M2RI Asymptotic Statistics Lecture 6: Asymptotic Normality of Z estimators

I. Intuition

Context: z-estimator $\hat{\theta}_n$ s.t.

$$z_n(\hat{\theta}_n) = \frac{1}{n}\sum_{i=1}^{n} z(X_i, \hat{\theta}_n) = 0$$

with $X_1, \dots$ i.i.d $\qquad z : \mathbb{R}^p \times \Theta \to \mathbb{R}$

$$\exists\, \theta_0 \in \Theta \text{ s.t. } \mathbb{E}\left(z(X_1, \theta_0)\right) = 0$$

$$\hat{\theta}_n \xrightarrow[n \to +\infty]{\mathbb{P}} \theta_0 \qquad \text{(cf lectures 4 and 5).}$$

Taylor around θ_0: (which makes sense since $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$)

$$0 = z_n(\hat{\theta}_n) \approx z_n(\theta_0) + (Jz_n)(\theta_0)\,(\hat{\theta}_n - \theta_0).$$

$$\text{"}\Rightarrow\text{"} \quad \sqrt{n}\,(\hat{\theta}_n - \theta_0) = -(Jz_n)(\theta_0)^{-1}\left(\sqrt{n}\, z_n(\theta_0)\right)$$

L.L.N $\quad (Jz_n)(\theta_0) \xrightarrow{\mathbb{P}} \mathbb{E}\left(Jz(\theta_1, \theta_0)\right) =: J$

and if $\det J \neq 0$, C.M $\Rightarrow (Jz_n)(\theta_0)^{-1} \xrightarrow{\mathbb{P}} J^{-1}$

C.L.T. $\quad \sqrt{n}\left[z_n(\theta_0) - \underbrace{\mathbb{E}(z(X_1, \theta_0))}_{=0}\right] \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \operatorname{cov}(z(X_1, \theta_0))\right)$

Slutsky

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, J^{-1}\operatorname{cov}(z(X_1, \theta_0))\, J^{-T}\right)$$

Objective of this lecture: Make this rigorous.

II. Main Theorem

Theorem (Maximal Inequality)

- $(X_i)_{i\geq 1}$ iid with distribution $\mathcal{L}(\mathbb{R}^d)$
- $\mathcal{F}$: set of functions from $\mathbb{R}^k$ to $\mathbb{R}$.
- $\exists F : \mathbb{R}^k \to \mathbb{R}$ s.t. $\forall f \in \mathcal{F}$, $|f(x)| \leq F(x)$ $\mathcal{L}$-a.e.,
  $C_F := \mathbb{E}(F(X_1)^2) < \infty$.

<u>Then</u>
$$\mathbb{E}^* \sup_{f \in \mathcal{F}} \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{n} f(X_i) - \mathbb{E}(f(X_1)) \right| \leq C_{MI} \int_0^{\sqrt{C_F}} \sqrt{\log\left(\mathcal{N}_{[\,]}(\mathcal{F}, L^2(\mathcal{L}), \varepsilon)\right)}\, d\varepsilon$$

<u>Proof:</u> Long and technical, see van der Vaart 2000.

<u>Remark:</u> $\mathbb{E}^*$ is the "outer expectation" in case of non-measurability.

<u>Excercice</u> If $X_1, \dots, \overset{iid}{\sim} \mathrm{Unif}([0,1])$, $\mathcal{F} = \{ \mathbb{1}_{(-\infty, t]} ; t \in [0,1] \}$,
find an upper-bound on

$$\underbrace{\mathbb{E}^* \sup_{f \in \mathcal{F}} \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{n} f(X_i) - \mathbb{E}(f(X_i)) \right|}_{*} .$$

<u>Solution</u> $\mathcal{N}_{[\,]}\left(\mathcal{F}, L^2(\mathrm{Unif}([0,1])), \varepsilon\right) \lesssim \frac{1}{\varepsilon^2}$ (cf last lecture).

$$F = \mathbb{1}_{\mathbb{R}}. \quad E(F(X_1)^2) = 1 =: C_F$$

so $$* \lesssim \int_0^1 \sqrt{\log(1/\varepsilon^2)}\, d\varepsilon \lesssim \int_0^1 \sqrt{-\log(\varepsilon)}\, d\varepsilon$$

$$\varepsilon = e \implies d\varepsilon = -e \ du$$

$$\lesssim \int_{\infty}^{0} e^{u/2}\left(-e^{-u}\right) du = \int_{0}^{\infty} e^{-u/2}\, du \lesssim 1 .$$

<u>Theorem (Asymptotic Normality of Z-estimators).</u> <u>If,</u>

- $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} z(X_i, \hat{\theta}_n) = o_{\mathbb{P}}(1)$ with $z : \mathbb{R}^k \times \Theta \to \mathbb{R}^p$ st $\mathbb{E}(\|z(X_1, \theta)\|^2) < \infty \ \forall \theta$,

  And $\exists \theta_0 \in \Theta$ st $\mathbb{E}(z(X_1, \theta_0)) = 0$,
  $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$.

- $\exists \mathcal{N}$ neighborhood of θ_0 s.t. $\theta \mapsto \mathbb{E}(z(X_1, \theta)) \in C^1(\mathcal{N})$.
  $J\mathbb{E}(X_1, \theta_0)$ is invertible.

- $\mathcal{F}_j := \{ x \mapsto z(x, \theta)_j \ ; \ \theta \in A \}$.
  $\forall j, \forall \delta > 0, \ \int_0^{\delta} \sqrt{\log \mathcal{N}_{[\,]}(\mathcal{F}_j, L^2(\mathcal{L}), \varepsilon)}\, d\varepsilon < \infty$

- $\mathbb{E} \sup_{\substack{\theta \in \mathcal{N} \\ \|\theta - \theta_0\| \leq \delta}} \|z(X_1, \theta) - z(X_1, \theta_0)\|^2 \xrightarrow[\delta \to 0]{} 0$.

<u>Then</u>

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) = -\left(J\mathbb{E}(z(X_1, \theta_0))\right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} z(X_i, \theta_0) + o_{\mathbb{P}}(1)$$

and thus

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}\left(0, J\mathbb{E}(z(X_1, \theta_0))^{-1} \operatorname{cov}(z(X_1, \theta_0)) J\mathbb{E}(z(X_1, \theta_0))^{-T}\right).$$

<u>Proof:</u> $V := J\mathbb{E}(z(X_1, \theta_0))$

Taylor on $\theta \mapsto \mathbb{E}(z(X_1, \theta))$ around θ_0

$$\int_{\mathbb{R}^p} z(x,\theta)\,d\mathcal{L}(x) = \int_{\mathbb{R}^p} z(x,\theta_0)\,d\mathcal{L}(x) + V(\theta-\theta_0) + \underbrace{o(\|\theta-\theta_0\|)}$$

And since $\hat{\theta}_n - \theta_0 = o_{\mathbb{P}}(1)$,

$$\int z(x,\hat{\theta}_n)\,d\mathcal{L}(x) = \int z(x,\theta_0)\,d\mathcal{L}(x) + \underbrace{V(\hat{\theta}_n-\theta_0) + o_{\mathbb{P}}(\|\hat{\theta}_n-\theta_0\|)}_{=\left(V+o_{\mathbb{P}}(1)\right)(\hat{\theta}_n-\theta_0)}.$$

Multiplication by $\sqrt{n}$, $\int z(x,\theta_0)\,d\mathcal{L}(x) = 0$ and the first hypothesis yield

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int z(x,\hat{\theta}_n)\,d\mathcal{L}(x) - z(X_i,\hat{\theta}_n)\right) = \left(V+o_{\mathbb{P}}(1)\right)\sqrt{n}\,(\hat{\theta}_n-\theta_0) + o_{\mathbb{P}}(1).$$

which rewrites as

$$\left(V+o_{\mathbb{P}}(1)\right)\sqrt{n}\,(\hat{\theta}_n-\theta_0) = o_{\mathbb{P}}(1) \overbrace{- \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(z(X_i,\theta_0) - \int_{\mathbb{R}^p} z(x,\theta_0)\,d\mathcal{L}(x)\right)}^{\text{Easy with CLT}}$$

$$\underbrace{+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\left(z(X_i,\theta_0) - z(X_i,\hat{\theta}_n)\right) - \int\left(z(x,\theta_0) - z(x,\hat{\theta}_n)\right)d\mathcal{L}(x)\right)}_{=:\, r_n}$$

$=: r_n$. We need to show that $r_n = o_{\mathbb{P}}(1)$.

Let's define: $\mathcal{F}_{j,\delta} := \{x \mapsto z(x,\theta)_j - z(x,\theta_0)_j \,;\, \theta \in B(\theta_0,\delta)\}$.

If $\|\hat{\theta}_n - \theta_0\| \leq \delta$, we have

$$\|r_n\| \leq \sqrt{p}\,\max_j \sup_{f\in\mathcal{F}_{j,\delta}} \frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n} f(X_i) - \mathbb{E}\left(f(X_1)\right)\right|.$$

Let's notice that $\forall \varepsilon>0$, $\mathcal{N}_{[\,]}(\mathcal{F}_{j,\delta}, L^2(\mathcal{L}), \varepsilon) \leq \mathcal{N}_{[\,]}(\mathcal{F}_j, L^2(\mathcal{L}), \varepsilon)$.

Then, $\forall \delta, \varepsilon > 0$ with $B(\theta_0, \delta) \subset A$,

$$\mathbb{P}(\|r_n\| \geq \varepsilon) \leq \mathbb{P}(\|\hat{\theta}_n - \theta_0\| \geq \delta) + \mathbb{P}\left(\sqrt{p}\max_j \sup_{f \in F_{j,\delta}} \frac{1}{\sqrt{n}}\left|\sum_{i=1}^n f(X_i) - \mathbb{E}f(X_1)\right| \geq \varepsilon\right)$$

And since $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$,

$$\limsup_n \mathbb{P}(\|r_n\| \geq \varepsilon) \leq \mathbb{P}\left(\sqrt{p}\max_j \sup_{f \in F_{j,\delta}} \frac{1}{\sqrt{n}}\left|\sum_{i=1}^n f(X_i) - \mathbb{E}f(X_1)\right| \geq \varepsilon\right).$$

$\forall f \in F_{j,\delta}$ and $x \in \mathbb{R}^b$, $|f(x)| \leq F_\delta(x)$

with $F_\delta(x) = \sup_{\substack{\theta \in A \\ \|\theta - \theta_0\| \leq \delta}} \|z(X_1, \theta) - z(X_1, \theta_0)\|$

So, by Markov and Maximal inequality,

$$\limsup_n \mathbb{P}(\|r_n\| \geq \varepsilon) \leq \sum_{j=1}^p \mathbb{P}\left(\sup_{f \in F_{j,\delta}} \frac{1}{\sqrt{n}}\left|\sum_{i=1}^n f(X_i) - \mathbb{E}f(X_1)\right| \geq \frac{\varepsilon}{\sqrt{p}}\right)$$

$$\leq \sum_{j=1}^p \frac{\sqrt{p}}{\varepsilon}\,\mathbb{E}\sup_{f \in F_{j,\delta}} \frac{1}{\sqrt{n}}\left|\sum_{i=1}^n f(X_i) - \mathbb{E}f(X_1)\right|$$

$$\leq \sum_{j=1}^p \frac{\sqrt{p}}{\varepsilon}\, C_{MI} \underbrace{\int_0^{\sqrt{\mathbb{E}(F_\delta(X_1)^2)}} \sqrt{\log \mathcal{N}_{[\,]}(F_{j}, L^2(\mathcal{L}), u)}\, du}_{\substack{\text{Can be made as small as we want} \\ \text{because } \mathbb{E}(F_\delta(X_1)^2) \xrightarrow[\delta \to 0]{} 0}}$$

$\square$

III. Application to the median.

$(X_i)_{i \in \mathbb{N}}$ iid with F_{X_1} and with density f

Assumption 1: $f(x) > 0 \quad \forall x \in \mathbb{R}$ almost surely.

Population median: θ_0 the only real number such that $F_{X_1}(\theta_0) = \frac{1}{2}$

Empirical median: $\hat{\theta}_n$ s.t. $\sum_{i=1}^{n} \text{sgn}(\hat{\theta}_n - X_i) = 0$

Remark: We might restrict n to even numbers and that all the X's are different (it happens almost surely).

$\Rightarrow$ By lecture 4, $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$ (The first point of the theorem is satisfied)

Assumption 2: f is continuous on a neighborhood of θ_0.

$\Rightarrow \mathbb{E}(\text{sgn}(\theta - X_1)) = 2F_{X_1}(\theta) - 1$ is C^1 on a neighborhood of θ_0 with derivatives $2f(\theta_0)$ at θ_0.

(The second point of the theorem is satisfied).

Next Step: $\mathcal{F} := \{x \mapsto \text{sgn}(\theta - x) ; \theta \in \mathbb{R}\}$

Control $\mathcal{N}_{[\,]}(\mathcal{F}, L^2(\mathcal{L}), \varepsilon)$

Exercice:

1) Show that $\forall N$ st $N+1 \geq \frac{1}{\varepsilon^2}$, $\exists t_1, \dots, t_N$ st

$\forall j = 0, \dots, N, \ \mathcal{L}\big((t_j, t_{j+1})\big) \leq \varepsilon^2$

with $t_0 = -\infty$ and $t_{N+1} = +\infty$.

2) By considering functions of the form $x \mapsto \mathbb{1}(t_j \leq x)$ and $x \mapsto \mathbb{1}(t_j < x)$,

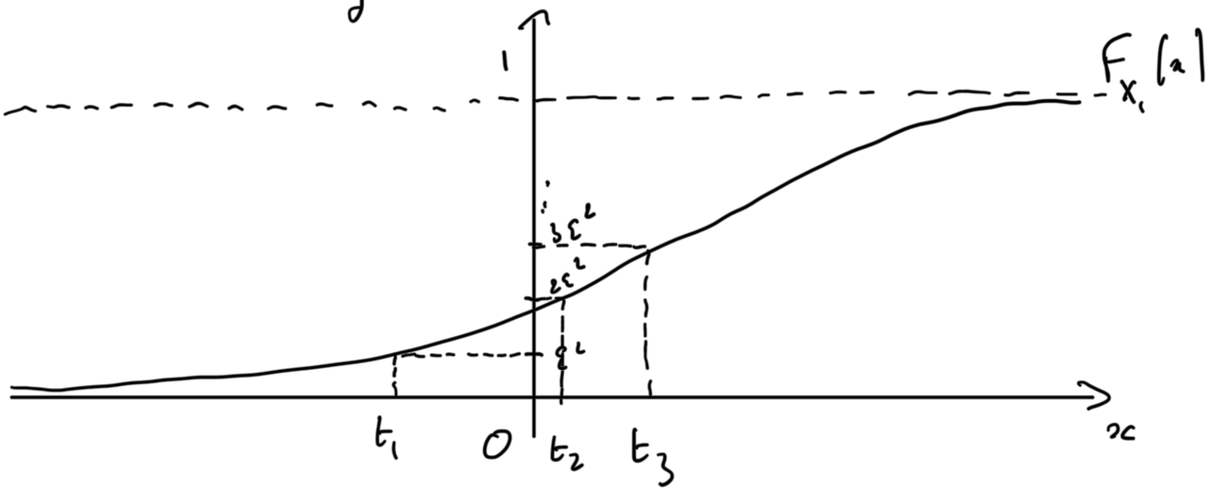
Build brackets of $\mathcal{F}_- = \{x \mapsto \mathbb{1}(\theta < x) \,;\, \theta \in \mathbb{R}\}$.

Do the same thing for $\mathcal{F}_+ = \{x \mapsto \mathbb{1}(x < \theta) \,;\, \theta \in \mathbb{R}\}$

3) After noticing that $\operatorname{sgn}(\theta - x) = \mathbb{1}(x < \theta) - \mathbb{1}(\theta < x)$, conclude.

Solution :

1) On a drawing



2) $\forall j = 1, \dots, N, \quad l_{-,j}(x) = \mathbb{1}(t_{j+1} \leq x)$

$u_{-,j}(x) = \mathbb{1}(t_j < x)$

then $\forall j \forall x, \ l_{-,j}(x) \leq u_{-,j}(x)$

and $\forall \theta, \exists j$ st $\forall x, \ l_{-,j}(x) \leq \mathbb{1}(\theta \leq x) \leq u_{-,j}(x)$.

Finally, $\int (u_{-,j} - l_{-,j})^2 \, d\mathcal{L} = \int \mathbb{1}\big(x \in (t_j, t_{j+1})\big) \, d\mathcal{L}(x)$

$$= \mathcal{L}\big((t_j, t_{j+1})\big)$$

$$\leq \varepsilon^2.$$

Same reasoning for $\mathcal{F}_+$ with $l_{+,j} = \mathbb{1}(x \leq t_j)$

$u_{+,j} = \mathbb{1}(x < t_{j+1})$.

3)

$$\operatorname{sgn}(\theta - x)$$

$$\shortparallel$$

$$\underbrace{\underbrace{l_{+,j}(x) - u_{-,j}(x)} \leq \mathbb{1}(x < \theta) - \mathbb{1}(\theta < x) \leq \underbrace{u_{+,j}(x) - l_{-,j}(x)}}_{\text{Brackets for } F.}$$

and by triangular inequality,

$$\big\| (l_{+,j} - u_{-,j}) - (u_{+,j} - l_{-,j}) \big\|_{L^2(\mathcal{L})}$$

$$= \big\| (l_{+,j} - u_{+,j}) + (l_{-,j} - u_{-,j}) \big\|_{L^2(\mathcal{L})}$$

$$\leq 2\varepsilon.$$

Thus, $N_{[\,]}(F, L^2(\mathcal{L}), 2\varepsilon) \leq N + 1$ and since we can chose $N+1 \leq \frac{1}{\varepsilon^2} + 1$,

$N_{[\,]}(F, L^2(\mathcal{L}), 2\varepsilon) \leq \frac{1}{\varepsilon^2} + 1$, which gives

$$\mathcal{N}_{[\,]}\left(\mathcal{F}, L^2(\mathcal{L}), \varepsilon\right) \leq \frac{4}{\varepsilon^2} + 1$$

$\square$

Furthermore, $\mathbb{E}\left(\operatorname{sgn}\left(\theta_0 - X_1\right)^2\right) = \mathbb{E}(1) = 1$.

thus
$$\int_0^{\sqrt{\mathbb{E}(\operatorname{sgn}(\theta_0 - X_1)^2)}} \sqrt{\log\left(\mathcal{N}_{[\,]}\left(F, L^2(\mathcal{L}), \varepsilon\right)\right)} < \infty.$$

(The third point of the theorem is satisfied).

- $$\mathbb{E}\left(\sup_{\substack{\theta \in \mathbb{R} \\ \|\theta - \theta_0\| \leq \delta}} \left(\operatorname{sgn}(\theta - X_1) - \operatorname{sgn}(\theta_0 - X_1)\right)^2\right)$$

$$= 2\,\mathbb{P}\left(X_1 \in [\theta_0 - \delta, \theta_0 + \delta]\right) \xrightarrow[\delta \to 0]{} 0$$

(The fourth hypothesis of the theorem is satisfied).

So,
$$\boxed{\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{4 f^2(\theta_0)}\right)}$$

IV. Application to Maximum Likelihood

Lemma
- $\mathcal{L}$: distribution on $\mathbb{R}^k$
- Θ : bounded set of $\mathbb{R}^p$
- $\mathcal{F} = \{ f_\theta \,;\, \theta \in \Theta \}$ with $\forall \theta, \int f_\theta^2 \, d\mathcal{L} < +\infty$.
- $\exists\, h : \mathbb{R}^k \to [0, \infty)$ with $1 \leq \int h^2 d\mathcal{L} < \infty$ s.t.

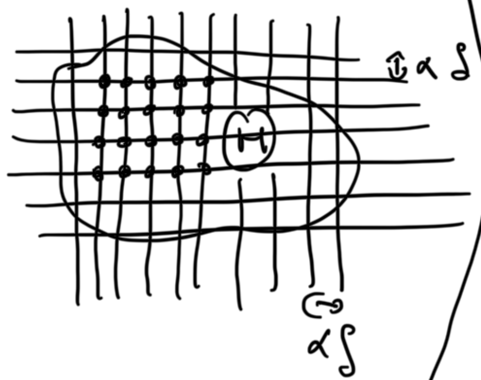
$$\forall \theta_1, \theta_2 \in \Theta, \ \forall x \in \mathbb{R}^k, \ |f_{\theta_1}(x) - f_{\theta_2}(x)| \leq \|\theta_1 - \theta_2\| \, h(x)$$

Then, $\forall \varepsilon > 0$, $N_{[\,]}(\mathcal{F}, L^2(\mathcal{L}), \varepsilon) \leq C_p \, \mathrm{diam}(\Theta)^p \left(\int h^2 d\mathcal{L} \right)^{\frac{p}{2}} \dfrac{1}{\varepsilon^p}$.

Proof There exists a constant C_p' such that, $\forall \delta > 0$, $\exists N \in \mathbb{N}$ with $N \leq C_p' \, \mathrm{diam}\,\Theta^p \dfrac{1}{\delta^p}$ and $\exists \theta_1, \dots, \theta_N \in \Theta$ with

$$\sup_{\theta \in \Theta} \min_{j = 1, \dots, N} \|\theta - \theta_j\| \leq \delta$$

(Left as exercice but



)

$$\forall j = 1, \dots, N, \qquad l_j(x) = f_{\theta_j}(x) - 2\delta h(x)$$
$$u_j(x) = f_{\theta_j}(x) + 2\delta h(x).$$

Then, $l_j(x) \leq u_j(x) \ \forall x$ and $\int (l_j - u_j)^2 d\mathcal{L} = 16 \delta^2 \int h^2 d\mathcal{L}$.

Furthermore, $\forall \theta \in \Theta$, $\exists j$ s.t. $\|\theta - \theta_j\| \leq 2\delta$.

Thus, $f_\theta(x) \geq f_{\theta_j}(x) - \|\theta - \theta_j\| h(x) \geq f_{\theta_j}(x) - 2\delta h(x) = \ell_j(x)$

and similarly, $f_\theta(x) \leq u_j(x)$.

So, $\mathcal{N}_{[\,]}\left(F, L^2(\mathcal{L}), 4\delta\sqrt{\int h^2 d\mathcal{L}}\right) \leq C_p' \operatorname{diam}(\Theta)^p \dfrac{1}{\delta^p}$.

$\left(\varepsilon = 4\delta\sqrt{\int h^2 d\mathcal{L}}\right)$
$\Rightarrow$ $\mathcal{N}_{[\,]}(F, L^2(\mathcal{L}), \varepsilon) \leq 4^p C_p' \operatorname{diam}(\Theta)^p \left(\int h^2 d\mathcal{L}\right)^{p/2} \dfrac{1}{\varepsilon^p}$.

Study of maximum likelihood

Model $\{\mathcal{L}_\theta ; \theta \in \Theta\}$ $\quad$ $\mathcal{L}_\theta$ with density f_θ w.r.t. Lebesgue.

$X_1, X_2, \dots \overset{i.i.d}{\sim} f_{\theta_0}$ $\quad$ $\theta_0 \in \mathring{\Theta}$.

<u>Assumptions</u> $\forall \theta, x$, $f_\theta(x) > 0$ and $\forall x$, $f_\cdot(x)$ is C^2.

$\forall \theta$, $\mathbb{E}\|\nabla_\theta(\log f_\theta(X_1))\|^2 < \infty$

Maximum likelihood

$\hat{\theta}_n$ s.t. $\dfrac{1}{n}\sum_{i=1}^{n} \nabla_\theta \log f_\theta(x) = \dfrac{1}{n}\sum_{i=1}^{n} \dfrac{\nabla_\theta f_\theta(x)}{f_\theta(x)} = 0$

<u>Assumption</u> $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$.

Notation: $z(x, \theta) := \dfrac{1}{f_\theta(x)} \nabla_\theta f_\theta(x)$.

Interversion of $\int$ and ∇:

$$\mathbb{E}\left(z(X_1,\theta_0)\right) = \mathbb{E}\left(\frac{\nabla_\theta f_{\theta_0}(X_1)}{f_{\theta_0}(X_1)}\right) = \int \nabla_\theta f_{\theta_0}(x)\,\frac{f_{\theta_0}(x)}{f_{\theta_0}(x)}\,dx$$

$$= \int \nabla_\theta f_{\theta_0}(x)\,dx$$

So, If $\displaystyle\int \sup_{\theta\in\Theta} \|\nabla f_\theta(x)\|\,dx < \infty$,

$$\mathbb{E}\left(z(X_1,\theta_0)\right) = \nabla_\theta\left(\int f_{\theta_0}(x)\,dx\right) = \nabla_\theta 1 = 0 .$$

(first point of Th holds)

Study of the Jacobian

$$J_\theta\left(\mathbb{E}\, z(X_1,\theta)\right) = J_\theta \int \nabla_\theta\left(\log f_\theta(x)\right) f_{\theta_0}(x)\,dx$$

and If $\forall a,b$, $\displaystyle\int \sup_{\theta\in\Theta} \left|\frac{\partial^2\left(\log\left(f_\theta(x)\right)\right)}{\partial\theta_a\,\partial\theta_b}\right| f_{\theta_0}(x)\,dx < \infty$,

$$J_\theta\left(\mathbb{E}\, z(X_1,\theta)\right) = \int \left(J_\theta \nabla_\theta\right)\left(\log f_\theta(x)\right) f_{\theta_0}(x)\,dx$$

Finally, $\forall a,b$,

$$\left(J_\theta \mathbb{E}\left(z(X_1,\theta_0)\right)\right)_{a,b} = \int \left(\frac{\frac{\partial^2 f_{\theta_0}}{\partial\theta_a\partial\theta_b} f_{\theta_0}(x) - \frac{\partial f_{\theta_0}(x)}{\partial\theta_a}\,\frac{\partial f_{\theta_0}(x)}{\partial\theta_b}}{f_{\theta_0}(x)^2}\right) f_{\theta_0}(x)\,dx$$

$$= \int \frac{\partial^2 f_{\theta_0}(x)}{\partial\theta_a\,\partial\theta_b}\,dx - \int \frac{\partial \log f_{\theta_0}(x)}{\partial\theta_a}\,\frac{\partial \log f_{\theta_0}(x)}{\partial\theta_b}\,f_{\theta_0}(x)\,dx$$

and if $\int \sup_{\theta \in \Theta} \frac{\partial^2 f_\theta(x)}{\partial\theta_a \partial\theta_b} \, dx < \infty$, the previous two terms make sense, and

$$\int \frac{\partial^2 f_{\theta_0}}{\partial\theta_a \partial\theta_b}(x)\, dx = \frac{\partial \int \frac{\partial f_{\theta_0}}{\partial \theta_a} dx}{\partial\theta_b} = \frac{\partial \frac{\partial 1}{\partial\theta_a} dx}{\partial\theta_b} = \frac{\partial 0}{\partial\theta_b} = 0.$$

So $J\,\mathbb{E}\left(z(X_1,\theta_0)\right) = -\operatorname{cov}\left(z(X_1,\theta_0)\right)$

Assumption : $\operatorname{cov}\left(z(X_1,\theta_0)\right)$ is invertible.

(second point of Th. holds.)

Furthermore,

$$\left| \frac{\partial \log f_{\theta_1}(x)}{\partial\theta_a} - \frac{\partial \log f_{\theta_2}(x)}{\partial\theta_a} \right| \leq \|\theta_1 - \theta_2\| \sup_\theta \left\| \left(\nabla^2 \log f_\theta(x)\right)_a \right\|_2$$

$$\leq \|\theta_1 - \theta_2\| \sqrt{p}\, \underbrace{\max_b \sup_\theta \left| \frac{\partial^2 \log f_\theta(x)}{\partial\theta_a \partial\theta_b} \right|}_{=: h(x)}$$

and Item 3 holds thanks to the last result.

Finally,

$$\mathbb{E} \sup_{\substack{\theta \\ \|\theta-\theta_0\| \leq \delta}} \|z(X_1,\theta) - z(X_1,\theta_0)\|^2$$

$$\leq \mathbb{E}\, \delta^2 \left(\sup_\theta \|J_\theta z(X_1,\theta)\|\right)^2$$

$$= \mathbb{E}\, \delta^2 \sup_\theta \|J_\theta z(X_1,\theta)\|^2 = \mathbb{E}\, \delta^2 \sum_{a,b} \sup_\theta \left(\frac{\partial^2 \log f_\theta(X_1)}{\partial\theta_a \partial\theta_b}\right)^2$$

$$\leq \delta \times C^{te} \text{ (by domination assumptions)}.$$

(So Item 4 holds).

So,

$$\boxed{\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}\left(0, \operatorname{cov}\left(z\left(X_1, \theta_0\right)\right)^{-1}\right)}$$