M1 – Statistique – 2020/2021
vincent.brault@univ-grenoble-alpes.fr

TD n°5 : Tests

Exemples du cours

Intuition

Étant donné un n-échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ iid de loi $\mathcal{N}(\theta^\star, 1)$ avec $\theta^\star \in \mathbb{R}$, nous voulons tester :

$$\begin{cases} \mathcal{H}_0 : & \theta^\star \geq 1, \\ \mathcal{H}_1 : & \theta^\star < 1. \end{cases}$$

1. Donner les ensembles Θ_0 et Θ_1 correspondants.

2. Nous prenons la variable $\varphi(X) = \mathbb{1}_{\{\overline{X}_n < 1\}}$. Est-ce une statistique de test ? Si oui, donner la région de rejet. Sinon, expliquer pourquoi et en proposer une nouvelle.

3. Calculer les fonctions $\underline{\alpha}$ et $\underline{\beta}$ en fonction de la fonction de répartition $F_{\mathcal{N}(0,1)}$.

4. Calculer la taille du test. Qu'en déduisez-vous ?

Construction d'un test

Étant donné un n-échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ iid de loi $\mathcal{N}(\theta^\star, 1)$ avec $\theta^\star \in \mathbb{R}$, nous voulons tester :

$$\begin{cases} \mathcal{H}_0 : & \theta^\star \geq 1, \\ \mathcal{H}_1 : & \theta^\star < 1. \end{cases}$$

1. Quelle forme de région de rejet vous paraît la plus logique ? $]-\infty, k_\alpha[$? $]k_\alpha, \ell_\alpha[$? $]k_\alpha, +\infty[$? Une autre forme ?

2. En déduire la statistique ϕ.

3. Calculer la taille $\alpha^\star$ en fonction de k_α.

4. En déduire la forme de la statistique de test.

Test du rapport de vraisemblance

Étant donné un n-échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ iid de loi $\mathcal{N}(\theta^\star, 1)$ avec $\theta^\star \in \mathbb{R}$, nous voulons tester :

$$\begin{cases} \mathcal{H}_0 : & \theta^\star \geq 1, \\ \mathcal{H}_1 : & \theta^\star < 1. \end{cases}$$

1. Calculer la vraisemblance du modèle et étudier son sens de variation en fonction de $\theta^\star$.

2. Montrer que le rapport de vraisemblance peut s'écrire de la forme $h(\mathbf{X}) = g(\overline{X}_n)$ où g est définie sur $\mathbb{R}$ par :

$$g(x) = \exp\left[\frac{n}{2}(x-1)^2\left(2\mathbb{1}_{\{x<1\}} - 1\right)\right].$$

3. Étudier le sens de variation de la fonction g.

4. En déduire que le test du rapport de vraisemblance peut s'écrire comme dans l'exercice précédent.

Exercice 1

Étant donné un n-échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ iid de loi $\mathcal{E}(\theta^\star)$ avec $\theta^\star \in \mathbb{R}_+^\star$, nous voulons tester :

$$\begin{cases} \mathcal{H}_0 : & \theta^\star \leq 1, \\ \mathcal{H}_1 : & \theta^\star > 1. \end{cases}$$

1. Calculer l'estimateur des moments.

Pour calculer l'estimateur des moments, nous commençons par calculer l'espérance :

$$\begin{aligned}
\mathbb{E}\left[X_1\right] &= \int_0^{+\infty} x\theta^\star e^{-x\theta^\star} \mathrm{d}x \\
& \quad u(x) = x,\ u'(x) = 1, \\
& \quad v'(x) = \theta^\star e^{-x\theta^\star} \text{ et } v(x) = -e^{-x\theta^\star}, \\
&= \left[-xe^{-x\theta^\star}\right]_0^{+\infty} + \int_0^{+\infty} e^{-x\theta^\star} \mathrm{d}x \\
&= \int_0^{+\infty} e^{-x\theta^\star} \mathrm{d}x \\
&= \left[-\frac{1}{\theta^\star} e^{-x\theta^\star}\right]_0^{+_infty} \\
&= \frac{1}{\theta^\star}.
\end{aligned}$$

La méthode des moments consiste à estimer le paramètre en estimant les moments par les estimateurs empiriques associés donc nous avons :

$$\overline{X}_n = \frac{1}{\widehat{\theta}_n} \Leftrightarrow \widehat{\theta}_n = \frac{1}{\overline{X}_n}.$$

2. Calculer la loi asymptotique de l'estimateur.

Par des calculs identiques, nous avons la variance de X_1 qui vaut $1/\theta^{\star 2}$ (je me permets de ne pas refaire le calcul mais il faudrait dans un examen). Donc, par le théorème de la limite centrale, nous avons :

$$\sqrt{n}\left(\overline{X}_n - 1/\theta^\star\right) \underset{n \to +\infty}{\longrightarrow} \mathcal{N}\left(0, 1/\theta^{\star 2}\right).$$

Nous observons que :

- $a_n = \sqrt{n} \to +\infty$,
- Nous avons $a_n \left(U - U*\right) \underset{n \to +\infty}{\longrightarrow} V$,
- La fonction $\ell(x) = 1/x$ qui est différentiable en $1/\theta^\star$ car $\theta^\star > 0$.

Donc, nous pouvons appliquer la méthode Delta. Nous commençons par calculer la défier de la fonction ℓ :

$$\ell'(x) = -\frac{1}{x^2}$$

que nous calculons au point $1\theta^\star$:

$$\ell'\left(\frac{1}{\theta^\star}\right) = -\theta^{\star 2}.$$

Ainsi, nous avons :

$$\begin{aligned}
\sqrt{n}\left(\widehat{\theta}_n - \theta^\star\right) \quad & \underset{n\to+\infty}{\longrightarrow} \quad \ell'\left(\frac{1}{\theta^\star}\right)\mathcal{N}\left(0, \frac{1}{{\theta^\star}^2}\right) \\
& \underset{n\to+\infty}{\longrightarrow} \quad -{\theta^\star}^2\mathcal{N}\left(0, \frac{1}{{\theta^\star}^2}\right) \\
& \underset{n\to+\infty}{\longrightarrow} \quad \mathcal{N}\left(0, \frac{(-1)^2{\theta^\star}^4}{{\theta^\star}^2}\right) \\
& \underset{n\to+\infty}{\longrightarrow} \quad \mathcal{N}\left(0, {\theta^\star}^2\right).
\end{aligned}$$

3. En déduire un test de taille asymptotique α.

À partir de la question précédente, nous obtenons que :

$$\begin{aligned}
\frac{\sqrt{n}}{\sqrt{{\theta^\star}^2}}\left(\widehat{\theta}_n - \theta^\star\right) \quad & = \quad \frac{\sqrt{n}}{\theta^\star}\left(\widehat{\theta}_n - \theta^\star\right) \\
& = \quad \sqrt{n}\left(\frac{\widehat{\theta}_n}{\theta^\star} - \frac{\theta^\star}{\theta^\star}\right) \\
& = \quad \sqrt{n}\left(\frac{\widehat{\theta}_n}{\theta^\star} - 1\right) \\
& \underset{n\to+\infty}{\longrightarrow} \quad \mathcal{N}\left(0, 1\right).
\end{aligned}$$

Donc, en notant $F_{\mathcal{N}(0,1)}$ la fonction de répartition d'une loi gaussienne centrée réduite et nous cherchons k_α de telle sorte que :

$$\begin{aligned}
\alpha_n^\star \quad & = \quad \sup_{\theta^\star\in\Theta_0} \mathbb{P}_{\theta^\star}\left(\widehat{\theta}_n \geq k_\alpha\right) \\
& = \quad \sup_{\theta^\star\leq 1} \mathbb{P}_{\theta^\star}\left(\sqrt{n}\left[\frac{\widehat{\theta}_n}{\theta^\star} - 1\right] \geq \sqrt{n}\left[\frac{k_\alpha}{\theta^\star} - 1\right]\right) \\
& = \quad \sup_{\theta^\star\leq 1} \left[1 - \mathbb{P}_{\theta^\star}\left(\sqrt{n}\left[\frac{\widehat{\theta}_n}{\theta^\star} - 1\right] < \sqrt{n}\left[\frac{k_\alpha}{\theta^\star} - 1\right]\right)\right] \\
& \approx \quad \sup_{\theta^\star\leq 1} \left(1 - F_{\mathcal{N}(0,1)}\left[\sqrt{n}\left(\frac{k_\alpha}{\theta^\star} - 1\right)\right]\right) \text{ pour } n \text{ assez grand,} \\
& \approx \quad 1 - F_{\mathcal{N}(0,1)}\left[\sqrt{n}\left(\frac{k_\alpha}{1} - 1\right)\right] \text{ car } \theta \mapsto F_{\mathcal{N}(0,1)}\left[\sqrt{n}\left(\frac{k_\alpha}{\theta} - 1\right)\right] \text{ est décroissante,} \\
& \approx \quad 1 - F_{\mathcal{N}(0,1)}\left[\sqrt{n}\left(k_\alpha - 1\right)\right].
\end{aligned}$$

Ainsi, nous avons pour tout $\alpha \in]0;1[$ et pour n assez grand :

$$\begin{aligned}
\alpha \approx \alpha_n^\star \quad & \Leftrightarrow \quad \alpha \approx 1 - F_{\mathcal{N}(0,1)}\left[\sqrt{n}\,(k_\alpha - 1)\right] \\
& \Leftrightarrow \quad 1-\alpha \approx F_{\mathcal{N}(0,1)}\left[\sqrt{n}\,(k_\alpha - 1)\right] \\
& \Leftrightarrow \quad q_{1-\alpha} \approx \sqrt{n}\,(k_\alpha - 1) \\
& \Leftrightarrow \quad k_\alpha \approx \frac{q_{1-\alpha}}{\sqrt{n}} + 1.
\end{aligned}$$

Donc si nous prenons la suite de test $\Phi_n = \mathbb{1}_{\left\{\widehat{\theta}_n \geq 1 + \frac{q_{1-\alpha}}{\sqrt{n}}\right\}} = \mathbb{1}_{\left\{\overline{X}_n \leq \frac{\sqrt{n}}{\sqrt{n}+q_{1-\alpha}}\right\}}$ est de niveau asymptotique α.

4. Calculer le test de rapport de vraisemblance. Qu'en déduisez-vous ?

Pour le test du rapport de vraisemblance, nous rappelons que la vraisemblance est :

$$\begin{aligned}
V_{\mathbf{X}}(\theta) &= \prod_{i=1}^{n}\left[\theta e^{-\theta X_i}\right] \\
&= \theta^n e^{-\theta \sum_{i=1}^n X_i}.
\end{aligned}$$

nous avons déjà vu que le maximum de vraisemblance est atteint pour $\theta = 1/\overline{X}_n$ pour une valeur :

$$\begin{aligned}
V_{\mathbf{X}}\left(\frac{1}{\overline{X}_n}\right) &= \left(\frac{1}{\overline{X}_n}\right)^n e^{-\frac{1}{\overline{X}_n}\sum_{i=1}^n X_i} \\
&= \frac{1}{\overline{X}_n^n} e^{-\frac{1}{\frac{1}{n}\sum_{i=1}^n X_i}\sum_{i=1}^n X_i} \\
&= \frac{1}{\overline{X}_n^n} e^{-n}
\end{aligned}$$

et nous avons le tableau suivant :

θ	0	$1/\overline{X}_n$	$+\infty$
$V_{\mathbf{X}}$	0 $\nearrow$	$\frac{1}{\overline{X}_n^n}e^{-n}$	$\searrow$ 0

Ainsi, la question se pose de savoir si $1/\overline{X}_n$ est plus petit ou plus grand que 1. Dans le cas où

$1 \leq 1/\overline{X}_n$, nous avons :

$$\begin{aligned}
h(X) &= \frac{\sup_{\theta \in \Theta_1} V_{\mathbf{X}}(\theta)}{\sup_{\theta \in \Theta_0} V_{\mathbf{X}}(\theta)} \\
&= \frac{\sup_{\theta > 1} V_{\mathbf{X}}(\theta)}{\sup_{\theta \leq 1} V_{\mathbf{X}}(\theta)} \\
&= \frac{V_{\mathbf{X}}\left(\frac{1}{\overline{X}_n}\right)}{V_{\mathbf{X}}(1)} \\
&= \frac{\frac{1}{\overline{X}_n^n} e^{-n}}{1^n e^{-\sum_{i=1}^n X_i}} \\
&= \frac{1}{\overline{X}_n^n} e^{-n} e^{\sum_{i=1}^n X_i} \\
&= e^{-n\left(\ln \overline{X}_n + 1 - \overline{X}_n\right)} \\
&= e^{-n g_n\left(\overline{X}_n\right)}.
\end{aligned}$$

De même, si $1 \geq 1/\overline{X}_n$, nous avons $h(X) = e^{n g_n\left(\overline{X}_n\right)}$ avec $g(x) = \ln(x) + 1 - x$. Du coup, nous obtenons :

$$\begin{aligned}
h(X) &= \begin{cases} e^{-n g_n\left(\overline{X}_n\right)} & \text{si } 1/\overline{X}_n \geq 1, \\ e^{n g_n\left(\overline{X}_n\right)} & \text{sinon,} \end{cases} \\
&= \begin{cases} e^{-n g_n\left(\overline{X}_n\right)} & \text{si } \overline{X}_n \leq 1, \\ e^{n g_n\left(\overline{X}_n\right)} & \text{sinon,} \end{cases} \\
&= h_n\left(\overline{X}_n\right).
\end{aligned}$$

Or, la fonction $x \mapsto g_n(x)$ est strictement croissante sur $]0;1[$ et strictement décroissante sur $]1;+\infty[$. Du coup, la fonction $x \mapsto h_n(x)$ est strictement décroissante sur $]0;+\infty[$. Du coup, il existe k_α tel que le test est donc :

$$\begin{aligned}
\mathbb{1}_{\{h(X) > k_\alpha\}} &= \mathbb{1}_{\{h_n\left(\overline{X}_n\right) > k_\alpha\}} \\
&= \mathbb{1}_{\{\overline{X}_n > h_n^{(-1)}(k_\alpha)\}} \\
&= \mathbb{1}_{\{\overline{X}_n > k'_\alpha\}}.
\end{aligned}$$

Le test est donc de la même forme que celui calculé précédemment et la valeur $k'_\alpha = \frac{\sqrt{n}}{\sqrt{n} + q_{1-\alpha}}$ de telle sorte que le niveau asymptotique soit α.

Exercice 2

Étant donné un n-échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ iid de loi $\mathcal{G}eo(\theta^\star)$ avec $\theta^\star \in]0;1[$, nous voulons tester :

$$\begin{cases} \mathcal{H}_0 : & \theta^\star = 1/2, \\ \mathcal{H}_1 : & \theta^\star = 3/4. \end{cases}$$

Proposez un test uniformément puissant de niveau α.

Exercice 3

Nous prenons dans cet exercice un n-échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ iid de loi $\mathcal{B}in(n, \theta^\star)$ avec $\theta^\star \in]0;1[$.

1. Soit $\theta_0, \theta_1 \in]0;1[$ avec $\theta_1 > \theta_0$, construire le test de rapport de vraisemblance de "$\theta^\star = \theta_0$" contre "$\theta^\star = \theta_1$" au niveau α.

2. Application numérique : $n = 10$, $\theta_0 = 0.3$ et $\alpha = 0.05$, quel est le niveau réel de ce test ?

3. Le test précédent à une zone de rejet de la forme $S > s_0$. Nous définissons un nouveau test à l'aide de la statistique T suivante :

$$T(S) = \begin{cases} 1 & si \quad S > s_0, \\ 0 & si \quad S < s_0, \end{cases}$$

et si $S = s_0$, T(S) est aléatoire, valant 1 avec probabilité γ et 0 sinon. Quel est le niveau de ce nouveau test ?

4. Application numérique : trouver γ lorsque $n = 10$, $p_0 = 0.3$ pour avoir un niveau réel de $\alpha = 0.05$.