Chapitre 4. Tests d’hypothèses

Pierre Duchesne

March 12, 2017

Soit X un caractère étudié. Soit $\{\mathcal{F}(x;\theta), \theta \in \Theta\}$ la famille des lois possibles de X, où Θ est ce que l'on a appelé l'espace paramétrique.

Notions sur les tests d'hypothèses

Pour les tests d'hypothèses, l'espace paramétrique est partitionné en deux sous-espaces, c'est-à-dire

$$\Theta = \Theta_0 \cup \Theta_1,$$

avec

$$\Theta_0 \cap \Theta_1 = \phi.$$

L'objectif est de savoir si $\theta \in \Theta_0$ ou encore si $\theta \in \Theta_1$.

Exemples

1. Le gouvernement du Québec s'intéresse aux cyanobactéries dans les lacs de l'Estrie. Les spécialistes du ministère du Développement durable, de l'Environnement et des Parcs veulent vérifier l'intensité de la prolifération des cyanobactéries. Supposons qu'en bas de 20 000 cellules par millilitre d'eau, la situation est considérée tolérable, alors qu'au dessus de ce seuil, le lac est considéré impropre à la baignade ou à la consommation.

Ici

$$\Theta = [0, \infty[,$$

avec

$$\Theta_0 = [0, 20000]$$

et

$$\Theta_1 =]20000, \infty[$$

Exemples (suite)

2. Supposons que votre ami possède deux pièces de monnaie. La première est une pièce régulière, donc la probabilité d'obtenir pile = probabilité d'obtenir face = 1/2. La seconde est pipée (truquée), de sorte que la probabilité d'obtenir face est 7/10, alors que la probabilité d'avoir pile est 3/10.
Votre ami, prend une des deux pièces, la lance n fois, et vous dit le nombre de faces obtenus sur les n lancers. Vous voulez savoir si la pièce choisie est la pièce régulière ou truquée. Ici le caractère $X \sim bin(n, \theta)$ et $\Theta = [0, 1]$, avec

$$\Theta_0 = 1/2,$$

et

$$\Theta_1 = \Theta - \Theta_0 = [0, 1/2[\cup]1/2, 1].$$

Définition

Une **hypothèse statistique** est un énoncé portant sur la loi du caractère X. Elle est dite **simple**, lorsqu'elle détermine complètement la loi de X. si ce n'est pas le cas, on dira que l'hypothèse est **composée**.

Exemples

1. $$\begin{cases} H_0 : \theta \in \Theta_0, & \text{i.e. } \theta \in [0, 20000] \Rightarrow \text{ hypothèse composée;} \\ H_1 : \theta \in \Theta_1, & \text{i.e. } \theta \in [20000, \infty[\Rightarrow \text{ hypothèse composée;} \end{cases}$$
2. $$\begin{cases} H_0 : \theta \in \Theta_0, & \text{i.e. } \theta = 1/2 \Rightarrow \text{ hypothèse simple;} \\ H_1 : \theta \in \Theta_1, & \text{i.e. } \theta \neq 1/2 \Rightarrow \text{ hypothèse composée.} \end{cases}$$

Effectuer un test d'hypothèses, c'est confronter deux hypothèses en vue d'en rejeter une, et donc, d'accepter l'autre (parfois faute de mieux).

L'hypothèse H_0 est appelée hypothèse principale ou hypothèse nulle. Souvent, c'est la plus simple des deux hypothèses.

L'hypothèse H_1 est appelée hypothèse alternative ou encore contre-hypothèse.

Soit $\mathcal{S}$ le support du caractère X. On rappelle que le support de X est l'ensemble des valeurs que X est susceptible de prendre. Ainsi, l'échantillon aléatoire $X_1, \ldots, X_n$ prend ses valeurs dans

$$\mathcal{S} \times \cdots \times \mathcal{S} = \mathcal{S}^n.$$

Formellement, nous avons alors un espace de décision:

$$\mathcal{D} = \{\text{accepter } H_0, \text{rejeter } H_0\}$$

Accepter H_0 revient à rejeter H_1; De même, rejeter H_0 revient à accepter H_1.
Remarque: on verra que lors de *l'acceptation* de H_0, on préfèrera ajouter dans ce cas et compte tenu de notre cadre théorique l'expression: *faute de mieux*.

Définition

Un test statistique $\mathcal{T}$ est une application:

$$\delta_n \longmapsto \mathcal{D},$$

c'est-à-dire qu'à chaque réalisation de l'échantillon aléatoire, T fait correspondre une et une seule décision.

Exemple

Le test qui consiste à rejeter H_0 si la moyenne échantillonnale $\bar{x} > 29$, sinon de conserver H_0, faute de mieux, est en fait le test statistique:

$$T(x_1, \ldots, x_n) = \begin{cases} \text{rejeter } H_0, & \text{si } (x_1, \ldots, x_n) \text{ est tel que } \bar{x} > 29; \\ \text{accepter } H_0, & \text{si } (x_1, \ldots, x_n) \text{ est tel que } \bar{x} \leq 29, \end{cases}$$

où $(x_1, \ldots, x_n) \in \mathcal{S}^n$.

Définition

La **région critique** d'un test T est l'ensemble qui points $(x_1, \ldots, x_n) \in \mathcal{S}^n$ qui mènent au rejet de l'hypothèse nulle. Plus précisément, c'est l'ensemble

$$\mathcal{C} = \{(x_1, \ldots, x_n) \in \mathcal{S}^n \mid T(x_1, \ldots, x_n) = \text{ 'rejeter } H_0\text{' }\}.$$

On appelle également la région critique la **région de rejet de l'hypothèse nulle** associée au test T.

À chaque test T correspond une région critique $\mathcal{C}$.
Inversement, pour tout sous-ensemble $\mathcal{C}$ de $\mathcal{S}^n$, on peut trouver un test T dont la région critique est $\mathcal{C}$.
En effet, il suffit de poser:

$$T(x_1, \ldots, x_n) = \begin{cases} \text{rejeter } H_0, & \text{si } (x_1, \ldots, x_n) \in \mathcal{C}, \\ \text{accepter } H_0, & \text{sinon.} \end{cases}$$

Il y a donc une correspondance biunivoque entre les tests T et les sous-ensembles $\mathcal{C}$ de $\mathcal{S}^n$.

Erreurs de type I et II

Lorsque l'on confronte deux hypothèses à l'aide d'un test T, on peut potentiellement commettre une de deux erreurs.

		Décision	
		Rejeter H_0	Accepter H_0
Vérité	H_0 est vraie	Erreur de première espèce	Bonne décision
	H_1 est vraie	Bonne décision	Erreur de seconde espèce

La probabilité de commettre les erreurs de type I et II sont les risques de première et seconde espèce, respectivement.

Soit T un test de région critique $\mathcal{C}$.

Définition: Risque de première espèce Le risque de première espèce associé à T est la probabilité de commettre l'erreur de première espèce. On le note traditionnellement α:

$$\begin{aligned} \alpha &= P(\text{ rejeter } H_0 \mid H_0 \text{ est vraie }), \\ &= P((X_1, \ldots, X_n) \in \mathcal{C} \mid H_0). \end{aligned}$$

Définition: Risque de seconde espèce Le risque de seconde espèce associé à T est la probabilité de commettre l'erreur de seconde espèce. On le note traditionnellement β:

$$\begin{aligned} \beta &= P(\text{ accepter } H_0 \mid H_1 \text{ est vraie }), \\ &= P((X_1, \ldots, X_n) \notin \mathcal{C} \mid H_1). \end{aligned}$$

Définition: Niveau d'un test T

La probabilité de commettre l'erreur de première espèce est appelée le niveau du test T. C'est α.

Définition: Puissance d'un test T

La quantité $1 - \beta$ est la puissance d'un test T. On note que par définition:

$$\begin{aligned} 1 - \beta &= P(\text{ rejeter } H_0 \mid H_1 \text{ est vraie }), \\ &= P((X_1, \ldots, X_n) \in \mathcal{C} \mid H_1). \end{aligned}$$

En d'autres termes, la probabilité de rejeter H_0 alors que H_0 devrait (idéalement) être rejetée est la puissance du test.

Tant que l'on a pas observé un échantillon, on ne sait pas quelle décision on est susceptible de prendre et on ne sait donc pas quelle erreur on risque de commettre.

Idéalement, on voudrait minimiser simultanément les deux risques.

Malheureusement, ceci est généralement impossible car on constate qu'un α petit va de pair avec un β grand, et un α grand va de pair avec un β petit.

Maximisation de la puissance pour un niveau donné

La stratégie consiste à fixer le niveau du test α à une valeur jugée acceptable. On fixe donc le risque de première espèce. Pour cette valeur de α spécifiée, on cherche alors à maximiser la puissance $1 - \beta$, c'est-à-dire à minimiser β pour α donné.

Exemple

Soit X un caractère tel que $X \sim bin(2, \theta)$, où $\theta \in \Theta = \{\frac{1}{2}, \frac{3}{4}\}$.
Posons $\Theta_0 = \{\frac{1}{2}\}$ et $\Theta_1 = \{\frac{3}{4}\}$.
On cherche donc à confronter:

$$\begin{cases} H_0 : \theta \in \Theta_0, & \text{c'est-à-dire } \theta = 1/2, \\ H_1 : \theta \in \Theta_1, & \text{c'est-à-dire } \theta = 3/4. \end{cases}$$

Tous les tests possibles (ou de manière toutes les régions critiques possibles) sont donc:

$$\{\phi, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\}.$$

On note que l'ensemble précédent est constitué de tous les sous-ensembles possibles de $\mathcal{S} = \{0, 1, 2\}$.

Ainsi, X est un caractère tel que $X \sim bin(2, \theta)$, avec $\theta \in \Theta = \{\frac{1}{2}, \frac{3}{4}\}$.
Ainsi la loi de X est:

$$\begin{aligned} p_X(x;\theta) &= \binom{2}{x}\theta^x(1-\theta)^{2-x}, \\ &= \binom{2}{x}(1-\theta)^2\left\{\frac{\theta}{1-\theta}\right\}^x, \; x = 0, 1, 2. \end{aligned}$$

On trouve

x	$p_X(x; 1/2)$	$p_X(x; 3/4)$
0	1/4	1/16
1	1/2	6/16
2	1/4	9/16

Rappel: On a que X est un caractère tel que $X \sim bin(2, \theta)$.
On désire tester:

$$\begin{cases} H_0 : \theta = 1/2, \\ H_1 : \theta = 3/4. \end{cases}$$

Les régions critiques (i.e. les zones de rejet de H_0) possibles sont:

$$\{\phi, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\}.$$

avec

x	$p_X(x; 1/2)$	$p_X(x; 3/4)$
0	1/4	1/16
1	1/2	6/16
2	1/4	9/16

x	$p_X(x;1/2)$	$p_X(x;3/4)$
0	1/4	1/16
1	1/2	6/16
2	1/4	9/16

Le risque de première espèce dans notre exemple est:

$$\alpha = P(\text{ rejeter } H_0 \mid H_0 \text{ est vraie}) = \sum_{x \in \mathcal{C}} p_X(x;1/2)$$

Le risque de seconde espèce est quant à lui:

$$\beta = P(\text{ accepter } H_0 \mid H_1 \text{ est vraie}) = \sum_{x \notin \mathcal{C}} p_X(x;3/4)$$

x	$p_X(x;1/2)$	$p_X(x;3/4)$
0	1/4	1/16
1	1/2	6/16
2	1/4	9/16

$$
\begin{aligned}
\alpha &= P(\text{ rejeter } H_0 \mid H_0 \text{ est vraie}) = \sum_{x \in \mathcal{C}} p_X(x;1/2), \\
\beta &= P(\text{ accepter } H_0 \mid H_1 \text{ est vraie}) = \sum_{x \notin \mathcal{C}} p_X(x;3/4)
\end{aligned}
$$

$\mathcal{C}$	ϕ	$\{0\}$	$\{1\}$	$\{2\}$	$\{0,1\}$	$\{0,2\}$	$\{1,2\}$	$\{0,1,2\}$
α	0/4	1/4	2/4	1/4	3/4	2/4	3/4	4/4
β	16/16	15/16	10/16	7/16	9/16	6/16	1/16	0/16

$\mathcal{C}$	ϕ	$\{0\}$	$\{1\}$	$\{2\}$	$\{0,1\}$	$\{0,2\}$	$\{1,2\}$	$\{0,1,2\}$
α	0/4	1/4	2/4	1/4	3/4	2/4	3/4	4/4
β	16/16	15/16	10/16	7/16	9/16	6/16	1/16	0/16
$1-\beta$	0/16	1/16	6/16	9/16	7/16	10/16	15/16	16/16

On remarque que α et β varient en sens inverse.

Si on fixe $\alpha = 1/4$, le meilleur test est associé à $\mathcal{C} = \{2\}$, et sa puissance est $1-\beta = 9/16$.

Si on fixe $\alpha = 2/4$, le meilleur test est associé à $\mathcal{C} = \{0,2\}$, et sa puissance est $1-\beta = 10/16$.

Si on fixe $\alpha = 3/4$, le meilleur test est associé à $\mathcal{C} = \{1,2\}$, et sa puissance est $1-\beta = 15/16$.

Confrontation de deux hypothèses simples

On cherche à développer une théorie nous permettant de dégager le test le plus puissant (test PP).
Dans un premier temps, on considère l'espace paramétrique le plus simple possible, composé de deux valeurs. Ainsi, l'espace paramétrique est

$$\Theta = \{\theta_0, \theta_1\}.$$

Les deux hypothèses que l'on confronte sont simples, et sont:

$$\begin{cases} H_0 : \theta = \theta_0, \\ H_1 : \theta = \theta_1. \end{cases}$$

Définition: test le plus puissance (PP) de niveau α

Le test T de région critique $\mathcal{C}$ est dit le test le plus puissance de niveau α pour confronter H_0 et H_1 si

a) Le test T est de niveau α, c'est-à-dire que

$$P((X_1, \ldots, X_n) \in \mathcal{C} \mid H_0) = \alpha;$$

b) Le test T est dans la classe des tests de niveau α dont la puissance est la plus élevée. Autrement dit:

$$1-\beta = P((X_1, \ldots, X_n) \in \mathcal{C} | H_1) \geq 1-\beta' = P((X_1, \ldots, X_n) \in \mathcal{C}' | H_1)$$

pour tout autre test T' de région critique $\mathcal{C}'$, et de niveau α.

Rapport de vraisemblance

On procède avec notre exemple: X est un caractère tel que $X \sim bin(2, \theta)$. On désire tester:

$$\begin{cases} H_0 : \theta = 1/2, \\ H_1 : \theta = 3/4. \end{cases}$$

x	$p_X(x; 1/2)$	$p_X(x; 3/4)$	$p_X(x; 1/2)/p_X(x; 3/4)$
0	1/4	1/16	4
1	1/2	6/16	4/3
2	1/4	9/16	4/9

Le rapport $p_X(x; 1/2)/p_X(x; 3/4)$ est appelé le rapport de vraisemblance (RV). Dans cet exemple, à mesure que x décroit, x devient plus favorable à H_0 (si vous observez $x = 0$, il semble plus vraisemblable que $\theta = 1/2$). Ici, le RV est monotone en x.

Lemme de Neyman-Pearson

Soit X un caractère étudié. Soit $\mathcal{E} : X_1, \ldots, X_n$ un échantillon aléatoire. Supposons que nous disposons de la réalisation $x_1, \ldots, x_n$. Si on confronte les deux hypothèses

$$\begin{cases} H_0 : \theta = \theta_0, \\ H_1 : \theta = \theta_1, \end{cases}$$

alors le test le plus puissant de niveau α est celui dont la région critique est donnée par

$$\mathcal{C} = \left\{ (x_1, \ldots, x_n) \mid \frac{L(\theta_0; x_1, \ldots, x_n)}{L(\theta_1; x_1, \ldots, x_n)} \leq c_\alpha \right\},$$

où la constante c_α est choisie de sorte que:

$$P((X_1, \ldots, X_n) \in \mathcal{C} \mid H_0) = \alpha.$$