Consistência / (vários conceitos)

Def: Dizemos q' a sequencia de variáveis aleatórias $\{Z_n\}$ converge em probabilidade até a v.a. Z; $Z_n \xrightarrow{P} Z$ se $P(|Z_n - Z| \geq \varepsilon) \underset{n\to\infty}{\to} 0$ $\forall \varepsilon > 0$.

- A sequência de vetores $Z_n = (Z_{n1}, \dots, Z_{nd})^T$ converge em probabilidade até o vetor aleatório $Z = (Z_1, \dots, Z_d)^T$. se $|Z_n - Z| \xrightarrow{P} 0$

onde $|\cdot|$: representa a distância euclideana.

ou equivalentemente, $Z_{nj} \xrightarrow{P} Z_j$, $1 \leq j \leq d$.

—o—

Suponho q' $X_1, \dots, X_n$ é uma amostra desde P_θ, $\theta \in \Theta$.

$g(\theta) \leftarrow$ quantidade a ser estimada; $\hat{g}_n(X_1, \dots, X_n)$: estimativa de $g(\theta)$ baseada na amostra $X_1, \dots, X_n$.

Qué é a tendencia de $\hat{g}_n(X_1, \dots, X_n)$ qdo $n \to \infty$.

- desejariamos por exemplo: $\hat{g}_n \xrightarrow[n\to\infty]{P_\theta} g(\theta)$. $\forall \theta$. (consistência)

ie $\forall \varepsilon > 0$ $\quad P_\theta(|\hat{g}_n(X_1, \dots, X_n) - g(\theta)| \geq \varepsilon) \underset{n\to\infty}{\to} 0$ $\quad \forall \theta$. $\quad$ -1-

- Uma condição frequentemente exigida é: Consistência Uniforme

$$\sup_\theta \left\{ P_\theta \{ |\hat{g}_n(X_1, \dots, X_n) - g(\theta)| \geq \varepsilon \} ; \theta \in \Theta \right\} \underset{n\to\infty}{\to} 0 \quad -2-$$

<u>Exemplo</u> (Médias). $X_1 \ldots X_n$ i.i.d. P; P desconhecido.

$E_P(|X_1|) < \infty$. (Lei fraca dos grandes números)

$$\frac{S_n}{n} = \frac{X_1 + \cdots + X_n}{n} \xrightarrow{P} \mu = E(X_1)$$

Desigualdade de Markov

ie $\bar{X} \xrightarrow{P} \mu(P) = E(X_1)$, logo $\bar{X}$ é consistente a $\mu(P)$

($P(g(\underline{X}) \geq k) \leq \frac{E(g(\underline{X}))}{k}$, se $g(\underline{X}) = (\bar{X} - \mu)^2$, $k = \varepsilon^2 \Rightarrow P((\bar{X}-\mu)^2 \geq \varepsilon^2) \leq \frac{E((\bar{X}-\mu)^2)}{\varepsilon^2} = \frac{Var(\bar{X})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$ e $P(|\bar{X}-\mu|^2 \geq \varepsilon^2) = P(|\bar{X}-\mu| \geq \varepsilon)$)

Que podemos dizer da consistência uniforme em geral? Exercício 5.2.2

- Caso especial Seja $\mathcal{P} = \{P : E_P(X_1^2) \leq M < \infty\}$, então $\bar{X}$ é uniformemente consistente a $\mu = \mu(P)$.

Vejamos: desde a desigualdade de Chebyshev.

$$P(|\bar{X} - \mu(P)| \geq \varepsilon) \leq \frac{Var(\bar{X})}{\varepsilon^2} \leq \frac{\tilde{M}}{n\varepsilon^2} \xrightarrow[n \to \infty]{} 0 \quad \#$$

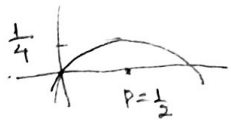
<u>Exemplo</u> (Binômio) Sejam $X_1 \ldots X_n$ indicadores $Ber(p)$, $P(X_1 = 1) = p$

Então $N = \sum_{i=1}^{n} X_i \sim Bin(n, p)$; $p \in [0,1]$.

$\hat{p} = \bar{X} = \frac{N}{n}$ é uniformemente consistente estimativa de p.

$Var(Ber) = p \cdot (1-p) = g(p)$, vejamos se $\underbrace{g(\hat{p}) \text{ é uniformemente consistente}}_{(*)}$

para $p(1-p)$:



$\tilde{M} = \frac{1}{4}$.

<u>Propriedade</u>: Se $Z_n \xrightarrow{P} z_0$ (cte) g contínua em z_0 então $g(Z_n) \xrightarrow{P} g(z_0)$

<u>dem</u>. $\forall \varepsilon > 0 \ \exists \underset{\delta(\varepsilon)}{\delta}$ tq se $|z - z_0| < \delta \Rightarrow |g(z) - g(z_0)| < \varepsilon$

logo $P(|g(Z_n) - g(z_0)| < \varepsilon) \geq P(|Z_n - z_0| < \delta) = 1 - \underbrace{P(|Z_n - z_0| \geq \delta)}_{\xrightarrow[n \to \infty]{} 0}$ #

$A \subseteq B \Rightarrow P(A) \leq P(B)$

cada vez que $\omega \in A \Rightarrow \omega \in B$

e (*) segue.

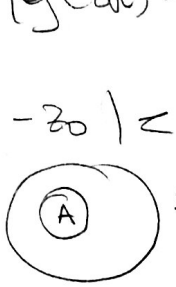
$$z_n \xrightarrow{P} z_0 \Rightarrow g(z_n) \xrightarrow[n\to\infty]{P} g(z_0)$$

g contínuo.

- Si g é contínuo $\forall \varepsilon > 0 \ \exists \ \delta = \delta(\varepsilon)$ tal que

se $|z_n - z_0| < \delta \Rightarrow |g(z_n) - g(z_0)| < \varepsilon$

Se $A = \{\omega : |z_n(\omega) - z_0| < \delta\}$; $B = \{\omega : |g(z_n(\omega)) - g(z_0)| < \varepsilon$

$\Rightarrow A \subseteq B \Rightarrow$ 

$\Rightarrow A^c \supseteq B^c$

$$P(B^c) = P(|g(z_n) - g(z_0)| \geq \varepsilon) \leq P(|z_n - z_0| \geq \delta) \xrightarrow[n\to\infty]{} 0$$

<u>Teorema</u>: Suponha q' $\mathcal{P} = S = \{ (p_1, \dots, p_k) : 0 \le p_j \le 1 ; 1 \le j \le k, \sum_{j=1}^{k} p_j = 1 \}$ é o k-dimensional simplex. ; $p_j = P(X_1 = x_j)$ $j \in \{1, \dots, k\}$. e $\{x_1, \dots, x_k\}$ valores de X_1. Seja $N_j = \sum_{i=1}^{n} 1_{(X_i = x_j)}$; $\hat{p}_j = \frac{N_j}{n}$ $p_n = (\hat{p}_1, \dots, \hat{p}_k) \in S$. Suponha q' $g : S \to \mathbb{R}^p$ é contínua. então $\hat{g}_n = g(p_n)$ é uniformemente consistente de $g(p)$.

<u>dem</u>: pela lei fraca dos grandes números $\forall$ $p, \delta > 0$

$$P_p(|\hat{p}_n - p| \geq \delta) \xrightarrow[n \to \infty]{} 0$$

Como g é contínua em S, S é compacto, então g é uniformemente contínua em S, logo, $\forall \varepsilon > 0$ $\exists$ $\delta = \delta(\varepsilon) > 0$ tq se $|p - p'| < \delta \Rightarrow$ $|g(p) - g(p')| \le \varepsilon$. ; então não depende de p e p' (no compacto)

$$P_p(|\hat{g}_n - g| \geq \varepsilon) \le P_p(|\hat{p}_n - p| \geq \delta(\varepsilon)).$$

Por outro lado $\sup\{P_p(|\hat{p}_n - p| \geq \delta), p \in S\} \le \frac{k}{4n\delta^2}$

$\hookrightarrow$ Problema 5.2.1

*

Seja g: aplicação, $g: \mathbb{R}^k \to \mathbb{R}^p$, definamos

$$\omega(g,\delta) = \sup\{|g(p)-g(p')| : |p-p'| \leq \delta\}$$

- notemos q' $\omega(g,\cdot)$ é crescente em $\delta \in [a,b)$.
- se g for contínua $\omega(g,\delta) \downarrow 0$ qdo $\delta \downarrow 0$.
- Podemos definir para g contínua $\omega^{-1}(\epsilon) \doteq \inf\{\delta : \omega(g,\delta) \geq \epsilon\}$

$\omega^{-1}: [a,b] \to \mathbb{R}^+$

<u>Afirmação</u>: (problema 5.2.3)

$$\sup\{P(|\hat{g}_n - g(p)| \geq \epsilon) : p \in \mathcal{P}\} \leq n^{-1}[\omega^{-1}(\epsilon)]^{-2} \cdot \frac{k}{2}$$

<u>Proposição</u>: Seja $g = (g_1, \ldots, g_d)$ um mapeamento de $\mathcal{X}$ em $\mathcal{Y} \subseteq \mathbb{R}^d$. Suponha ainda q' $E_\theta(|g_j(X_1)|) < \infty$; $1 \leq j \leq d$ $\forall \theta$. Seja $m_j(\theta) = E_\theta g_j(X_1)$ $1 \leq j \leq d$ e seja $g(\theta) = h(m(\theta))$; $m(\theta) = (m_1(\theta), \ldots, m_d(\theta))$. onde $h: \mathcal{Y} \to \mathbb{R}^p$. Então, se h é contínua.

$$\hat{g} = h(\bar{g}) = h\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right)$$

é estimador consistente de $g(\theta)$.

<u>dem</u>: Sendo P caracterizada por θ. e aplicando a lei fraca. e tendo $X_1 \ldots X_n$ a a, resulta

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i) \xrightarrow[n\to\infty]{P} E_\theta(g(X_1)) \quad \text{se } E_\theta(|g(X_1)|) < \infty.$$

pag 512

<u>Proposição</u>: (exercício) Se $z_n \xrightarrow[n\to\infty]{P} z$, g contínua $g: \mathbb{R}^d \to \mathbb{R}^p \Rightarrow g(z_n) \xrightarrow[n\to\infty]{P} g(z)$

aplicando a proposição, a $h: \mathcal{Y} \to \mathbb{R}^p$.

resulta $h(\bar{g}) \xrightarrow[n\to\infty]{P} h\left(E_\theta(g(x_1))\right) = g(\theta)$ ✱

<u>Exemplo</u> (Variâncias e correlações) Seja $x_i = (U_i, V_i)$; $1 \le i \le n$

iid $N_2(\overbrace{\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho}^{\theta})$, $\sigma_i^2 > 0$, $|\rho| < 1$.

Seja $g(u,v) = (u, v, u^2, v^2, u.v)$ logo $g: \mathbb{R}^2 \to \mathbb{R}^5$

$\sum_{i=1}^{n} g(U_i, V_i)$: estatística 5-dimensional

se $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$; $m(\theta) = (m_1(\theta), \dots, m_5(\theta))$

$$= \left(E_\theta(g_1(U_1,V_1)); \dots, E_\theta(g_5(U_1,V_1))\right)$$

$$= \left(\mu_1, \mu_2, \sigma_1^2 + \mu_1^2, \sigma_2^2 + \mu_2^2, \rho\sigma_1\sigma_2 + \mu_1 \cdot \mu_2\right)$$

Seja $h = m^{-1}$ (para recuperar θ $\theta = m^{-1} m(\theta)$)

$$h(m_1, m_2, m_3, m_4, m_5) = \left(m_1; m_2; m_3 - m_1^2; m_4 - m_2^2; (m_5 - m_1 \cdot m_2)(m_3 - m_1^2)^{-1/2} \cdot (m_4 - m_2^2)^{-1/2}\right)$$

Contínua.

aplicando a proposição anterior, resulta $h(\bar{g}) = m^{-1}(\bar{g})$ é estimativa consistente de $g(\theta) = h(m(\theta)) = \theta$

ie as médias, variâncias e correlações empíricas são estimativas consistentes das medias, variâncias e correlação do modelo

$g(u,v) = (u, v, u^2, v^2, u.v)$

$g_1(u,v) = u$, $g_2(u,v) = v$, $g_3(u,v) = u^2$, $g_4(u,v) = v^2$

$g_5(u,v) = u.v$

$\bar{g}_1(\{u_i, v_i\}_{i=1}^n) = \frac{1}{n}\sum_{i=1}^n u_i$; ... ; $\bar{g}_3(\{u_i, v_i\}_{i=1}^n) = \frac{1}{n}\sum_{i=1}^n u_i^2$; ...

$\bar{g}_5(\{u_i, v_i\}_{i=1}^n) = \frac{1}{n}\sum_{i=1}^n u_i \cdot v_i$

$$E\left(\bar{g}_1(\{u_i, v_i\})\right) = E\left(\frac{\sum u_i}{n}\right) = \mu_1$$

$$E\left(\bar{g}_3(\{u_i, v_i\})\right) = \frac{1}{n}\sum_{i=1}^{n} E(U_i^2) = \sigma_1^2 + \mu_1^2$$

$$U_i \sim N(\mu_1, \sigma_1^2) \Rightarrow \underbrace{Var(U_i)}_{E(U_i^2) - [E(U_i)]^2} = \sigma_1^2$$

ie $E(U_i^2) = \sigma_1^2 + \mu_1^2$

se $m_1 = \mu_1$; $m_3 = \sigma_1^2 + \mu_1^2 \Rightarrow m_3 - m_1 = \sigma_1^2$.

$$E\left(\bar{g}_5(\{u_i, v_i\})\right) = \frac{1}{n} \cdot n E(U \cdot V)$$

$$Cov(U,V) = E(U \cdot V) - E(U)E(V) \Rightarrow E(U \cdot V) = Cov(U,V) + E(U)E(V)$$

$$Cov(U,V) = correloc(U,V) \cdot \sqrt{Var(U) \cdot Var(V)}$$

$$= \rho \cdot \sigma_1 \cdot \sigma_2$$

$$E(U \cdot V) = \rho \sigma_1 \sigma_2 + \mu_1 \cdot \mu_2 \Rightarrow E\left(\bar{g}_5(\{u_i, v_i\})\right) = \rho \sigma_1 \cdot \sigma_2 + \mu_1 \cdot \mu_2$$

$m_1 = \mu_1$

$m_2 = \mu_2$

$m_3 = \sigma_1^2 + \mu_1^2$

$m_4 = \sigma_2^2 + \mu_2^2$

$$\Rightarrow (m_5 - m_1 \cdot m_2) \cdot (m_3 - m_1^2)^{-1/2} (m_4 - m_2^2)^{-1/2}$$

Teorema 5.2.2 Suponha q' $\mathcal{P}$ é a flia de representação canonica da flia exponencial de posto d gerada por T. Seja η, $\mathcal{E}$ e $A(\cdot)$ os elementos q definem $\mathcal{P}$. Suponha q' $\mathcal{E}$ é aberto. Então, se $x_1, \dots, x_n$ são amostras desde $P_\eta \in \mathcal{P}$,

(i) $P_\eta(\hat{\eta} \text{ EMV } \exists) \to 1$

(ii) $\hat{\eta}$ é consistente para η.

$$h(x_1)\dots h(x_n)\exp\{\eta T(x_1) - A(\eta)\}\dots \exp\{\eta T(x_n) - A(\eta)\}$$
$$= \left(\prod_{i=1}^{n} h(x_i)\right)\exp\left\{\eta \sum_{i=1}^{n} T(x_i) - nA(\eta)\right\}$$

$E(T(x_i)) = \dot{A}(\eta)$

Dem: Pelo corolario 2.31 do teorema 2.31 pag 122 Bickel.

$\hat{\eta}(x_1, \dots, x_n) \exists$ se $\frac{1}{n}\sum_{i=1}^{n} T(x_i) = \bar{T}_n \in \text{int}(C_T^\circ)$ onde C_T° é o suporte convexo da distribuição de $\bar{T}_n$.

Se η_0 é o verdadeiro valor de η, ainda pelo Teorema 2.3.1 $E_{\eta_0}(T(X_1)) \in$ interior de suporte convexo C_T° pa q' η_0 é solução da equação $\dot{A}(\eta) = t_0$, onde $t_0 = \dot{A}(\eta_0) = E_{\eta_0}(T(X_1))$

Se $E_{\eta_0}(T(X_1))$ está no interior de C_T°, por definição de interior de suporte convexo, resulta q' $\exists$ $S_\delta \subseteq C_T^\circ$.

$$S_\delta = \{t : |t - E_{\eta_0}(T(X_1))| < \delta\}$$

Pela lei fraca, $\frac{1}{n}\sum_{i=1}^{n} T(X_i) \xrightarrow[n\to\infty]{P_{\eta_0}} E_{\eta_0}(T(X_1))$

Então $P_{\eta_0}\left\{\frac{1}{n}\sum_{i=1}^{n} T(X_i) \in C_T^\circ\right\} \to 1$

Mas, $\hat{\eta}$ é solução de $\dot{A}(\eta) = \frac{1}{n}\sum_{i=1}^{n} T(X_i)$; logo (i) é demonstrado.

Mostramos com anteceder cia, no teorema 2.31 q' no espaço C_T° a aplicação $\eta \to \dot{A}(\eta)$ e' contínua e 1:1 em $\mathcal{E}$.

logo $\dot{A}^{-1} : \dot{A}(\mathcal{E}) \to \mathcal{E}$ e' contínua em cada bola aberta S_δ do convexo C_T°. Aplicando a proposição anterior $g = \dot{A}$, $h = \dot{A}^{-1}$

Resulta a consistência de $\hat{\eta}$ a η.

$$\dot{A}(\eta) = \frac{1}{n}\sum_{i=1}^{n} T(X_i) \xrightarrow{P_{\eta_0}} E_{\eta_0}(T(X_1)) = \dot{A}(\eta_0)$$

$\hookrightarrow \dot{A}$ cont

<u>Definições</u>: Notação:

Dizemos:

(a) $U_n = o_p(1)$ se $U_n \xrightarrow{P} 0$

(b) $U_n = O_p(1)$ se $\forall \varepsilon > 0\ \exists M < \infty$ tq $\forall n\ P(|U_n| \geq M) \leq \varepsilon$

(c) $U_n = o_p(V_n)$ se $\frac{|U_n|}{|V_n|} = o_p(1)$

(d) $U_n = O_p(V_n)$ se $\frac{|U_n|}{|V_n|} = O_p(1)$

(e) $U_n \asymp_p V_n$ se $U_n = O_p(V_n)$ e $V_n = O_p(U_n)$

<u>Aplicação</u> ① Suponho q' $Z_1, Z_2, \ldots, Z_n$ são iid réplicas de Z e $E(|Z|) < \infty$

Seja $\mu = E(Z)$, então $\bar{Z}_n = \mu + o_p(1)$ pela Lei fraca.

② Se $E(|Z|^2) < \infty$. Então $\bar{Z}_n = \mu + O_p(n^{-1/2})$ pelo teorema c L. e porque convergência em lei $\Rightarrow O_p(1)$

ie $U_n \xrightarrow{\mathcal{L}} U \Rightarrow U_n = O_p(1)$

pag 320 Bickel

Teorema 5.3.4 Suponho $Y_1, \dots, Y_n$ (d-dim vetores) independentes e identicamente distribuidas tais q $E(|Y_1|^2) < \infty$. Se $E(Y_1) = m$, $Var(Y_1) = \Sigma$ e $h: \underset{\text{aberto}}{\mathcal{O}} \subseteq \mathbb{R}^d \to \mathbb{R}^p$; $h = (h_1, \dots, h_p)$ e h diferenciável, $h^{(1)}(m) = \left\| \frac{\partial h_i}{\partial x_j}(m) \right\|_{p \times d}$

Então $h(\bar{Y}) = h(m) + h^{(1)}(m)(\bar{Y} - m) + o_p(n^{-1/2})$ (a)

$$\sqrt{n} \cdot [h(\bar{Y}) - h(m)] \xrightarrow{\mathcal{L}} N(0, h^{(1)}(m) \Sigma [h^{(1)}(m)]^T) \quad (b)$$

dem exercício.

Normalidade assintótica dos estimadores de MV

Teorema 5.3.5: Suponho q' $\mathcal{P}$ representa uma familia exponencial canônica de posto d gerado por T, com $\mathcal{E}$ aberto. Então, se $X_1, \dots, X_n$ é amostra desde $P_\eta \in \mathcal{P}$ e $\hat{\eta}$ é o EMV,

(i) $\hat{\eta} = \eta + \frac{1}{n} \sum_{i=1}^{n} \dot{A}^{-1}(\eta) \{T(X_i) - \dot{A}(\eta)\} + o_{P_\eta}(n^{-1/2})$

(ii) $\mathcal{L}_\eta(\sqrt{n} \cdot (\hat{\eta} - \eta)) \xrightarrow[n \to \infty]{} N_d(0; \ddot{A}^{-1}(\eta))$.

dem: Desde o teorema 5.2.2 resulta q' se $\bar{T} = \frac{1}{n} \sum_{i=1}^{n} T(X_i)$,

$P_\eta(\bar{T} \in \dot{A}(\mathcal{E})) \to 1$ i.e. $P_\eta(\hat{\eta} = \dot{A}^{-1}(\bar{T})) \to 1$.

$\dot{A}(\hat{\eta}) = \bar{T}$ solução

Assumindo h (do teorema 5.3.4) como $\dot{A}^{-1}$

m (" " 5.3.4) como $\dot{A}(\eta)$

e se $t = \dot{A}(\eta)$.

$$\dot{A}^{-1}(\bar{T}) = \dot{A}^{-1}(\dot{A}(\eta)) + (\dot{A}^{-1})^{(1)}(\dot{A}(\eta)) \cdot (\bar{T} - \dot{A}(\eta)) + o_p(n^{-1/2})$$

$$\sqrt{n}\left(\underbrace{\dot{A}^{-1}(\bar{T})}_{\hat{\eta}} - \underbrace{\dot{A}^{-1}(\dot{A}(\eta))}_{\eta}\right) \xrightarrow{\mathcal{L}} N(0, (\dot{A}^{-1})^{(1)}(\dot{A}(\eta)) \Sigma [(\dot{A}^{-1})^{(1)}(\dot{A}(\eta))]^T$$

<u>Resultado</u> Se $h: \mathbb{R}^d \to \mathbb{R}^d$ é 1:1, derivável num entorno de x e $Dh(x) = \left\| \frac{\partial h_i}{\partial x_j}(x) \right\|_{d \times d}$ é não singular, então h^{-1} é derivável num entorno e se $y = h(x)$,

$$D h^{-1}(h(x)) = [D h(x)]^{-1}.$$

Então $D\dot{A}^{-1}(\underset{\dot{A}(\eta)}{\underset{\parallel}{t}}) \overset{(*)}{=} [D\dot{A}(\eta)]^{-1}$ e como $D\dot{A} = \ddot{A}$

(*) aplicando o resultado anterior $h = \dot{A}$.

logo, se $h = \dot{A}^{-1}$ (como antes) $h^{(1)}(m) \overset{m = \dot{A}(\eta)}{=} D\{\dot{A}^{-1}(\dot{A}(\eta))\}$

$\dot{A}^{-1}(t) = \eta$ $= D\{\dot{A}^{-1}(t)\} = [D\dot{A}(\eta)]^{-1}$

$= (\ddot{A}(\eta))^{-1}$

sabemos por (a) teo anterior q' $h(\bar{T}) = h(m) + h^{(1)}(m)(\bar{T} - m) + o_p(n^{-1/2})$

$$\underbrace{\dot{A}^{-1}(\bar{T})}_{\hat{\eta}} = \dot{A}^{-1}(m) + [\ddot{A}(\eta)]^{-1}(\bar{T} - \underset{\sim}{m}) + o_p(n^{-1/2})$$

$$= \eta + \frac{1}{n}\sum_{i=1}^{n} [\ddot{A}(\eta)]^{-1}(T(X_i) - \dot{A}(\eta)) + o_p(n^{-1/2})$$

e vale (i).

Por outra parte $\Sigma = Var(T(X_1)) = \ddot{A}(\eta)$ (corolario 1.6.1)

Então, $h^{(1)}(m)\Sigma[h^{(1)}(m)]^T = [\ddot{A}(\eta)]^{-1} \cdot \Sigma ([\ddot{A}(\eta)]^{-1})^T$

$$= \ddot{A}^{-1}\ddot{A}\ddot{A}^{-1}(\eta) = \ddot{A}^{-1}(\eta)$$

desde (b) teo anterior, $\sqrt{n}(\hat{\eta} - \eta) \xrightarrow[n \to \infty]{d} N(0, \ddot{A}^{-1}(\eta))$ ※