Statistique Asymptotique

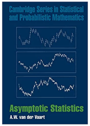
Université Hassiba Benbouali de Chlef

Plan du Cours

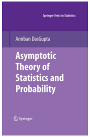
- Modes de convergence
- Méthode Delta
- M et Z-estimateurs

Références

- Van der Vaart, A. W. (2000). Asymptotic statistics. Cambridge university press.



- DasGupta, A. (2008). Asymptotic theory of statistics and probability. Springer.



M-estimateurs

- Soit $X_1, \ldots, X_n$ une suite de variables aléatoires réelles i.i.d et une famille de fonctions mesurables $m_\theta(x)$. Un estimateur $\widehat{\theta}_n$ est obtenu en maximisant

$$\theta \mapsto M_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} m_\theta(X_i)$$

L'estimateur

$$\widehat{\theta}_n = \arg\max_{\theta \in \Theta} M_n(\theta)$$

est appelé un M-estimateur (M pour maximum)

M-estimateurs (suite)

- Pour que cette approche soit pertinente, il faut que la fonction $\theta \mapsto M(\theta) := \mathbb{E}\left[m_\theta(X)\right]$ soit maximale en $\theta = \theta_0$. On espère ainsi que $\widehat{\theta}_n$ ne soit pas trop éloigné de θ_0 (loi des grands nombres).

Plus généralement,

Définition 1.1

Soit M_n une fonction définie sur Θ, à valeurs réelles, et dépendant des observations. On dit que $\widehat{\theta}_n$ est un M-estimateur si

$$M_n(\widehat{\theta}_n) \geq \sup_{\theta \in \Theta} M_n(\theta) - o_P(1)$$

Exemples classiques de M-estimateurs

Exemples

- L'exemple le plus simple est celui de la moyenne empirique $\overline{X}_n$ qui est un M-estimateur pour $m_\theta(x) = -(x-\theta)^2$.
- L'estimateur du maximum de vraisemblance avec $m_\theta(x) = \ln f_\theta(x)$
- La médiane est un M-estimateur associé à la fonction $m_\theta(x) = -|x-\theta|$.

Z-estimateurs

- Un estimateur $\widehat{\theta}_n$ est un Z-estimateur (Z pour zéro) associé à la fonction $\psi_\theta(x)$ qui vérifie $\Psi_n(\widehat{\theta}_n) = 0$, avec

$$\Psi_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \psi_\theta(X_i).$$

Plus généralement,

Définition 1.2

Soit Ψ_n une fonction définie sur Θ, à valeurs dans un espace vectoriel normé, et qui dépend des observations. On dit que $\widehat{\theta}_n$ est un Z-estimateur si

$$\left\|\Psi_n(\widehat{\theta}_n)\right\| = o_P(1)$$

Z-estimateurs (suite)

Remarque

- Étudier un M-estimateur peut se ramener à étudier un Z-estimateur lorsque $\theta \mapsto m_\theta(x)$ est différentiable pour tout x. On a alors $\psi_\theta(x) = \dot{m}_\theta(x)$ (gradient de $\theta \mapsto m_\theta(x)$ au point θ).
- $\psi_\theta(x)$ est appelée la fonction de score.

Exemples classiques de Z-estimateurs

Exemples

- L'exemple le plus simple est celui de la moyenne empirique $\overline{X}_n$ qui est un Z-estimateur pour $\psi_\theta(x) = x - \theta$.
- L'estimateur du maximum de vraisemblance avec $\psi_\theta(x) = \dfrac{\partial}{\partial\theta} \ln f_\theta(x) = \dfrac{\dot{f}_\theta(x)}{f_\theta(x)}$
- La médiane est un Z-estimateur associé à la fonction $\psi_\theta(x) = sign(x - \theta)$.

Consistance des M-estimateurs

Théorème 1.3

On suppose que

1. $\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| \xrightarrow[n \to +\infty]{P} 0$
2. $\forall \epsilon > 0, \quad \sup_{\theta :\ |\theta - \theta_0| \geq \epsilon} M(\theta) < M(\theta_0)$

Alors toute suite d'estimateurs telle que $M_n(\widehat{\theta}_n) \geq \sup_{\theta \in \Theta} M_n(\theta) - o_P(1)$ *vérifie* $\widehat{\theta}_n \xrightarrow[n \to +\infty]{P} \theta_0$

Remarque

Consistance des M-estimateurs (suite)

1. La condition 1 est une loi (faible) des grands nombres uniforme et se montre grâce aux techniques de processus empiriques.
2. Si Θ est compact M est continue, alors la condition 2 est vérifiée. Cette condition demande une "bonne séparation" du point maximum.

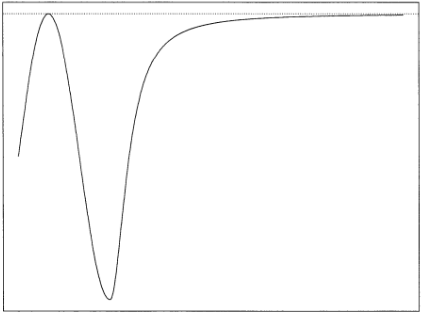


Figure 1 – Exemple de fonction dont le point de maximum n'est pas bien séparé.

Preuve

Soit $\theta_0 \in \Theta$. Comme θ_0 est le maximum de la fonction $\theta \mapsto M(\theta)$, nous avons

$$
\begin{aligned}
0 \leq M\left(\theta_0\right) - M\left(\widehat{\theta}_n\right) &= M\left(\theta_0\right) - M_n\left(\theta_0\right) + M_n\left(\theta_0\right) - M_n\left(\widehat{\theta}_n\right) \\
&\quad + M_n\left(\widehat{\theta}_n\right) - M\left(\widehat{\theta}_n\right), \\
&= M\left(\theta_0\right) - M_n\left(\theta_0\right) + M_n\left(\widehat{\theta}_n\right) - M\left(\widehat{\theta}_n\right) \\
&\quad + \underbrace{M_n\left(\theta_0\right) - M_n\left(\widehat{\theta}_n\right)}_{\leq o_P(1)}, \\
&\leq 2 \sup_{\theta \in \Theta} \left| M_n(\theta) - M(\theta) \right| + o_P(1).
\end{aligned}
$$

Preuve (suite)

Par conséquent, pour tout $\eta > 0$, nous avons

$$\mathbb{P}\left(M\left(\theta_0\right) - M\left(\widehat{\theta}_n\right) \geq \eta\right) = 0.$$

Soit $\epsilon > 0$. D'après la condition 2, il existe $\eta > 0$ tel que $M(\theta) \leq M\left(\theta_0\right) - \eta$ pour tout $\theta \in \Theta$ tels que $\left|\theta - \theta_0\right| \geq \epsilon$, ce qui implique

$$\left\{\left|\widehat{\theta}_n - \theta_0\right| \geq \epsilon\right\} \subset \left\{M\left(\widehat{\theta}_n\right) \leq M(\theta) - \eta\right\}.$$

Par conséquent, nous avons

$$\begin{aligned} \mathbb{P}\left(\left|\widehat{\theta}_n - \theta_0\right| \geq \epsilon\right) &\leq \mathbb{P}\left(M\left(\widehat{\theta}_n\right) < M\left(\theta_0\right) - \eta\right) \\ &\leq \mathbb{P}\left(M\left(\theta_0\right) - M\left(\widehat{\theta}_n\right) > \eta\right) \underset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Consistance des Z-estimateurs

Théorème 1.4

On suppose que

1. $\sup_{\theta \in \Theta} |\Psi_n(\theta) - \Psi(\theta)| \xrightarrow[n \to +\infty]{P} 0$
2. $\forall \epsilon > 0, \quad \inf_{\theta :\ |\theta - \theta_0| \geq \epsilon} \|\Psi(\theta)\| > 0 = \|\Psi(\theta_0)\|$

Alors toute suite d'estimateurs telle que $\left\|\Psi_n(\widehat{\theta}_n)\right\| = o_P(1)$ *vérifie* $\widehat{\theta}_n \xrightarrow[n \to +\infty]{P} \theta_0$

Consistance des Z-estimateurs (suite)

Remarque

Le résultat découle du théorème 1.3 avec $M_n(\theta) = -\|\Psi_n(\theta)\|$ et $M(\theta) = -\|\Psi(\theta)\|$.

Les conditions énoncées dans les deux théorèmes précédents peuvent être affaiblies de multiples façons. En voici un exemple lorsque $\Theta \subset \mathbb{R}$, où l'on affaiblit l'hypothèse de convergence uniforme de Ψ_n vers Ψ.

Consistance des Z-estimateurs (suite)

Proposition 1.5

Supposons que :

1. $\forall \theta \in \Theta, \Psi_n(\theta) \xrightarrow{\mathrm{P}} \Psi(\theta)$,
2. $\forall \theta \in \Theta, \theta \longmapsto \Psi_n(\theta)$ *est continue et s'annule seulement en* $\widehat{\theta}_n$*, ou :*
   (2') $\theta \longmapsto \Psi_n(\theta)$ est croissante, telle que $\Psi_n\left(\widehat{\theta}_n\right) = o_P(1)$,
3. *il existe θ_0 tel que :* $\forall \epsilon > 0, \Psi\left(\theta_0 - \epsilon\right) < 0 < \Psi\left(\theta_0 + \epsilon\right)$.

Alors $\widehat{\theta}_n$ converge en probabilité vers θ_0.

Preuve

Supposons tout d'abord la condition 2. Soit $\epsilon > 0$. Si $\Psi_n(\theta_0 - \epsilon) < 0 < \Psi_n(\theta_0 + \epsilon)$, alors $\theta_0 - \epsilon < \widehat{\theta}_n < \theta_0 + \epsilon$, d'après la condition (2) et le théorème des valeurs intermédiaires. D'où

$$\mathbb{P}\left(\Psi_n(\theta_0 - \epsilon) < 0, \Psi_n(\theta_0 + \epsilon) > 0\right) \leq \mathbb{P}\left(\theta_0 - \epsilon < \widehat{\theta}_n < \theta_0 + \epsilon\right).$$

On montre facilement que
$\mathbb{P}\left(\Psi_n(\theta_0 - \epsilon) < 0, \Psi_n(\theta_0 + \epsilon) > 0\right) \longrightarrow 1$.
En effet, d'après les conditions (1) et (3),
$\Psi_n(\theta_0 - \epsilon) \xrightarrow{\mathrm{P}} \Psi(\theta_0 - \epsilon) < 0$, i.e.

$$\forall \eta > 0, \quad \mathbb{P}\left(\Psi(\theta_0 - \epsilon) - \eta < \Psi_n(\theta_0 - \epsilon) < \Psi(\theta_0 - \epsilon) + \eta\right) \to 1$$

Preuve (suite)

Posons $\eta = -\Psi\left(\theta_0 - \epsilon\right) > 0$. Il vient
$\mathbb{P}\left(2\Psi\left(\theta_0 - \epsilon\right) < \Psi_n\left(\theta_0 - \epsilon\right) < 0\right) \to 1$. Or

$$\left\{2\Psi\left(\theta_0 - \epsilon\right) < \Psi_n\left(\theta_0 - \epsilon\right) < 0\right\} \subset \left\{\Psi_n\left(\theta_0 - \epsilon\right) < 0\right\}$$

d'où $\mathbb{P}\left(\Psi_n\left(\theta_0 - \epsilon\right) < 0\right) \to 1$.
De même, $\mathbb{P}\left(\Psi_n\left(\theta_0 + \epsilon\right) > 0\right) \to 1$. On en déduit que

$$\begin{aligned}\mathbb{P}\left(\Psi_n\left(\theta_0 - \epsilon\right) < 0, \Psi_n\left(\theta_0 + \epsilon\right) > 0\right) = &\\ \mathbb{P}\left(\Psi_n\left(\theta_0 - \epsilon\right) < 0\right) + \mathbb{P}\left(\Psi_n\left(\theta_0 + \epsilon\right) > 0\right)&\\ - \mathbb{P}\left(\left\{\Psi_n\left(\theta_0 - \epsilon\right) < 0\right\} \cup \left\{\Psi_n\left(\theta_0 + \epsilon\right) > 0\right\}\right)&\\ \longrightarrow 1&\end{aligned}$$

Preuve (suite)

D'où $\mathbb{P}\left(\left|\widehat{\theta}_n - \theta_0\right| < \epsilon\right) \longrightarrow 1$, ce qui achève la démonstration.
Supposons maintenant la condition (2'). Soit $\eta > 0$. On a

$$\begin{aligned}
\left\{\left|\widehat{\theta}_n - \theta_0\right| > \eta\right\} &= \left\{\widehat{\theta}_n > \theta_0 + \eta\right\} \cup \left\{\widehat{\theta}_n < \theta_0 - \eta\right\} \\
&\subset \left\{\Psi_n\left(\widehat{\theta}_n\right) \geq \Psi_n\left(\theta_0 + \eta\right)\right\} \\
&\quad \cup \left\{\Psi_n\left(\widehat{\theta}_n\right) \leq \Psi_n\left(\theta_0 - \eta\right)\right\}
\end{aligned}$$

car Ψ_n est croissante. De là, on déduit

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{\theta}_n - \theta_0\right| > \eta\right) &\leq \mathbb{P}\left(\Psi_n\left(\theta_0 + \eta\right) - \Psi_n\left(\widehat{\theta}_n\right) \leq 0\right) \\
&\quad + \mathbb{P}\left(0 \leq \Psi_n\left(\theta_0 - \eta\right) - \Psi_n\left(\widehat{\theta}_n\right)\right).
\end{aligned}$$

Preuve (suite)

Or $\Psi_n(\theta_0 - \eta) - \Psi_n\left(\widehat{\theta}_n\right) \xrightarrow{\mathrm{P}} \Psi(\theta_0 - \eta) < 0$.
Par le même raisonnement que ci-dessus,
$\mathbb{P}\left(0 \leq \Psi_n(\theta_0 - \eta) - \Psi_n\left(\widehat{\theta}_n\right)\right) \longrightarrow 0$.
De même, $\mathbb{P}\left(\Psi_n(\theta_0 + \eta) - \Psi_n\left(\widehat{\theta}_n\right) \leq 0\right) \longrightarrow 0$. D'où
$\mathbb{P}\left(\left|\widehat{\theta}_n - \theta_0\right| > \eta\right) \longrightarrow 0$

Loi limite des Z-estimateurs : principe

- Loi des grands nombres

$$\Psi_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \psi_\theta(X_i) \xrightarrow{P} \Psi(\theta) = \mathbb{E}\left[\psi_\theta(X)\right]$$

- Principe. Développement de Taylor autour de θ_0 :

$$0 = \Psi_n(\widehat{\theta}_n) = \Psi_n(\theta_0) + (\widehat{\theta}_n - \theta_0)\dot{\Psi}_n(\theta_0) + \tfrac{1}{2}(\widehat{\theta}_n - \theta_0)^2 \ddot{\Psi}_n(\widetilde{\theta}_n).$$

- On néglige le reste :

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \approx \frac{-\sqrt{n}\Psi_n(\theta_0)}{\dot{\Psi}_n(\theta_0)}$$

Loi limite des Z-estimateurs : principe

- Convergence du numérateur

$$\sqrt{n}\Psi_n(\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{\theta_0}(X_i) \xrightarrow{d} \mathcal{N}\left(0, \mathbb{E}\left[\psi^2_{\theta_0}(X)\right]\right)$$

si $\mathbb{E}\left[\psi_{\theta_0}(X)\right] = 0$ et $\mathbb{E}\left[\psi^2_{\theta_0}(X)\right] < +\infty$.

- Convergence du dénominateur

$$\dot{\Psi}_n(\theta_0) = \frac{1}{n}\sum_{i=1}^{n}\dot{\psi}_{\theta_0}(X_i) \xrightarrow{P} \mathbb{E}\left[\dot{\psi}_{\theta_0}(X)\right]$$

$\neq 0$ (à supposer).

- $+$ hypothèses techniques pour contrôler le reste (besoin de la convergence de $\widehat{\theta}_n$).

Normalité Asymptotique des Z-estimateurs

Théorème 1.6

Supposons que $\theta \mapsto \psi_\theta(x)$ *est* C^2 *pour tout* x, *et que, pour tout* θ *dans un voisinage de* θ_0, $\left|\ddot{\psi}_\theta(x)\right| \leq h(x)$, *avec* $\mathbb{E}[h(X)] < \infty$.
Supposons que $\Psi(\theta_0) = 0$, *que* $\mathbb{E}\left[\left|\psi_{\theta_0}(X)\right|^2\right] < \infty$ *et que* $\mathbb{E}\left[\dot{\psi}_{\theta_0}(X)\right]$ *est inversible. Soit une suite* $\widehat{\theta}_n$ *telle que* $\forall n, \Psi_n\left(\widehat{\theta}_n\right) = o_P\left(\frac{1}{\sqrt{n}}\right)$, *et* $\widehat{\theta}_n \xrightarrow{\mathrm{P}} \theta_0$. *Alors*

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\mathbb{E}\left[\psi_{\theta_0}^2(X)\right]}{\left(\mathbb{E}\left[\dot{\psi}_{\theta_0}(X)\right]\right)^2}\right)$$