STATISTIQUE MATHÉMATIQUE - MS ENSAI
RÉSUMÉ CHAP.2 - 2025

1 Estimateurs: définitions et propriétés

Soit $(X_1,...,X_n)$ un n-échantillon de loi $\mathbb{P}_\theta$ inconnue. On cherche à estimer θ.

Définition

Un estimateur $S = S_n = S_n(X_1,...X_n)$ de θ est une statistique (fonction des seules données et pas du paramètre).

Définition

- $b(\theta) = \mathbb{E}[S] - \theta$ est le biais.
- $\mathbb{E}_\theta[(S-\theta)^2]$ est l'erreur quadratique (EQM).

Théorème

$\mathbb{E}_\theta[(S-\theta)^2] = \mathbb{V}_\theta(S) + b(\theta)^2$

Si $b(\theta) \neq 0$, S est biaisé.

Ex: moyenne empirique.
Ex: variance empirique.

Définition

- S_n CV en moyenne quadratique si $S_n \xrightarrow{L^2} \theta$
- S_n fortement consistant pour θ si $S_n \xrightarrow{\text{p.s.}} \theta$
- S_n consistant pour θ si $S_n \xrightarrow{\mathbb{P}_\theta} \theta$

Rappel: on dit que S_n converge vers S en distribution (ou en loi) et on note $S_n \rightsquigarrow S$ si la fonction de répartition de S_n converge vers celle de S quand n tend vers $+\infty$ en tout point de continuité.

Théorème

- Loi forte des grands nombres (LFGN):
si $\mathbb{E}[X_i] = m < \infty$,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\mathbb{P}-\text{p.s.}} m \qquad (1)$$

- Th. de la limite centrale (TLC):
si $\mathbb{V}(X) = \sigma^2 < \infty$,

$$\sqrt{n}(\overline{X} - m) \rightsquigarrow \mathcal{N}\left(0, \sigma^2\right) \qquad (2)$$

2 Estimateur plug-in

Définition

- Mesure empirique:

$$\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i} \qquad (3)$$

- Fonction de répartition empirique:

$$F_n(x) = \mathbb{P}_n(]-\infty, x]) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[X_i \le x]} \qquad (4)$$

Rappel important: $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$

$F_n(x)$ est la moyenne empirique de n vaiid de Bernoulli de paramètre $F(x)$, donc LFGN $F_n(x) \xrightarrow{\mathbb{P}-\text{p.s.}} F(x)$ pour tout $x \in \mathbb{R}$. Cette convergence est même uniforme en x, ce qui constitue le théorème de Glivenko-Cantelli ou théorème fondamentale de la statistique.

Définition

Fonctionnelle statistique linéaire: c'est une fonction de $\mathbb{P}$ de la forme:

$$\phi(\mathbb{P}) = \int a(x)\, \mathbb{P}(dx) = \mathbb{E}[a(X)] \qquad (5)$$

Ex: moyenne, variance, médiane d'une va de loi $\mathbb{P}$.

Définition

Un estimateur plug-in est obtenu en substituant $\mathbb{P}_n$ à $\mathbb{P}$ dans une fonctionnelle:
estimateur plug-in associé à $\phi(\mathbb{P}) = \int f(x)\, \mathbb{P}(dx)$:

$$\phi(\mathbb{P}_n) = \frac{1}{n}\sum_{i=1}^n f(X_i) \qquad (6)$$

Théorème

- Si $f \in L^1$, LFGN: $\phi(\mathbb{P}_n) \xrightarrow{\mathbb{P}-\text{p.s.}} \phi(\mathbb{P})$
- Si $f \in L^2$, TLC: $\sqrt{n}(\phi(\mathbb{P}_n) - \phi(\mathbb{P})) \rightsquigarrow \mathcal{N}\left(0, \tau(\mathbb{P})^2\right)$

3 Estimateur par substitution

Définition

Soit S un estimateur de θ et $\psi : \Theta \longrightarrow \mathbb{R}^d$ une fonction mesurable.
Un estimateur par substitution de $\psi(\theta)$ est un estimateur de la forme $\psi(S)$.

Théorème — de l'application continue

Soit S un estimateur (fortement) consistant de θ et $\psi : \Theta \subset \mathbb{R}^p \longrightarrow \mathbb{R}^d$ fonction continue.
Alors $\psi(S)$ est (fortement) consistant pour $\psi(\theta)$.

A.N. : asymptotiquement normal.

Théorème — Méthode delta en dimension 1

S_n estimateur de θ tel que $\sqrt{n}(S_n - \theta) \rightsquigarrow S$
$\psi : \Theta \subset \mathbb{R} \longrightarrow \mathbb{R}$ dérivable. Alors

$$\sqrt{n}\left(\psi(S_n) - \psi(\theta)\right) \rightsquigarrow \psi'(S) \tag{7}$$

Si S_n A.N. et $\sqrt{n}(S_n - \theta) \rightsquigarrow \mathcal{N}(0, \sigma^2)$, alors $\psi(S_n)$ A.N. et

$$\sqrt{n}\left(\psi(S_n) - \psi(\theta)\right) \rightsquigarrow \mathcal{N}\left(0, \Psi'(\theta)^2\sigma^2\right) \tag{8}$$

Théorème — Méthode delta multidim.

S_n estimateur de θ avec $\sqrt{n}(S_n - \theta) \rightsquigarrow S$
$\psi : \Theta \subset \mathbb{R}^p \longrightarrow \mathbb{R}^d$ fonction $\neq$tiable de jacobienne J. Alors

$$\sqrt{n}\left(\psi(S_n) - \psi(\theta)\right) \rightsquigarrow d\psi_\theta(S) \tag{9}$$

et si S_n A.N avec

$$\sqrt{n}(S_n - \theta) \rightsquigarrow \mathcal{N}(0, \Sigma_\theta) \tag{10}$$

alors $\psi(S_n)$ A.N. et

$$\sqrt{n}\left(\psi(S_n) - \psi(\theta)\right) \rightsquigarrow \mathcal{N}\left(0, J\Sigma_\theta J^T\right) \tag{11}$$

Dans le cas de la dim. 1, on a $J\Sigma J^T = \psi'(\theta)^2\sigma_\theta^2$, où σ_θ^2 est la variance de la loi normale asymptotique.

4 Estimateur des moments

Définition

Soit $m_k(\theta) = \mathbb{E}_\theta\left[X^k\right]$ et $\widehat{m}_k(\theta) = \frac{1}{n}\sum_{i=1}^n X_i^k$.

L'estimateur des moments de $\theta \in \Theta \subset \mathbb{R}^p$ est la solution $\hat{\theta}$ du syst. à p éq. d'inconnue θ:

$$\begin{cases} m_1(\theta) = \widehat{m}_1(\theta) \\ \vdots \\ m_p(\theta) = \widehat{m}_p(\theta) \end{cases} \tag{12}$$

Matriciellement: $M(\theta) = \widehat{M}(\theta)$

Ex: échantillon de Poisson, exponentiel, uniforme (dim.1).
Ex: échantillon normal pour $\theta = (m, \sigma^2)$ (dim.2).

Propriété

- Si $M : \Theta \longrightarrow \mathbb{R}^p$ bijective, alors: $\widehat{\theta} = M^{-1} \circ \widehat{M}(\theta)$ et $\widehat{M}(\theta) \xrightarrow[n\to+\infty]{\mathbb{P}_\theta\text{-p.s.}} M(\theta)$.
- Si M bijective et M^{-1} continue: $\widehat{\theta} \xrightarrow[n\to+\infty]{\mathbb{P}_\theta\text{-p.s.}} \theta$
- Si M^{-1} $\neq$tiable et $\mathbb{P}_\theta$ a un moment d'ordre 2: $\sqrt{n}\left(\widehat{M}(\theta) - M(\theta)\right) \rightsquigarrow \mathcal{N}(0, \Sigma_\theta)$ et $\sqrt{n}\left(\widehat{\theta} - \theta\right) \rightsquigarrow \mathcal{N}(0, \tilde{\Sigma}_\theta)$ avec $\Sigma_\theta = \left(\mathrm{cov}(X^i, X^j)\right)_{i,j=1..p}$ et $\tilde{\Sigma}_\theta = J_M^{-1}\Sigma_\theta\left(J_M^{-1}\right)'$

5 Estimateur du max de vraisemblance

Définition

$$\widehat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmax}}\ L(X, \theta) \tag{13}$$

$\widehat{\theta}$ est solution de:

$$\frac{\partial}{\partial\theta}\ln L(X,\theta) = \frac{\partial}{\partial\theta}l(X,\theta) = \nabla l(X,\theta) = S(X,\theta) = 0$$

Théorème — Condition suffisante d'$\exists$ de l'e.m.v.

Si Θ espace compact et l continue sur Θ, l'emv existe.

Théorème — de Zehna

Soit $\widehat{\theta}$ l'e.m.v. de θ et ψ une fonction quelconque. Alors l'e.m.v. de $\psi(\theta)$ est $\psi(\widehat{\theta})$

Théorème — Consistance forte de l'e.m.v.

On considère le modèle d'échantillonnage $X_1, \ldots, X_n$ de v.a.i.i.d. $\sim L(x, \theta)$. Sous les hypothèses de régularité ($\star\star$) ci-dessous,

- Le modèle est identifiable et homogène.
- Θ ouvert ou compact de $\mathbb{R}^p$.
- Pour tout x, $L(x, \theta)$ $\neq$tiable en θ sur Θ.

Alors l'emv est fortement consistant.

6 Estimateur bayésien

$$L(\theta|x) \propto L(x|\theta) \times \Pi(\theta) \tag{14}$$

$\propto$ signifie « proportionnel à ».

- Moyenne *a posteriori*

$$\mathbb{E}[H|X=x] = \int_{\Theta} \theta \times L(\theta|x) d\theta \tag{15}$$

- Médiane *a posteriori*

$$\mathbb{P}[H \leq \theta_{me}] = \int_{-\infty}^{\theta_{me}} L(\theta|x) d\theta = 1/2 \tag{16}$$

- Mode *a posteriori*

$$\underset{\theta}{\operatorname{argmax}}\, L(\theta|x) \tag{17}$$