n)}[\phi] \geq \alpha$ pour tout $\theta \in \Theta_1$.

Définition 14

Le test ϕ_* est à PUM dans la classe des tests sans biais au niveau α
$\Leftrightarrow$ (i) ϕ_* est sans biais au niveau α, et
(ii) pour tout ϕ sans biais au niveau α, on a $\mathrm{E}_\theta[\phi_*] \geq \mathrm{E}_\theta[\phi]$ pour tout $\theta \in \Theta_1$.

Familles exponentielles

Théorème 15 (prouvé en STAT-F404)

Soit un modèle statistique dont les vraisemblances sont de la forme

$$L_\theta^{(n)}(x^{(n)}) = c_\theta h(x^{(n)}) \exp(\theta\, T(x^{(n)})).$$

Soit $\alpha \in]0,1[$. *Alors (i) il existe* $t_{\alpha 1}, t_{\alpha 2} \in \mathbb{R}$ *et* $\gamma_{\alpha 1}, \gamma_{\alpha 2} \in [0,1]$ *tels que le test* ϕ_{*T} *défini par*

$$\phi_{*T}(x^{(n)}) = \begin{cases} 1 & \textit{si} \quad T(x^{(n)}) \notin [t_{\alpha 1}, t_{\alpha 2}] \\ \gamma_{\alpha 1} & \textit{si} \quad T(x^{(n)}) = t_{\alpha 1} \\ \gamma_{\alpha 2} & \textit{si} \quad T(x^{(n)}) = t_{\alpha 2} \\ 0 & \textit{si} \quad T(x^{(n)}) \in]t_{\alpha 1}, t_{\alpha 2}[\end{cases}$$

vérifie $\mathrm{E}_{\theta_0}[\phi_{*T}] = \alpha$ *et* $\mathrm{E}_{\theta_0}[\phi_{*T} T] = \alpha \mathrm{E}_{\theta_0}[T]$ *(où on a écrit* $T = T(X^{(n)})$*);* *(ii)* ϕ_{*T} *est à PUM dans la classe des tests sans biais au niveau* α *pour le problème de test bilatéral en* θ_0.

Le cas gaussien à variance fixée

Soit $X^{(n)} = (X_1, \ldots, X_n)$, où les X_i sont i.i.d. $\mathcal{N}(\mu, \sigma_0^2)$, avec $\sigma_0^2 > 0$ fixé. Puisque

$$L_\mu^{(n)}(x^{(n)}) = \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i - \mu)^2\right)$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \exp\left(-\frac{n\mu^2}{2\sigma_0^2}\right)}_{c_\mu} \times \underbrace{\exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2\right)}_{h(x^{(n)})} \times \underbrace{\exp\left(\frac{\mu}{\sigma_0^2}\sum_{i=1}^n x_i\right)}_{\exp(\mu T(x^{(n)}))},$$

la vraisemblance est de la forme mentionnée dans le théorème, avec

$$T(x^{(n)}) = \frac{1}{\sigma_0^2}\sum_{i=1}^n x_i.$$

Le cas gaussien à variance fixée

Exercice : montrer que le test ϕ_{*T} optimal alors est donné par

$$\phi_{*T}(x^{(n)}) = \begin{cases} 1 & \text{si} \quad |\bar{X} - \mu_0| > \frac{\sigma_0}{\sqrt{n}} \Phi^{-1}\big(1 - \frac{\alpha}{2}\big) \\ 0 & \text{si} \quad |\bar{X} - \mu_0| \leq \frac{\sigma_0}{\sqrt{n}} \Phi^{-1}\big(1 - \frac{\alpha}{2}\big) \end{cases}$$

et que la fonction de puissance de ϕ_{*T} est

$$\mu \mapsto \mathrm{E}_\mu^{(n)}[\phi_{*T}] = 1 - \Phi\Big(\Phi^{-1}\Big(1 - \frac{\alpha}{2}\Big) + \frac{\sqrt{n}}{\sigma_0}(\mu_0 - \mu)\Big)$$
$$+\Phi\Big(- \Phi^{-1}\Big(1 - \frac{\alpha}{2}\Big) + \frac{\sqrt{n}}{\sigma_0}(\mu_0 - \mu)\Big).$$

Remarques (voir aussi les illustrations dans R): 1. Le risque de 1ère espèce de ϕ_{*T} vaut α. 2. Sa fonction de puissance est symétrique par rapport à μ_0. 3. ϕ_{*T} est convergent: en tout $\mu \neq \mu_0$, la puissance $\mathrm{E}_\mu^{(n)}[\phi_{*T}] \to 1$ si $n \to \infty$.

Contenu du chapitre

Introduction

Soit un modèle statistique paramétrique

$$\left(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)} = \left\{\mathrm{P}_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^k\right\}\right),$$

où les mesures de probabilité $\mathrm{P}_\theta^{(n)}$ représentent les distributions possibles de $X^{(n)} = (X_1, \ldots, X_n)$. Notons $L_\theta^{(n)}(x^{(n)})$ la vraisemblance associée à $\mathrm{P}_\theta^{(n)}$.

Nous considérons le problème de test général

$$\begin{cases} \mathcal{H}_0 : \theta \in \Theta_0 \\ \mathcal{H}_1 : \theta \in \Theta_1, \end{cases}$$

où Θ_0 et Θ_1 forment une partition de Θ.

Tests de rapport de vraisemblance

Soit

$$\Lambda(x^{(n)}) = \frac{\sup_{\theta \in \Theta_0} L_\theta^{(n)}(x^{(n)})}{\sup_{\theta \in \Theta} L_\theta^{(n)}(x^{(n)})}.$$

Une petite valeur de $\Lambda(x^{(n)})$ indique que l'observation $x^{(n)}$ est beaucoup moins vraisemblable sous $\mathcal{H}_0$ que sous $\mathcal{H}_1$.

Le test de rapport de vraisemblance est alors de la forme

$$\phi_{\mathrm{RV}}(x^{(n)}) = \mathbb{I}\big[\Lambda(x^{(n)}) < c_\alpha\big],$$

où c_α est choisi de telle façon que

$$\sup_{\theta \in \Theta_0} \mathrm{E}_\theta^{(n)}[\phi_{\mathrm{RV}}] = \alpha.$$

Comportement asymptotique

S'il est en général difficile de déterminer c_α pour un n fixé, on peut construire un test asymptotique grâce au résultat suivant.

Théorème 16

Sous certaines conditions de régularité, on a que, pour tout $\theta \in \Theta_0$,

$$-2 \ln \Lambda(X^{(n)}) \xrightarrow{\mathcal{L}} \chi^2_{k-r}$$

sous $\mathrm{P}_\theta^{(n)}$ quand $n \to \infty$, où r est le nombre de paramètres fonctionnellement indépendants dans Θ_0.

En particulier, si Θ_0 est un espace affin, alors r est la dimension correspondante.

Test du rapport de vraisemblance asymptotique

Le test de rapport de vraisemblance asymptotique

$$\phi^{\text{as}}_{\text{RV}}(x^{(n)}) = \mathbb{I}\Big[-2\ln\Lambda(x^{(n)}) > \chi^2_{k-r,1-\alpha}\Big],$$

où $\chi^2_{k-r,1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ de la loi χ^2_{k-r}, vérifie

$$\lim_{n\to\infty} \sup_{\theta\in\Theta_0} \mathrm{E}^{(n)}_\theta[\phi_{\text{RV}}] \leq \alpha$$

(et est donc "asymptotiquement de niveau α").

Sous certaines conditions de régularité, ce test jouit de certaines propriétés d'optimalité asymptotiques (// MLE).

Contenu du chapitre

Variable aléatoire multinomiale

Soit E une expérience aléatoire dont les résultats possibles sont $A_1, \ldots, A_I$. Notons $p_1, \ldots, p_I$ les probabilités correspondantes.

Définition 17

La loi multinomiale de paramètres $n, p_1, \ldots, p_{I-1}$ est la loi de

$$N = (N_1, \ldots, N_I),$$

où N_i compte le nombre d'occurences du résultat A_i dans n répétitions indépendantes de E.

Notation: $N \sim \mathrm{Mult}(n, p_1, \ldots, p_{I-1})$.
Remarque : $p_I = 1 - \sum_{i=1}^{I-1} p_i$ n'est pas inclus dans les paramètres.

Moments

Si $N \sim \text{Mult}(n, p_1, \ldots, p_{I-1})$, alors $N_i \sim \text{Bin}(n, p_i)$ pour tout i. Donc

$$\text{E}[N_i] = np_i \quad \text{et} \quad \text{Var}[N_i] = np_i(1 - p_i).$$

Puisque $N_i + N_j \sim \text{Bin}(n, p_i + p_j)$ pour tout $i \neq j$,

$$\begin{aligned} n(p_i + p_j)(1 - (p_i + p_j)) &= \text{Var}[N_i + N_j] \\ &= \text{Var}[N_i] + \text{Var}[N_j] + 2\text{Cov}[N_i, N_j] \\ &= np_i(1 - p_i) + np_j(1 - p_j) + 2\text{Cov}[N_i, N_j], \end{aligned}$$

ce qui livre $\text{Cov}[N_i, N_j] = -np_ip_j < 0$ pour tout $i \neq j$.

Les N_i sont donc dépendants et négativement corrélés (c'est surprenant?)

Résultats asymptotiques

Théorème 18 (non démontré)

Soit $N \sim \text{Mult}(n, p_1, \ldots, p_{I-1})$. *Alors*

$$\sum_{i=1}^{I} \frac{(N_i - np_i)^2}{np_i} \xrightarrow{\mathcal{L}} \chi^2_{I-1}$$

quand $n \to \infty$.

En pratique, l'approximation χ^2 ci-dessus est considérée satisfaisante quand

(i) $n \geq 30$,

(ii) $np_i \geq 5$ pour au moins 80% des valeurs de i, et

(iii) $np_i \geq 1$ pour toutes les valeurs de i.

Résultats asymptotiques

Théorème 19 (non démontré)

Soit $N \sim \text{Mult}(n, p_1(\theta), \ldots, p_{I-1}(\theta))$, *avec* $\theta \in \Theta \subset \mathbb{R}^k$, *où* $\text{int}(\Theta) \neq \emptyset$. *Alors, en désignant par* $\hat{\theta}$ *le MLE de* θ, *on a que, pour tout* $\theta \in \Theta$,

$$\sum_{i=1}^{I} \frac{(N_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})} \xrightarrow{\mathcal{L}} \chi^2_{I-1-k}$$

sous $\mathrm{P}_\theta^{(n)}$ *quand* $n \to \infty$.

On supposera l'approximation satisfaisante sous les mêmes conditions que ci-dessus.

Test d'adéquation : introduction

Soit E une expérience aléatoire dont les résultats possibles sont $A_1, \ldots, A_I$. Notons $p_1, \ldots, p_I$ les probabilités correspondantes. Soit le problème de test

$$\begin{cases} \mathcal{H}_0 : p_i = p_{i0} \text{ pour tout } i = 1, \ldots, I \\ \mathcal{H}_1 : p_i \neq p_{i0} \text{ pour au moins un } i = 1, \ldots, I, \end{cases}$$

où les p_{i0}, $i = 1, \ldots, I$, sont des nombres positifs fixés tels que $\sum_{i=1}^{I} p_{i0} = 1$.

Un test naturel rejette $\mathcal{H}_0$ pour les grandes valeurs de

$$Q^{(n)} = \sum_{i=1}^{I} \frac{(N_i - np_{i0})^2}{np_{i0}},$$

une "distance" entre les effectifs observés N_i et les effectifs attendus np_{i0} sous $\mathcal{H}_0$.

Test d'adéquation

Par le Théorème 18, le test qui rejette $\mathcal{H}_0$ quand

$$Q^{(n)} = \sum_{i=1}^{I} \frac{(N_i - np_{i0})^2}{np_{i0}} > \chi^2_{I-1,1-\alpha},$$

est asymptotiquement de niveau α.

Ici, $\chi^2_{I-1,1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ de la loi χ^2_{I-1}.

Notations pour les quantiles

- $\chi^2_{k,\alpha}$ désignera le quantile d'ordre α de la loi χ^2_k.
- $t_{k,\alpha}$ désignera le quantile d'ordre α de la loi t_k.
- $\ldots$
- Exception: z_α désignera le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{N}(0, 1)$, i.e. $\Phi^{-1}(1 - \alpha)$. Donc le quantile d'ordre α correspondant est $-z_\alpha$.

De nombreuses tables de ces quantiles existent (cfr annexe du syllabus).

Exemple : “vérification” des lois de Mendel

Lors de croisements de certains pois, les lois de Mendel prévoient d’obtenir des descendants jaunes lisses, jaunes ridés, verts lisses et verts ridés en proportions respectives $\frac{9}{16}$, $\frac{3}{16}$, $\frac{3}{16}$ et $\frac{1}{16}$.

Pour confronter la théorie aux données, Mendel considère le problème

$$\begin{cases} \mathcal{H}_0 : (p_1, p_2, p_3, p_4) = (\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}) \\ \mathcal{H}_1 : (p_1, p_2, p_3, p_4) \neq (\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}) \end{cases}$$

et effectue $n = 556$ croisements, qui livrent les effectifs $n_1 = 315$, $n_2 = 101$, $n_3 = 108$ et $n_4 = 32$.

Exemple : “vérification” des lois de Mendel

La statistique de test $Q^{(n)}$ prend la valeur

$$\begin{aligned} q^{(n)} &= \frac{(315 - 556 \times \frac{9}{16})^2}{556 \times \frac{9}{16}} + \frac{(101 - 556 \times \frac{3}{16})^2}{556 \times \frac{3}{16}} \\ &\quad + \frac{(108 - 556 \times \frac{3}{16})^2}{556 \times \frac{3}{16}} + \frac{(32 - 556 \times \frac{1}{16})^2}{556 \times \frac{1}{16}} \\ &\approx 0.47. \end{aligned}$$

Au niveau $\alpha = 1\%$, on ne rejette pas $\mathcal{H}_0$ puisque $0.47 \leq \chi^2_{3,0.99} \approx 11.34$.
Au niveau $\alpha = 5\%$, on ne rejette pas $\mathcal{H}_0$ puisque $0.47 \leq \chi^2_{3,0.95} \approx 7.81$.
Au niveau $\alpha = 10\%$, on ne rejette pas $\mathcal{H}_0$ puisque $0.47 \leq \chi^2_{3,0.90} \approx 6.25$.

p-valeur

Différents niveaux α demandent d'utiliser différentes valeurs critiques $\chi^2_{3,1-\alpha}$.

Mais si on pose

$$\text{p-valeur} = \mathrm{P}^{(n)}_{\mathcal{H}_0}[Q^{(n)} > q^{(n)}]$$

(où $q^{(n)}$ désigne encore la valeur observée de $Q^{(n)}$), le test rejette $\mathcal{H}_0$ si et seulement si la p-valeur est strictement inférieure à α.

Donc la p-valeur permet d'effectuer le test pour toutes les valeurs de α!

Dans l'exemple de Mendel, la p-valeur vaut 0.9254(!)

Tests d'ajustement

Il est courant d'avoir un "modèle" qui spécifie les probabilités p_i à un ou plusieurs paramètre(s) près. Le problème de test est alors

$$\begin{cases} \mathcal{H}_0 : \exists\theta \in \Theta \text{ tel que } p_i = p_i(\theta) \text{ pour tout } i = 1, \ldots, I \\ \mathcal{H}_1 : \forall\theta \in \Theta,\ p_i \neq p_i(\theta) \text{ pour au moins un } i = 1, \ldots, I, \end{cases}$$

où $\Theta \subset \mathbb{R}^k$ est d'intérieur non vide.

Par le Théorème 19, le test qui rejette $\mathcal{H}_0$ quand

$$\sum_{i=1}^{I} \frac{(N_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})} > \chi^2_{I-1-k,1-\alpha}$$

est asymptotiquement de niveau α (pour rappel, $\hat{\theta}$ désigne le MLE de θ sous $\mathcal{H}_0$).

Exemple : groupes sanguins

Le groupe sanguin MN se compose de trois phénotypes distincts: M, MN, et N.

Une population appartenant à ce groupe est dite *"stable dans sa composition"* si les proportions relatives à ces trois phénotypes sont de la forme

$$\begin{aligned} \mathrm{P}[\mathrm{M}] &= p_1(\theta) = \theta^2, \\ \mathrm{P}[\mathrm{MN}] &= p_2(\theta) = 2\theta(1-\theta), \\ \mathrm{P}[\mathrm{N}] &= p_3(\theta) = (1-\theta)^2, \end{aligned}$$

pour un certain $\theta \in \,]0,1[$.

Un échantillon de taille $n = 500$ est prélevé dans une population, et donne lieu à $n_1 = 125$ individus présentant le phénotype M, $n_2 = 225$ individus présentant le phénotype MN, et $n_3 = 150$ individus présentant le phénotype N.

Sur base de cet échantillon, testons au niveau $\alpha = 5\%$ l'hypothèse nulle que la population est stable dans sa composition.

Exemple : groupes sanguins

On vérifiera que

$$\hat{\theta} = \frac{2n_1 + n_2}{2n} = \frac{19}{40}$$

et

$$q^{(n)} \approx 4.777.$$

La p-valeur associée au test est donc

$$P[Y > 4.777] = 0.029, \qquad \text{où } Y \sim \chi^2_{3-1-1}.$$

Par conséquent, ...

Tests d'adéquation pour des lois générales

Si on souhaite tester l'hypothèse nulle que les observations proviennent

- d'une loi commune discrète avec un nombre infini de valeurs possibles, ou
- d'une loi continue,

alors se ramener au schéma multinomial nécessite de constituer des classes.

Exemple : soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d. représentant les nombres d'accidents de voiture par heure dans une région donnée. On peut vouloir tester $\mathcal{H}_0 : X_1 \sim \mathcal{P}(\lambda)$ pour un certain $\lambda > 0$.

Tronquer le nombre d'accidents à $I - 1$ ($\rightsquigarrow I$ valeurs possibles: $0, 1, \ldots, I - 1$) permet de rendre le problème compatible avec le schéma multinomial.

Test d'adéquation de Poisson

Soit $N = (N_1, \ldots, N_I)$, où

$$N_i = \sum_{j=1}^{n} \mathbb{I}[X_j = i-1], \quad i = 1, \ldots, I-1 \;\text{ et }\; N_I = \sum_{j=1}^{n} \mathbb{I}[X_j \geq I-1].$$

Sous $\mathcal{H}_0$, ceci est associé à des probabilités de type Poisson données par

$$p_i(\lambda) = e^{-\lambda} \frac{\lambda^{i-1}}{(i-1)!}, \quad i = 1, \ldots, I-1 \;\text{ et }\; p_I(\lambda) = 1 - \sum_{i=1}^{I-1} p_i(\lambda).$$

Le test rejettera $\mathcal{H}_0$ au niveau asymptotique α si

$$\sum_{i=1}^{I} \frac{(N_i - np_i(\hat{\lambda}))^2}{np_i(\hat{\lambda})} > \chi^2_{I-2,1-\alpha},$$

où $\hat{\lambda} (= \bar{X})$ est le MLE de λ.

Tests d'homogénéité : introduction

On observe ici J multinomiales mutuellement indépendantes

$$N_j = (N_{1j}, \ldots, N_{Ij}) \sim \mathrm{Mult}(n_j, p_{1j}, \ldots, p_{I-1,j}), \qquad j = 1, \ldots, J.$$

Le nombre de modalités I est le même pour chaque multinomiale, mais les autres paramètres peuvent dépendre de j.

On considère le problème de test

$$\begin{cases} \mathcal{H}_0 : \forall i \in \{1, \ldots, I\},\ p_{i1} = \ldots = p_{iJ} \\ \mathcal{H}_1 : \exists i \in \{1, \ldots, I\} \text{ tels que } p_{ij} \neq p_{ij'} \text{ pour certains } j \neq j'. \end{cases}$$

Tests d'homogénéité : exemple

Par exemple, les multinomiales N_j peuvent représenter les nombres de personnes qui, parmi les n_j personnes interrogées dans le pays j, se sont prononcées

- en faveur de l'extrême gauche ($i = 1$),
- en faveur de la gauche ($i = 2$),
- $\ldots$
- en faveur de l'extrême droite ($i = I$).

L'hypothèse nulle considérée signifie alors que les poids des différents courants politiques sont les mêmes dans les J pays.

Tests d'homogénéité

Les résultats asymptotiques livrent, que pour tout $j = 1, \ldots, J$,

$$\sum_{i=1}^{I} \frac{(N_{ij} - n_j p_{ij})^2}{n_j p_{ij}} \xrightarrow{\mathcal{L}} \chi^2_{I-1} \quad \text{quand } n \to \infty.$$

Les multinomiales étant mutuellement indépendantes, ceci livre

$$\sum_{j=1}^{J} \sum_{i=1}^{I} \frac{(N_{ij} - n_j p_{ij})^2}{n_j p_{ij}} \xrightarrow{\mathcal{L}} \chi^2_{J(I-1)} \quad \text{quand } n \to \infty.$$

Sous $\mathcal{H}_0$, ceci se réécrit

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(N_{ij} - n_j p_i)^2}{n_j p_i} \xrightarrow{\mathcal{L}} \chi^2_{J(I-1)} \quad \text{quand } n \to \infty,$$

où p_i est la valeur commune des probabilités $p_{i1}, \ldots, p_{iJ}$.

Tests d'homogénéité

Sous $\mathcal{H}_0$, les MLE de p_i, $i = 1, \ldots, I$ sont

$$\hat{p}_i = \frac{N_{i.}}{n} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{j=1}^{J} N_{ij}, \qquad i = 1, \ldots, I.$$

Le nombre de paramètres indépendants estimés valant $I - 1$ (pourquoi?),

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(N_{ij} - n_j \hat{p}_i)^2}{n_j \hat{p}_i} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(N_{ij} - \frac{N_{i.} n_j}{n})^2}{\frac{N_{i.} n_j}{n}} \xrightarrow{\mathcal{L}} \chi^2_{(I-1)(J-1)}$$

sous $\mathcal{H}_0$ quand $n \to \infty$.

Le test qui en résulte rejette donc $\mathcal{H}_0$ au niveau asymptotique α quand

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(N_{ij} - \frac{N_{i.} n_j}{n})^2}{\frac{N_{i.} n_j}{n}} > \chi^2_{(I-1)(J-1), 1-\alpha}.$$

Tests d'homogénéité : exemple

Lors d'une période de récession économique, 100 chefs d'entreprises, 100 économistes et 100 hommes politiques ont été interrogés sur les décisions qui, selon eux, seraient les plus favorables à une sortie de la récession.

	CEO	Economistes	Politiciens
Accroître le déficit budgétaire	10	15	29
Réduire les impôts	37	37	33
Réduire les taux d'intérêts	24	34	25
Offrir des avantages fiscaux aux entreprises	29	14	13
	$n_1 = 100$	$n_2 = 100$	$n_3 = 100$

Peut-on conclure, au niveau asymptotique $\alpha = 5\%$, que les opinions sont réparties de façon différente dans ces trois groupes?

Tests d'homogénéité : exemple

Dans ce contexte,

- la statistique de test prend la valeur 21.877, et
- la valeur critique (au niveau asymptotique 5%) vaut 12.59.

$\Rightarrow R\mathcal{H}_0$: on conclut, au niveau asymptotique $\alpha = 5\%$, que les opinions sont réparties de façon différente dans les trois groupes

Ceci est confirmé par la p-valeur, qui vaut 0.0013.

Tests d'indépendance : introduction

Soit un vecteur aléatoire (X, Y), où X est à valeurs dans $\{x_1, \ldots, x_I\}$ et Y est à valeurs dans $\{y_1, \ldots, y_J\}$. La distribution jointe de (X, Y) est décrite par

$$p_{ij} = \mathrm{P}[X = x_i, Y = y_j], \quad i = 1, \ldots, I, \;\; j = 1, \ldots, J.$$

Les distributions marginales de X et de Y sont associées à

$$p_{i.} = \mathrm{P}[X = x_i] = \sum_{j=1}^{J} p_{ij}, \quad i = 1, \ldots, I$$

et

$$p_{.j} = \mathrm{P}[Y = y_j] = \sum_{i=1}^{I} p_{ij}, \quad j = 1, \ldots, J.$$

Tests d'indépendance : introduction

Sur la base d'observations i.i.d. $(X_1, Y_1), \ldots, (X_n, Y_n)$, on veut tester l'hypothèse nulle $\mathcal{H}_0 : X \perp\!\!\!\perp Y$. Le problème de test est donc

$$\begin{cases} \mathcal{H}_0 : \forall i, j,\ p_{ij} = p_{i.}p_{.j} \\ \mathcal{H}_1 : \exists i, j \text{ tel que } p_{ij} \neq p_{i.}p_{.j}. \end{cases}$$

Si on note $N_{ij} = \sum_{\ell=1}^{n} \mathbb{I}[(X_\ell, Y_\ell) = (x_i, y_j)]$ pour tout i, j, alors

$$N = (N_{11}, N_{12}, \ldots, N_{IJ})$$

est de loi multinomiale de paramètres $n, p_{11}, p_{12}, \ldots, p_{IJ-1}$. Donc

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(N_{ij} - np_{ij})^2}{np_{ij}} \xrightarrow{\mathcal{L}} \chi^2_{IJ-1}$$

quand $n \to \infty$.

Statistique de test sous $\mathcal{H}_0$ et MLE

Sous $\mathcal{H}_0$,

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(N_{ij}-np_{i.}p_{.j})^2}{np_{i.}p_{.j}} \overset{\mathcal{L}}{\to} \chi^2_{IJ-1} \quad \text{quand } n\to\infty.$$

Sous $\mathcal{H}_0$, les MLE des $p_{i.}$ et des $p_{.j}$ sont

$$\hat{p}_{i.} = \frac{N_{i.}}{n} \overset{\text{def}}{=} \frac{1}{n}\sum_{j=1}^{J} N_{ij}, \qquad i=1,\ldots,I$$

et

$$\hat{p}_{.j} = \frac{N_{.j}}{n} \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{I} N_{ij}, \qquad j=1,\ldots,J.$$

Tests d'indépendance

Le nombre de paramètres indépendants estimés valant $(I-1)+(J-1)$,

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(N_{ij}-n\hat{p}_{i.}\hat{p}_{.j})^2}{n\hat{p}_{i.}\hat{p}_{.j}}=\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(N_{ij}-\frac{N_{i.}N_{.j}}{n})^2}{\frac{N_{i.}N_{.j}}{n}}\xrightarrow{\mathcal{L}}\chi^2_{(I-1)(J-1)}$$

sous $\mathcal{H}_0$ quand $n\to\infty$.

Le test qui en résulte rejette donc $\mathcal{H}_0$ au niveau asymptotique α quand

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(N_{ij}-\frac{N_{i.}N_{.j}}{n})^2}{\frac{N_{i.}N_{.j}}{n}}>\chi^2_{(I-1)(J-1),1-\alpha}.$$

Tests d'indépendance : exemple

Un ensemble de $n = 2100$ certificats de décès d'adultes, enregistrés dans les hôpitaux d'une grande ville, ont mené à la classification suivante.

$Y \setminus X$	Gros fumeurs	Fumeurs légers	Non-fumeurs
Origine respiratoire	55	120	162
Origine cardio-vasculaire	49	388	315
Autre origine	61	300	650

Peut-on conclure, au niveau asymptotique $\alpha = 5\%$, que les habitudes tabagiques et la cause de décès ne sont pas indépendantes?

Tests d'indépendance : exemple

Dans ce contexte,

- la statistique de test prend la valeur 134.12, et
- la valeur critique (au niveau asymptotique 5%) vaut 9.49.

$\Rightarrow R\mathcal{H}_0$: on conclut, au niveau asymptotique $\alpha = 5\%$, que les habitudes tabagiques et la cause de décès ne sont pas indépendantes.

Ceci est confirmé par la p-valeur, qui est inférieure à 10^{-15}.

Différence entre tests d'homogénéité et d'indépendance

Les expressions des statistiques de test pour les tests d'homogénéité et d'indépendance sont identiques.

La différence entre le test d'homogénétié et le test d'indépendance est donc dans l'interprétation du problème de test...

- Dans le premier cas, le statisticien se demande si un facteur donné se répartit de la même manière dans différentes *populations*.
- Le second test cherche à mesurer l'indépendance entre deux facteurs (qui jouent là un rôle symétrique).