Tests d'hypothèses

L'objectif d'un test d'hypothèse paramétrique est de répondre à la question que l'on forme de la manière suivante: Est-ce qu'au vu de l'observation d'un échantillon $\boldsymbol{X} = (X_1, ..., X_n)$, les X_j sont des v.a. iid, on peut décider entre les deux possibilités H_0: $\theta \in \Theta_0$ et H_1: $\theta \in \Theta_1$? Ici $\Theta_i \subset \Theta$ pour $i = 0, 1$ et $\Theta_0 \cap \Theta_1 = \emptyset$.

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-à-d que l'on doit s'attendre à: $\theta = 1/2$. ($X \sim \mathrm{Ber}(\theta), \theta \in]0, 1[$)

Au vu de l'observation d'un échantillon (X_1, ..., X_n) où les X_j sont iid $\sim \mathrm{Ber}(\theta)$ on pourrait tester les hypothèses suivantes:

- H_0: $\theta = 1/2$ contre H_1: $\theta = 1/3$
- H_0: $\theta = 1/2$ contre H_1: $\theta = 2/3$
- H_0: $\theta = 1/2$ contre H_1: $\theta < 1/2$
- H_0: $\theta = 1/2$ contre H_1: $\theta > 1/2$
- 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 $X_1, ..., X_n$ sont des v.a. iid $\sim \mathcal{N}(\mu, \sigma^2)$. On va tester

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

H_0 représente l'éventualité que le médicament n'a aucun effet, H_1 celle de l'événement que en moyenne la pression artérielle baisse.

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

On appelle test d'hypothèse ne règle de décision qui permet de décider entre les hypothèses H_0: $\theta \in \Theta_0$ et H_1: $\theta \in \Theta_1$. H_0 est appelée l'hypothèse nulle et H_1 l'hypothèse alternative.

Un test est une statistique $\phi(X_1, ..., X_n)$ à valeurs dans $\{0, 1\}$:

$$\phi(X_1, ..., X_n) = \begin{cases} 1 & \text{si on rejette } H_0 \text{ (et donc accepte } H_1) \\ 0 & \text{si on accepte } H_0 \end{cases}$$

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

$$W = \{\boldsymbol{x} = (x_1, ..., x_n) \text{ tel que } \phi(x_1, ..., x_n) = 1\} \subseteq \mathcal{X}^n.$$

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

$$\mathbb{P}_\theta(\boldsymbol{X} \in W) = \mathbb{P}_\theta(\text{rejeter } H_0) \quad \text{pour } \theta \in \Theta_0.$$

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

$$\mathbb{P}_\theta(\boldsymbol{X} \notin W) = \mathbb{P}_\theta(\text{accepter } H_0) \quad \text{pour } \theta \in \Theta_1.$$

On appelle *puissance du test* la fonction $\beta\colon \Theta_1 \to [0,1]$ donnée par $\beta(\theta) = \mathbb{P}_\theta(\text{rejeter } H_0)$ pour $\theta \in \Theta_1$.

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

On dira qu'un test est de niveau α (où $\alpha \in]0,1[$ fixé) si $\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\boldsymbol{X} \in W) = \alpha$.

On dira qu'un test est de seuil α (où $\alpha \in]0,1[$ fixé) si $\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\boldsymbol{X} \in W) \leqslant \alpha$.

On dira qu'un test de seuil α est uniformément le plus puissant (UPP) si sa puissance est maximale pour tout $\theta \in \Theta_1$ parmi tous les test de seuil α.

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

Exemple 3. Soit $(X_1, ..., X_n)$ un échantillon issu de la loi $\mathcal{N}(\mu, \sigma^2)$ où σ_0 est connu. On considère le problème de comparer

$$H_0\colon \mu = \mu_0 \quad \text{contre} \quad H_1\colon \mu = \mu_1$$

avec $\mu_0 < \mu_1$ fixés.

Un estimateur naturel de l'espérance μ 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

$$W = \{(x_1, ..., x_n) : \bar{x}_n > k\}$$

et donc il accepte H_0 si $\bar{X}_n \leqslant k$ et rejette H_0 si $\bar{X}_n > k$. Autrement dit $\phi(X_1, ..., X_n) = \mathbb{I}_{\bar{X}_n > k}$. Il nous reste de déterminer k.

Si on veut construire un test de niveau α on doit avoir $\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\bar{X}_n > k) = \alpha$. Ici $\Theta_0 = \{\mu_0\}$ et donc

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\bar{X}_n > k) = \mathbb{P}_{\mu_0}(\bar{X}_n > k) = \alpha.$$

Sous l'hypothèse nulle $\bar{X}_n \sim \mathcal{N}(\mu_0, \sigma_0^2/n)$:

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

$$\alpha = \mathbb{P}_{\mu_0}(\bar{X}_n > k) = \mathbb{P}_{\mu_0}\left(\frac{\bar{X}_n - \mu_0}{\sigma_0/\sqrt{n}} > \frac{k - \mu_0}{\sigma_0/\sqrt{n}}\right) = \mathbb{P}\left(Z > \frac{k - \mu_0}{\sigma_0/\sqrt{n}}\right)$$

$$\Leftrightarrow \ \mathbb{P}\left(Z \leqslant \frac{k - \mu_0}{\sigma_0/\sqrt{n}}\right) = 1 - \alpha$$

On en déduit que

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

où z_α 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 $\beta(\mu_1)$ de ce test. Par définition

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

Sous l'hypothèse alternative

$$Z' = \frac{\bar{X}_n - \mu_1}{\sigma_0/\sqrt{n}} \sim \mathcal{N}(0,1)$$

et donc

$$\beta(\mu_1) = \mathbb{P}_{\mu_1}\left(\frac{\bar{X}_n - \mu_1}{\sigma_0/\sqrt{n}} > \frac{\mu_0-\mu_1}{\sigma_0/\sqrt{n}} + z_{1-\alpha}\right) = 1 - \mathbb{P}\left(Z' \leqslant \frac{\mu_0-\mu_1}{\sigma_0/\sqrt{n}} + z_{1-\alpha}\right)$$

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

$$\lim_{n\to\infty} \mathbb{P}\left(Z' \leqslant \frac{\mu_0-\mu_1}{\sigma_0/\sqrt{n}} + z_{1-\alpha}\right) = 0$$

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

Les différentes catégories d'hypothèses

Les hypothèses simples

Définition 4. *On dira qu'un test est un test d'hypothèses simples si les hypothèses sont du type* $H_0\colon \theta = \theta_0$ *et* $H_1\colon \theta = \theta_1$ *où* $\theta_0 \neq \theta_1$ *(*$\Theta_0 = \{\theta_0\}$ *et* $\Theta_1 = \{\theta_1\}$*).*

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

$$\frac{L_n(\boldsymbol{x},\theta_1)}{L_n(\boldsymbol{x},\theta_0)} > c_\alpha \qquad \text{où} \qquad \mathbb{P}_{\theta_0}\left(\frac{L_n(\boldsymbol{X},\theta_1)}{L_n(\boldsymbol{X},\theta_0)} > c_\alpha\right) = \alpha$$

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

Exemple 6. Reprenons l'exemple 3.

$$\frac{L_n(\boldsymbol{x},\mu_1)}{L_n(\boldsymbol{x},\mu_0)} = \frac{(2\pi\sigma_0^2)^{-n/2}\exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i-\mu_1)^2\right)}{(2\pi\sigma_0^2)^{-n/2}\exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i-\mu_0)^2\right)} = \exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n [(x_i-\mu_1)^2 - (x_i-\mu_0)^2]\right)$$

$$= \exp\left(\frac{(\mu_1-\mu_0)}{\sigma_0^2}\sum_{i=1}^n x_i - \frac{n}{2\sigma_0^2}(\mu_1^2-\mu_0^2)\right) = \exp\left(n\frac{(\mu_1-\mu_0)}{\sigma_0^2}\bar{x}_n - \frac{n}{2\sigma_0^2}(\mu_1^2-\mu_0^2)\right)$$

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

$$\frac{L_n(\boldsymbol{x},\mu_1)}{L_n(\boldsymbol{x},\mu_0)} = \exp\left(n\frac{(\mu_1-\mu_0)}{\sigma_0^2}\,\bar{x}_n - \frac{n}{2\sigma_0^2}(\mu_1^2-\mu_0^2)\right) > c_\alpha$$

$$\Leftrightarrow \quad \bar{x}_n > \frac{\sigma_0^2\log c_\alpha + \frac{n}{2}(\mu_1^2-\mu_0^2)}{(\mu_1-\mu_0)n} = k_\alpha$$

où k_α est déterminée par la condition

$$\mathbb{P}_{\mu_0}\left(\frac{L_n(\boldsymbol{X},\mu_1)}{L_n(\boldsymbol{X},\mu_0)} > c_\alpha\right) = \alpha = \mathbb{P}_{\mu_0}(\bar{X}_n > k_\alpha).$$

Lien avec l'exhaustivité

Proposition 7. *On considère les hypothèses simples* $H_0\colon \theta=\theta_0$ *et* $H_1\colon \theta=\theta_1$ $(\theta_0\neq\theta_1)$. *Si* $S(\boldsymbol{X})$ *est exhaustive pour* θ *et*

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

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

$$\frac{g(S(\boldsymbol{x}),\theta_1)}{g(S(\boldsymbol{x}),\theta_0)} > d_\alpha$$

avec d_α *déterminée par la condition*

$$\mathbb{P}_{\theta_0}\left(\frac{g(S(\boldsymbol{X}),\theta_1)}{g(S(\boldsymbol{X}),\theta_0)} > d_\alpha\right) = \alpha.$$

Exemple 8. Reprenons encore l'exemple 3. On sait que $S(\boldsymbol{X}) = \bar{X}_n$ est exhaustive pour μ et que

$$g(s,\mu) = f_{S(\boldsymbol{X})}(s,\mu) = \frac{1}{(2\pi\sigma_0^2)^{1/2}} e^{-\frac{n}{2\sigma_0^2}(s-\mu)^2}.$$

Donc

$$\frac{g(s,\theta_1)}{g(s,\theta_0)} = e^{-\frac{n}{2\sigma_0^2}(s-\mu_1)^2 - \frac{n}{2\sigma_0^2}(s-\mu_0)^2} > d_\alpha$$

$$\Leftrightarrow -\frac{n}{2\,\sigma_0^2}(s-\mu_1)^2 - \frac{n}{2\,\sigma_0^2}(s-\mu_0)^2 > \log d_\alpha$$

$$\Leftrightarrow \quad s > \frac{1}{2}\left(\mu_0+\mu_1+\frac{2\,\sigma_0^2}{n(\mu_0-\mu_1)}\log d_\alpha\right) = k_\alpha.$$

Le test rejette H_0 si $S(\boldsymbol{X}) = \bar{X}_n > k_\alpha$ et donc on retrouve que

$$k_\alpha = \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_{1-\alpha}.$$

Hypothèse simple contre hypothèse composite

Définition 9. *On appelle test entre hypothèse simple et hypothèse composite tout test ou les hypothèses sont du type* $H_0: \theta = \theta_0$ *et* $H_1: \theta \in \Theta_1$ *où* $\theta_0 \notin \Theta_1$ *et* $\mathrm{card}(\Theta_1) > 1$.

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 dépends 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 test de seuil* α.

Exemple 11. Reprenons l'exemple 3. On considère les hypothèses H_0: $\mu = \mu_0$ contre H_1: $\mu > \mu_0$. On a vu que le test de Neyman et Pearson de niveau α pour les hypothèses simples H_0: $\mu = \mu_0$ contre H_1: $\mu = \mu_1$ où $\mu_1 > \mu_0$ est de rejeter H_0 si

$$\bar{X}_n > \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_{1-\alpha} .$$

La région critique de ce test est

$$W = \{\boldsymbol{x} : \bar{x}_n > \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_{1-\alpha}\}$$

et ne dépends pas de μ_1. D'après la proposition précédente, le test de Neyman et Pearson est le test UPP pour les hypothèses H_0: $\mu = \mu_0$ contre H_1: $\mu > \mu_0$ parmi toues les test de seuil α.

Remarque 12. Le test UPP de niveau α pour les hypothèses H_0: $\mu = \mu_0$ contre H_1: $\mu < \mu_0$ est la règle de décision qui rejette H_0 si

$$\bar{X}_n < \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_\alpha = \mu_0 - \frac{\sigma_0}{\sqrt{n}} z_{1-\alpha} .$$

Remarque 13. (Exercice) Il n'existe pas un test UPP pour les hypothèses H_0: $\mu = \mu_0$ contre H_1: $\mu \neq \mu_0$.

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* $H_1: \theta \in \Theta_1$ *où* $\Theta_0 \cap \Theta_1 = \emptyset$, $\mathrm{card}(\Theta_0) > 1$ *et* $\mathrm{card}(\Theta_1) > 1$.

Théorème 15. *(de Karlin-Rubin) On considère les hypothèses* H_0: $\theta \leqslant \theta_0$ *contre* H_1: $\theta > \theta_0$. *Si* $S(\boldsymbol{X})$ *est une statistique exhaustive pour* θ *telle que, pour tout* $\theta_2 > \theta_1$ *la fonction*

$$s \mapsto \frac{g(s, \theta_2)}{g(s, \theta_1)}$$

est croissante, alors le test qui rejette H_0 *si* $S(\boldsymbol{X}) > s_\alpha$ *est le test UPP parmi tous les test de seuil* $\alpha = \mathbb{P}_{\theta_0}(S(\boldsymbol{X}) > s_\alpha)$.

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

$$\frac{g(s, \mu_2)}{g(s, \mu_1)} = e^{-\frac{n}{2}(s-\mu_2)^2 + \frac{n}{2}(s-\mu_1)^2} = e^{\frac{n}{2}(2s(\mu_2-\mu_1) - \mu_2^2 + \mu_1^2)}$$

est une fonction croissante de s pour tout $\mu_2 > \mu_1$. D'après le théorème de Karlin-Rubin, le test qui rejette H_0 si $S(\boldsymbol{X}) > s_\alpha$ est le test UPP parmi tous les tests de seuil $\alpha = \mathbb{P}_{\theta_0}(\bar{X}_n > s_\alpha)$ et donc

$$s_\alpha = \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_{1-\alpha}\,.$$

Test du rapport de vraisemblances

On considère en général les hypothèses H_0: $\theta \in \Theta_0$ contre H_1: $\theta \in \Theta_1$. On pose

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

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

$$\lambda_n(\boldsymbol{X}) \leqslant \lambda_\alpha\,.$$

Exemple 18. Reprenons l'exemple 3 avec H_0: $\mu = \mu_0$ contre H_1: $\mu \neq \mu_0$. Ici $\Theta_0 = \{\mu_0\}$ et $\Theta_1 = \mathbb{R} \backslash \Theta_0$. Donc $\Theta_1 \cup \Theta_0 = \mathbb{R}$.

$$\lambda_n(\boldsymbol{x}) = \frac{L_n(\boldsymbol{x}, \mu_0)}{\sup_{\mu \in \mathbb{R}} L_n(\boldsymbol{x}, \mu)} = \frac{L_n(\boldsymbol{x}, \mu_0)}{L_n(\boldsymbol{x}, \bar{x}_n)}$$

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

$$\frac{\sqrt{n}|\bar{X}_n - \mu_0|}{\sigma_0} > z_{1-\alpha/2}$$

Exemple 19. Dans le cas où $X \sim \mathcal{N}(\mu, \sigma^2)$ avec μ et σ inconnus et H_0: $\mu = \mu_0$ contre H_1: $\mu \neq \mu_0$. on trouve que le test de niveau α du rapport de vraisemblances rejette H_0 si

$$\frac{\sqrt{n}|\bar{X}_n - \mu_0|}{S_n} > t_{1-\alpha/2, n-1}$$

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

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