Statistiques (STA1)

Cours III – Théorèmes asymptotiques, théorème de Neyman-Pearson

Luca Ganassali

Laboratoire de Mathématiques d'Orsay, Université Paris-Saclay

Jeudi 2 octobre 2025

Previously in STA1...

Si les $(X_1, \ldots, X_n)$ sont i.i.d. de loi $\mathrm{Exp}(\lambda)$ (densité $f(x) = \lambda e^{-\lambda x}, x \geq 0$), quel est l'estimateur du maximum de vraisemblance de λ ?

La vraisemblance s'écrit

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \exp\Big(- \lambda \sum_{i=1}^{n} x_i \Big).$$

La log-vraisemblance est

$$\ell(\lambda) = \ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i.$$

En dérivant et en annulant :

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \quad \Longrightarrow \quad \hat{\lambda}_{\mathrm{MLE}} = \frac{n}{\sum_{i=1}^{n} x_i} = 1/\bar{X}.$$

De plus $\ell''(\lambda) = -n/\lambda^2 < 0$, donc ℓ est concave : il s'agit bien d'un maximum.

Théorèmes asymptotiques – Lemme de Slutsky

Lemme de Slutsky : exemple

Exemple : $X_1, \ldots, X_n$ i.i.d., moyenne $\mu \in \mathbb{R}$ inconnue, variance finie $\sigma^2 > 0$ inconnue. On veut un IC asymptotique pour μ.

Méthode pivotale. Théorème central limite :

$$\sqrt{n}\frac{\bar{X} - \mu}{\sigma} \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,1)\,.$$

On a donc

$$\mathbb{P}\left(\mu \in \left[\bar{X} \pm \frac{\sigma q_{1-\alpha/2}}{\sqrt{n}}\right]\right) = \mathbb{P}\left(\sqrt{n}\frac{\bar{X}-\mu}{\sigma} \in [\pm q_{1-\alpha/2}]\right)$$
$$\xrightarrow[n\to\infty]{} \mathbb{P}\left(\mathcal{N}(0,1) \in [\pm q_{1-\alpha/2}]\right) = 1 - \alpha\,.$$

IC asymptotique de proba de couverture $1-\alpha$: $\left[\bar{X} \pm \frac{\sigma q_{1-\alpha/2}}{\sqrt{n}}\right]$.

Problème. Cet "IC" dépend de σ qui est inconnu ! $\to$ ce n'est pas un IC.

Lemme de Slutsky : exemple

Idée. Remplacer σ par un estimateur $\to \sqrt{\hat{\sigma}^2}$ ($\hat{\sigma}^2$ estimateur non biaisé de la variance). On a que

$$\sqrt{n}\frac{\bar{X}-\mu}{\sqrt{\hat{\sigma}^2}} = \underbrace{\sqrt{n}\frac{\bar{X}-\mu}{\sigma}}_{\xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,1)} \times \underbrace{\frac{\sigma}{\sqrt{\hat{\sigma}^2}}}_{\xrightarrow[n\to\infty]{p.s.} 1}$$

Question : a-t-on le droit de multiplier des limites en loi ?

Réponse : Non, en général. Prendre $X \sim \mathcal{N}(0,1)$ et $Y = X$, $XY = X^2 \geq 0$ p.s., mais le produit de deux $\mathcal{N}(0,1)$ indépendantes n'est pas positif p.s. !

Mais ici on peut le faire, c'est le Lemme de Slutsky qui permet de l'établir !

Lemme de Slutsky : énoncé

Lemme de Slutsky. Si on a les convergences en loi $X_n \underset{n\to\infty}{\longrightarrow} X$ et $Y_n \underset{n\to\infty}{\longrightarrow} c$ où c est constante, alors $(X_n, Y_n) \underset{n\to\infty}{\longrightarrow} (X, c)$ en tant que couple, et en particulier, $X_n + Y_n \underset{n\to\infty}{\longrightarrow} X + c$, $X_nY_n \underset{n\to\infty}{\longrightarrow} cX$ et si c est réel et $c \neq 0$ alors $X_n/Y_n \underset{n\to\infty}{\longrightarrow} X/c$.

(Lemme admis)

Dans notre cas, $\hat{\sigma}^2 \to \sigma^2$ p.s. par la LGN donc $\sqrt{\hat{\sigma}^2}/\sigma \to 1$ p.s. par continuité et donc en loi aussi. La limite est une constante.

D'autre part $\sqrt{n}\frac{\bar{X}-\mu}{\sigma} \underset{n\to\infty}{\xrightarrow{(d)}} \mathcal{N}(0,1)$ par le TCL.

Par le Lemme de Slutsky,

$$\sqrt{n}\frac{\bar{X}-\mu}{\sqrt{\hat{\sigma}^2}} = \underbrace{\sqrt{n}\frac{\bar{X}-\mu}{\sigma}}_{\underset{n\to\infty}{\xrightarrow{(d)}} \mathcal{N}(0,1)} \times \underbrace{\frac{\sigma}{\sqrt{\hat{\sigma}^2}}}_{\underset{n\to\infty}{\xrightarrow{p.s.}} 1} \to \mathcal{N}(0,1)\,.$$

L'IC asymptotique devient $\left[\bar{X} \pm \frac{\sqrt{\hat{\sigma}^2}q_{1-\alpha/2}}{\sqrt{n}}\right] \to$ ne dépend plus de paramètres inconnus !

Théorèmes asymptotiques – Méthode Delta

Méthode Delta : exemple

Exemple : $X_1, \ldots, X_n$ i.i.d. $\mathrm{Exp}(\lambda)$ (densité $x \to \mathbf{1}_{x \geq 0} \lambda e^{-\lambda x}$), $\lambda > 0$ inconnu.

On a que (LGN) $\bar{X} \to \mathbb{E}[X_1] = 1/\lambda$, on pose donc $\hat{\lambda} = 1/\bar{X}$. On a par le TCL que

$$\sqrt{n}(\bar{X} - 1/\lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0, \mathrm{Var}(X_1) = 1/\lambda^2)\,.$$

Question :

$$\sqrt{n}(\hat{\lambda} - \lambda) = \sqrt{n}(1/\bar{X} - \lambda) \xrightarrow[n\to\infty]{(d)} ?$$

$\bar{X}$ est asymptotiquement normal. Est-ce que $1/\bar{X}$ aussi... ?

Problème. $\frac{1}{\text{Gaussienne}}$ c'est pas super Gaussien...

Idée. Le TCL c'est un DL d'ordre 1 ! Donc avec Taylor on pourrait avoir quelque chose ! C'est la méthode Delta.

Méthode Delta (cas réel). Soient $(X_n)_{n\geq 1}$ des variables aléatoires et $(v_n)_{n\geq 1}$ une suite de nombres réels positifs telle que $v_n \to +\infty$. On suppose qu'il existe un réel a et une variable aléatoire X tels que

$$v_n(X_n - a) \xrightarrow[n\to\infty]{(d)} X.$$

Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction différentiable en a. Alors

$$v_n\big(g(X_n) - g(a)\big) \xrightarrow[n\to\infty]{(d)} g'(a) \times X.$$

Méthode Delta : preuve

Puisque g est différentiable en a, on peut écrire un développement de Taylor de $g(x)$ en $x = a$:

$$g(x) = g(a) + (x-a)g'(a) + (x-a)\varepsilon(x),$$

avec ε continue sur $\mathbb{R} \setminus \{a\}$ et $\varepsilon(x) \underset{x\to a}{\longrightarrow} 0$. On peut alors prolonger ε par continuité en a avec $\varepsilon(a) = 0$.

Comme $X_n \xrightarrow[n\to\infty]{(d)} a$ en loi par le Lemme de Slutsky, et donc en probabilité, par continuité on a $\varepsilon(X_n) \xrightarrow[n\to\infty]{(d)} \varepsilon(a) = 0$. Ainsi, pour tout n,

$$g(X_n) = g(a) + (X_n - a)\big(g'(a) + \varepsilon(X_n)\big),$$

avec $g'(a) + \varepsilon(X_n) \xrightarrow[n\to\infty]{(d)} g'(a)$ (constante). On applique encore le lemme de Slutsky pour conclure :

$$v_n\big(g(X_n) - g(a)\big) = \big(g'(a) + \varepsilon(X_n)\big)\, v_n(X_n - a) \xrightarrow[n\to\infty]{(d)} g'(a) \times X.$$

$\square$

On avait

$$\sqrt{n}(\bar{X} - 1/\lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0, \mathrm{Var}(X_1) = 1/\lambda^2)\,.$$

En on veut étudier $g(\bar{X}) = 1/\bar{X}$. g est différentiable, $g'(x) = -1/x^2$, et donc d'après la méthode Delta

$$\sqrt{n}(\hat{\lambda}-\lambda) = \sqrt{n}(1/\bar{X}-\lambda) \xrightarrow[n\to\infty]{(d)} g'(1/\lambda)\times\mathcal{N}(0,1/\lambda^2) \overset{(d)}{=} -\lambda^2\times\mathcal{N}(0,1/\lambda^2) \overset{(d)}{=} \mathcal{N}(0,\lambda^2)\,.$$

Tests uniforméments plus puissants, Théorème de Neyman-Pearson

Tests uniforméments plus puissants

Rappel sur les tests. Pour un test, on contrôle l'erreur de première espèce (niveau) α, mais pas sa puissance β.

Question : A niveau α fixé, quel est le test ayant la meilleure puissance β ?

On se place dans un modèle paramétrique $(\mathcal{Z}, (\mathbb{P}_\theta)_{\theta \in \Theta})$. Un test de région de rejet $\mathcal{R}$ est uniformément plus puissant de niveau α (UPP(α)) pour tester $H_0 : \theta \in \Theta_0$ contre $H_1 : \theta \in \Theta_1$ si:

(*i*) il est de niveau α (pour tout $\theta \in \Theta_0$), i.e.

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\mathcal{R}) \leq \alpha.$$

(*ii*) Pour tout $\theta \in \Theta_1$, sa puissance $\pi(\theta) = \mathbb{P}_\theta(\mathcal{R})$ est supérieure à la puissance de tout autre test de niveau α.

On se place dans le cas où $H_0 : \theta = \theta_0$ et $H_1 : \theta = \theta_1$. On parle d'hypothèses simples. Dans ce cas, on sait construire des tests UPP(α): c'est le Théorème de Neyman-Pearson.

Théorème de Neyman-Pearson (cas de deux hyp. simples) On se place dans un modèle $(\mathcal{Z}, (\mathbb{P}_\theta)_{\theta\in\Theta})$ que l'on suppose dominé. On note L sa vraisemblance. Supposons qu'il existe $k_\alpha > 0$ tel que le test de région de rejet

$$\mathcal{R} = \left\{ z \in \mathcal{Z} : \frac{L(\theta_1; z)}{L(\theta_0; z)} > k_\alpha \right\}$$

soit de niveau α. Alors, ce test est UPP(α) pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$.

Preuve du Théorème de Neyman-Pearson (cas de deux hyp. simples)

Supposons que k_α existe. On note

$$\mathcal{R} = \left\{ z \in \mathcal{Z} : \frac{L(\theta_1; z)}{L(\theta_0; z)} > k_\alpha \right\}.$$

Soit $\mathcal{R}_\alpha$ une autre région de rejet de niveau α.

On note

$$A := \mathcal{R}_\alpha \setminus \mathcal{R} \quad \text{et} \quad B := \mathcal{R} \setminus \mathcal{R}_\alpha.$$

En général, pour tout θ, on a :

$$\begin{aligned}
&\mathbb{P}_\theta(Z \in \mathcal{R}_\alpha) - \mathbb{P}_\theta(Z \in \mathcal{R}) \\
&= (\mathbb{P}_\theta(\mathcal{R}_\alpha \setminus \mathcal{R}) + \mathbb{P}_\theta(\mathcal{R}_\alpha \cap \mathcal{R})) - (\mathbb{P}_\theta(\mathcal{R} \setminus \mathcal{R}_\alpha) + \mathbb{P}_\theta(\mathcal{R}_\alpha \cap \mathcal{R})) \\
&= \mathbb{P}_\theta(A) - \mathbb{P}_\theta(B).
\end{aligned}$$

Par ailleurs, comme $B \subseteq \mathcal{R} = \left\{z : \frac{L(\theta_1;z)}{L(\theta_0;z)} > k_\alpha\right\}$,

$$\mathbb{P}_{\theta_1}(B) = \int_B L(\theta_1, z)\, d\xi(z) \geq k_\alpha \int_B L(\theta_0, z)\, d\xi(z) = k_\alpha \mathbb{P}_{\theta_0}(B),$$

et de même, comme $A \subseteq \mathcal{R}^c$,

$$\mathbb{P}_{\theta_1}(A) = \int_A L(\theta_1, z)\, d\xi(z) \leq k_\alpha \int_A L(\theta_0, z)\, d\xi(z) = k_\alpha \mathbb{P}_{\theta_0}(A)$$

On en déduit :

$$\begin{aligned}
\mathbb{P}_{\theta_1}(Z \in \mathcal{R}_\alpha) - \mathbb{P}_{\theta_1}(Z \in \mathcal{R}) &= \mathbb{P}_{\theta_1}(A) - \mathbb{P}_{\theta_1}(B) \leq k_\alpha\big(\mathbb{P}_{\theta_0}(A) - \mathbb{P}_{\theta_0}(B)\big) \\
&= k_\alpha\big(\mathbb{P}_{\theta_0}(Z \in \mathcal{R}_\alpha) - \mathbb{P}_{\theta_0}(Z \in \mathcal{R})\big) \\
&= k_\alpha(\alpha - \alpha) = 0.
\end{aligned}$$

Ainsi, la puissance de $\mathcal{R}_\alpha$ est nécessairement inférieure ou égale à celle de $\mathcal{R}$. □

Question : existe-t-il toujours $k_\alpha > 0$ tel que

$$\mathcal{R} = \left\{ z \in \mathcal{Z} : \frac{L(\theta_1; z)}{L(\theta_0; z)} > k_\alpha \right\}$$

définisse un test de niveau α?

Cas continu : oui ! Théorème des valeurs intermédiaires avec l'application

$$h : \mathbb{R}_+ \to [0,1], \qquad k \longmapsto \mathbb{P}_{\theta_0}\left(\frac{L(\theta_1; z)}{L(\theta_0; z)} > k \right).$$

Cas continu : non. En général k_α n'existe pas pour tout α...

Merci !

Rdv en TD pour les questions et la pratique de ces notions.

(contenu du cours disponible sur ma page web : `lganassali.github.io`)