

Tests d'hypothèses

L'objectif d'un test d'hypothèse paramétrique est de répondre à la question que l'on se pose sur la forme de la loi sous-jacente. Est-ce qu'au vu d'une observation d'un échantillon $\mathbf{X} = (X_1, \dots, X_n)$, les X_j sont des v.a. iid, on peut décider entre les deux possibilités H_0 et H_1 ? Ici i pour $i = 0, 1$ et $\mathcal{H} = \{H_0, H_1\}$.

Exemple 1. Dans un jeu au pile ou face, on définit la variable aléatoire

$$X = \begin{cases} 1 & \text{si on obtient face} \\ 0 & \text{si on obtient pile} \end{cases}$$

Si la pièce n'est pas truquée, on doit s'attendre à avoir la même probabilité d'obtenir pile ou face, c'est-à-dire $P(X=1) = P(X=0) = 1/2$. ($X \sim \text{Ber}(1/2)$, $X \in \{0, 1\}$)

Au vu de l'observation d'un échantillon (X_1, \dots, X_n) où les X_j sont iid $\text{Ber}(1/2)$ on pourrait tester les hypothèses suivantes:

$$H_0: P(X=1) = 1/2 \text{ contre } H_1: P(X=1) = 1/3$$

$$H_0: P(X=1) = 1/2 \text{ contre } H_1: P(X=1) = 2/3$$

$$H_0: P(X=1) = 1/2 \text{ contre } H_1: P(X=1) < 1/2$$

$$H_0: P(X=1) = 1/2 \text{ contre } H_1: P(X=1) > 1/2$$

H_0 etc...

Exemple 2. Supposons que l'on s'intéresse à déterminer si un médicament fait baisser la tension artérielle chez certains patients. Soit X_j la différence entre les deux mesures de la tension artérielle du patient j . (2 mesures: avant et après l'administration du médicament). Supposons que X_1, \dots, X_n sont des v.a. iid $N(\mu, \sigma^2)$. On va tester

$$H_0: \mu = 0 \text{ contre } H_1: \mu < 0$$

H_0 représente l'hypothèse que le médicament n'a aucun effet, H_1 celle de l'effet négatif, c'est-à-dire que l'administration du médicament entraîne une baisse de la tension artérielle.

On suppose que (X_1, \dots, X_n) est un échantillon d'une loi appartenant au modèle $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Soient Θ_0 et Θ_1 deux sous-ensembles de Θ tels que $\Theta_0 \cap \Theta_1 = \emptyset$.

On appelle test d'hypothèse une règle de décision qui permet de décider entre les hypothèses H_0 et H_1 . H_0 est appelée hypothèse nulle et H_1 l'hypothèse alternative.

Un test est une statistique $T(\mathbf{X}_1, \dots, \mathbf{X}_n)$ qui prend des valeurs dans $\{0, 1\}$:

$$T(\mathbf{X}_1, \dots, \mathbf{X}_n) = \begin{cases} 1 & \text{si on rejette } H_0 \text{ (et donc accepte } H_1) \\ 0 & \text{si on accepte } H_0 \end{cases}$$

Un test est déterminé par sa région critique (ou zone de rejet) W :

$$W = \{\mathbf{x} = (x_1, \dots, x_n) \text{ tel que } T(\mathbf{x}) = 1\} \subset \mathcal{X}^n.$$

On appelle *erreur (risque) de 1ere espece* le rejet de H_0 fort. Cette erreur de 1ere espece est mesurée par la probabilité

$$P(X = W) = P(\text{rejeter } H_0) \text{ pour } \theta_0.$$

On appelle *erreur (risque) de 2eme espece* le rejet de H_1 fort. Cette erreur est mesurée par la probabilité

$$P(X \neq W) = P(\text{accepter } H_0) \text{ pour } \theta_1.$$

On appelle *puissance du test* la fonction $\beta : \theta \in [0, 1] \rightarrow [0, 1]$ donnée par $\beta(\theta) = P(\text{rejeter } H_0) \text{ pour } \theta$.

	H_0 est vraie	H_1 est vraie
Accepter H_0	OK	Erreur de 2eme espece
Rejeter H_1	Erreur de 1ere espece	OK

On dira qu'un test est de niveau α ($\alpha \in [0, 1]$) si $\sup_{\theta \in \theta_0} P(X = W) = \alpha$.

On dira qu'un test est de seuil α ($\alpha \in [0, 1]$) si $\sup_{\theta \in \theta_0} P(X = W) \leq \alpha$.

On dira qu'un test de seuil α est uniformément le plus puissant (UPP) si sa puissance est maximale pour tout θ_1 parmi tous les test de seuil α .

En général on choisira $\alpha \in \{0.01, 0.05, 0.1\}$.

Exemple 3. Soit (X_1, \dots, X_n) un échantillon issu de la loi $N(\theta, \sigma^2)$ où σ^2 est connu. On considère le problème de comparer

$$H_0: \theta = \theta_0 \text{ contre } H_1: \theta = \theta_1$$

avec $\theta_0 < \theta_1$.

Un estimateur naturel de θ est la moyenne empirique \bar{X}_n . Il semble naturel de vouloir rejeter H_0 si \bar{X}_n est grande. Le test admet pour région critique

$$W = \{(X_1, \dots, X_n) : \bar{X}_n > k\}$$

et donc il accepte H_0 si $\bar{X}_n \leq k$ et rejette H_0 si $\bar{X}_n > k$. Autrement dit $(X_1, \dots, X_n) \in W$ si $\bar{X}_n > k$. Il nous reste de déterminer k .

Si on veut construire un test de niveau α on doit avoir $\sup_{\theta \in \theta_0} P(\bar{X}_n > k) = \alpha$. Ici $\theta_0 = \{\theta_0\}$ et donc

$$\sup_{\theta \in \theta_0} P(\bar{X}_n > k) = P_{\theta_0}(\bar{X}_n > k) = \alpha.$$

Sous l'hypothèse nulle $\bar{X}_n \sim N(\theta_0, \sigma^2/n)$:

$$\begin{aligned} Z &= \frac{\bar{X}_n - \theta_0}{\sigma/\sqrt{n}} \sim N(0, 1); \\ &= P_{\theta_0}(\bar{X}_n > k) = P_{\theta_0}\left(\frac{\bar{X}_n - \theta_0}{\sigma/\sqrt{n}} > \frac{k - \theta_0}{\sigma/\sqrt{n}}\right) = P\left(Z > \frac{k - \theta_0}{\sigma/\sqrt{n}}\right) \\ &= P\left(Z > \frac{k - \theta_0}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(\frac{k - \theta_0}{\sigma/\sqrt{n}}\right) = \alpha \end{aligned}$$

On en déduit que

$$\frac{k - \mu_0}{\sigma_0 / \sqrt{n}} = Z_{1-\alpha}$$

où z est le quantile de la loi Gaussienne standard. Ce qui permet de donner la formule suivante pour k :

$$k = \mu_0 + \frac{\sigma_0}{\sqrt{n}} Z_{1-\alpha}.$$

Calculons la puissance β de ce test. Par définition

$$\beta = P_{\mu_1}(\bar{X}_n > k) = P_{\mu_1}\left(\bar{X}_n > \mu_0 + \frac{\sigma_0}{\sqrt{n}} Z_{1-\alpha}\right)$$

Sous l'hypothèse alternative

$$Z = \frac{\bar{X}_n - \mu_0}{\sigma_0 / \sqrt{n}} \sim N(0, 1)$$

et donc

$$\beta = P_{\mu_1}\left(\frac{\bar{X}_n - \mu_0}{\sigma_0 / \sqrt{n}} > \frac{\mu_0 - \mu_1}{\sigma_0 / \sqrt{n}} + Z_{1-\alpha}\right) = 1 - P\left(Z \leq \frac{\mu_0 - \mu_1}{\sigma_0 / \sqrt{n}} + Z_{1-\alpha}\right)$$

Or, si $\mu_0 - \mu_1 < 0$,

$$\lim_n P\left(Z \leq \frac{\mu_0 - \mu_1}{\sigma_0 / \sqrt{n}} + Z_{1-\alpha}\right) = 0$$

et donc la puissance converge vers 1 lorsque la taille de l'échantillon $n \rightarrow \infty$.

Les différentes catégories de hypothèses

Les hypothèses simples

Définition 4. On dira qu'un test est un test d'hypothèses simples si les hypothèses sont du type $H_0: \theta = \theta_0$ et $H_1: \theta = \theta_1$ où $\theta_0, \theta_1 \in \Theta$ ($\theta_0 = \{\theta_0\}$ et $\theta_1 = \{\theta_1\}$).

Théorème 5. (Lemme fondamentale de Neyman-Pearson) Soit $P = \{P_\theta : \theta \in \Theta\}$ un modèle paramétrique. On considère le test d'hypothèses simples $H_0: \theta = \theta_0$ et $H_1: \theta = \theta_1$. Le test qui rejette H_0 si

$$\frac{L_n(\mathbf{x}, \theta_1)}{L_n(\mathbf{x}, \theta_0)} > c \quad \text{où} \quad P_{\theta_0}\left(\frac{L_n(\mathbf{X}, \theta_1)}{L_n(\mathbf{X}, \theta_0)} > c\right) = \alpha$$

est un test UPP parmi tous les test de seuil α .

Exemple 6. Reprenons l'exemple 3.

$$\begin{aligned} \frac{L_n(\mathbf{x}, \theta_1)}{L_n(\mathbf{x}, \theta_0)} &= \frac{(2^{-n/2}) \exp\left\{-\frac{1}{2} \sum_{i=1}^n (x_i - \theta_1)^2\right\}}{(2^{-n/2}) \exp\left\{-\frac{1}{2} \sum_{i=1}^n (x_i - \theta_0)^2\right\}} = \exp\left\{-\frac{1}{2} \sum_{i=1}^n [(x_i - \theta_1)^2 - (x_i - \theta_0)^2]\right\} \\ &= \exp\left\{-\frac{(\theta_1 - \theta_0)}{2} \sum_{i=1}^n x_i + \frac{n}{2} (\theta_1^2 - \theta_0^2)\right\} = \exp\left\{-n \frac{(\theta_1 - \theta_0)}{2} \bar{x} + \frac{n}{2} (\theta_1^2 - \theta_0^2)\right\} \end{aligned}$$

D'après le lemme de Neyman-Pearson on le test UPP est de la forme

$$\frac{L_n(\mathbf{X}, \mu_1)}{L_n(\mathbf{X}, \mu_0)} = \exp \left[-n \frac{(\mu_1 - \mu_0)^2}{2\sigma^2} \right] \sum_{j=0}^n \binom{n}{j} \mu_1^j \mu_0^{n-j} > c$$

$$\sum_{j=0}^n \binom{n}{j} \mu_1^j \mu_0^{n-j} > \frac{\frac{2}{\sigma^2} \log c + \frac{n}{2} (\mu_1 - \mu_0)^2}{(\mu_1 - \mu_0)^2} = k$$

où k est déterminé par la condition

$$P_{\mu_0} \left(\frac{L_n(\mathbf{X}, \mu_1)}{L_n(\mathbf{X}, \mu_0)} > c \right) = P_{\mu_0} \left(\sum_{j=0}^n \binom{n}{j} \mu_1^j \mu_0^{n-j} > k \right).$$

Lien avec l'exhaustivité

Proposition 7. On considère les hypothèses simples $H_0: \mu = \mu_0$ et $H_1: \mu = \mu_1$ ($\mu_0 < \mu_1$). Si $S(\mathbf{X})$ est exhaustive pour μ et

$$g(s, \mu) = \begin{cases} p_{S(\mathbf{X})}(s, \mu) & \text{dans le cas discret} \\ f_{S(\mathbf{X})}(s, \mu) & \text{dans le cas continu} \end{cases}$$

alors le test de Neyman et Pearson se réduit au test qui rejette H_0 si

$$\frac{g(S(\mathbf{X}), \mu_1)}{g(S(\mathbf{X}), \mu_0)} > d$$

avec d déterminé par la condition

$$P_{\mu_0} \left(\frac{g(S(\mathbf{X}), \mu_1)}{g(S(\mathbf{X}), \mu_0)} > d \right) = \alpha.$$

Exemple 8. Reprenons encore l'exemple 3. On sait que $S(\mathbf{X}) = \sum_{i=1}^n X_i$ est exhaustive pour μ et que

$$g(s, \mu) = f_{S(\mathbf{X})}(s, \mu) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{\mu - \frac{n}{2}\sigma^2}{2\sigma^2} (s - \frac{n}{2}\sigma^2)^2}.$$

Donc

$$\frac{g(s, \mu_1)}{g(s, \mu_0)} = e^{\frac{\mu_0 - \frac{n}{2}\sigma^2}{2\sigma^2} (s - \frac{n}{2}\sigma^2)^2 - \frac{\mu_1 - \frac{n}{2}\sigma^2}{2\sigma^2} (s - \frac{n}{2}\sigma^2)^2} > d$$

$$\frac{\mu_0 - \frac{n}{2}\sigma^2}{2\sigma^2} (s - \frac{n}{2}\sigma^2)^2 - \frac{\mu_1 - \frac{n}{2}\sigma^2}{2\sigma^2} (s - \frac{n}{2}\sigma^2)^2 > \log d$$

$$s > \frac{1}{2} \mu_0 + \frac{1}{2} \mu_1 + \frac{\frac{2}{\sigma^2} \log d}{n(\mu_1 - \mu_0)} = k.$$

Le test rejette H_0 si $S(\mathbf{X}) = \sum_{i=1}^n X_i > k$ et donc on retrouve que

$$k = \mu_0 + \frac{\sigma^2}{n} Z_{1-\alpha}.$$

Hypothèse simple contre hypothèse composite

Définition 9. On appelle *test entre hypothèse simple et hypothèse composite* tout test où les hypothèses sont du type $H_0: \theta = \theta_0$ et $H_1: \theta \in \Theta_0 / \theta_0$ et $\text{card}(\Theta_0) > 1$.

Proposition 10. On considère le test d'hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta > \theta_0$. Si la région critique du test de Neyman et Pearson pour le test avec hypothèses $H_0: \theta = \theta_0$ et $H_1: \theta = \theta_1$ avec $\theta_1 > \theta_0$ ne dépend pas de θ_1 alors le test de Neyman et Pearson (de niveau α) est le test UPP pour les hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta > \theta_0$ parmi tous les tests de seuil α .

Exemple 11. Reprenons l'exemple 3. On considère les hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta > \theta_0$. On a vu que le test de Neyman et Pearson de niveau α pour les hypothèses simples $H_0: \theta = \theta_0$ contre $H_1: \theta = \theta_1$ ou $\theta_1 > \theta_0$ est de rejeter H_0 si

$$\bar{X}_n > \theta_0 + \frac{\alpha}{n} Z_{1-\alpha}.$$

La région critique de ce test est

$$W = \{x: \bar{x} > \theta_0 + \frac{\alpha}{n} Z_{1-\alpha}\}$$

et ne dépend pas de θ_1 . D'après la proposition précédente, le test de Neyman et Pearson est le test UPP pour les hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta > \theta_0$ parmi tous les tests de seuil α .

Remarque 12. Le test UPP de niveau α pour les hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta > \theta_0$ est la règle de décision qui rejette H_0 si

$$\bar{X}_n < \theta_0 + \frac{\alpha}{n} Z = \theta_0 + \frac{\alpha}{n} Z_{1-\alpha}.$$

Remarque 13. (Exercice) Il n'existe pas un test UPP pour les hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta \neq \theta_0$.

Tests entre hypothèses composites

Définition 14. On appelle *test entre hypothèses composites* tout test où les hypothèses sont du type $H_0: \theta \in \Theta_0$ et $H_1: \theta \in \Theta_1 / \theta \in \Theta_0$, $\text{card}(\Theta_0) > 1$ et $\text{card}(\Theta_1) > 1$.

Théorème 15. (de Karlin-Rubin) On considère les hypothèses $H_0: \theta \in \Theta_0$ contre $H_1: \theta \in \Theta_1$. Si $S(X)$ est une statistique exhaustive pour θ telle que, pour tout $\theta_1 \in \Theta_1$ la fonction

$$s \mapsto \frac{g(s, \theta_1)}{g(s, \theta_0)}$$

est croissante, alors le test qui rejette H_0 si $S(X) > s$ est le test UPP parmi tous les tests de seuil $\alpha = P_{\theta_0}(S(X) > s)$.

Exemple 16. Reprenons encore l'exemple 3. $S(X) = \bar{X}_n$ est exhaustive et

$$\frac{g(s, \theta_1)}{g(s, \theta_0)} = e^{\frac{n}{2}(s - \theta_0)^2 + \frac{n}{2}(s - \theta_1)^2} = e^{\frac{n}{2}(2s(\theta_0 + \theta_1) - \theta_0^2 - \theta_1^2)}$$

est une fonction croissante de s pour tout $\theta_0 > 0$. D'après le théorème de Karlin-Rubin, le test qui rejette H_0 si $S(\mathbf{X}) > s$ est le test UPP parmi tous les tests de seuil $\alpha = P_{\theta_0}(X_{(n)} > s)$ et donc

$$s = \theta_0 + \frac{0}{n} Z_{1-\alpha/\theta_0}.$$

Test du rapport de vraisemblances

On considère en général les hypothèses $H_0: \theta = \theta_0$ contre $H_1: \theta \neq \theta_0$. On pose

$$\lambda_n(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L_n(\mathbf{x}, \theta)}{\sup_{\theta \in \Theta_1} L_n(\mathbf{x}, \theta)}.$$

Définition 17. On appelle test du rapport de vraisemblances le test qui rejette H_0 si

$$\lambda_n(\mathbf{X}) \leq c_n.$$

Exemple 18. Reprenons l'exemple 3 avec $H_0: \theta = \theta_0$ contre $H_1: \theta \neq \theta_0$. Ici $\Theta_0 = \{\theta_0\}$ et $\Theta_1 = \mathbb{R} \setminus \{\theta_0\}$. Donc $\sup_{\theta \in \Theta_1} L_n(\mathbf{x}, \theta) = \sup_{\theta \in \mathbb{R}} L_n(\mathbf{x}, \theta)$.

$$\lambda_n(\mathbf{x}) = \frac{L_n(\mathbf{x}, \theta_0)}{\sup_{\theta \in \mathbb{R}} L_n(\mathbf{x}, \theta)} = \frac{L_n(\mathbf{x}, \theta_0)}{L_n(\mathbf{x}, \hat{\theta}_n)}$$

car le sup est atteint pour $\theta = \hat{\theta}_n$. On trouve que le test de niveau α du rapport de vraisemblances rejette H_0 si

$$\frac{\bar{n} |X_{(n)} - \theta_0|}{\sigma} > Z_{1-\alpha/2}$$

Exemple 19. Dans le cas où $X \sim N(\theta, \sigma^2)$ avec θ et σ inconnus et $H_0: \theta = \theta_0$ contre $H_1: \theta \neq \theta_0$ on trouve que le test de niveau α du rapport de vraisemblances rejette H_0 si

$$\frac{\bar{n} |X_{(n)} - \theta_0|}{S_n} > t_{1-\alpha/2, n-1}$$

où $t_{\alpha, n}$ est le quantile α de la loi t de Student avec n degrés de liberté et S_n^2 est la variance empirique

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