統計数学特別講義 第二 (2025/07/09)

本日の講義内容

- 混合モデルの最尤推定量の一致性
- 主張：$\mathcal{G}$は全有界 + 漸近同等連続 ⇒ $\mathcal{G}$はP-Donsker
- 主張2：P-Donskerの定理

混合モデルの復習

2

X: 観測

Y: 潜在変数, $Y \sim F_0$

$f(x|y)$: $Y=y$ ($y \in \mathbb{R}$) を与えたときの X の p.d.f. (既知)

X の p.d.f.

$$p_{F_0}(x) := \int f(x|y)\, dF_0(y) \qquad (x \in \mathbb{R})$$

目標 X の独立複製 $X_1, X_2, \ldots, X_n$ に基づいて

F_0 を推定

事実

定理 6.15 $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} P$ とする

$\forall \varepsilon > 0$ に対して

$$H_B(\varepsilon, \mathcal{G}, \|\cdot\|_{L_1(P)}) < \infty$$

ならば、$\mathcal{G}$ は P-GC。すなわち

$$\sup_{g \in \mathcal{G}} \left| \frac{1}{n} \sum_{j=1}^{n} g(X_j) - Pg \right| \xrightarrow{a.s.} 0 \quad (n \to \infty)$$

補題6.47 $P \geq 1$, P は $\mathbb{R}$ 上の確率分布, **4**

(Θ, d) はコンパクト距離空間,

$$\mathcal{G} = \{ g_\theta : \mathbb{R} \to \mathbb{R} \,;\, \theta \in \Theta \}$$

は以下をみたす.

- $\forall x \in \mathbb{R}$ において, 写像

$$\Theta \ni \theta \longmapsto g_\theta(x) \in \mathbb{R}$$

は連続

$\{\theta_n\}_{n=0}^{\infty} \subset \Theta$ s.t. $d(\theta_n, \theta_0) \xrightarrow{n \to \infty} 0$

ならば $\lim_{n \to \infty} g_{\theta_n}(x) = g_\theta(x)$ (各 x)

- $G(x) := \sup_{\theta \in \Theta} |g_\theta(x)| \in L_p(P)$

$\Rightarrow$ $H_B(\varepsilon, \mathcal{G}, \|\cdot\|_{L_p(P)}) < \infty$, $(\forall \varepsilon > 0)$.

F_0 の MLE

$$\hat{F}_n^{mle} \in \arg\max_{F \in \mathcal{F}} \hat{P}_n \log p_F$$

ただし、$\hat{P}_n(B) = \frac{1}{n} \# \{ X_j \in B ; j = 1, \dots, n \}$

$(\forall B \in \mathcal{B}(\mathbb{R}))$

$\mathcal{F}$: $\mathbb{R}$ 上の分布関数の全体

$$p_F(x) = \int q(x \mid y) \, dF(y)$$

注意 一般に、$\hat{F}_n^{mle}$ は経験分布関数 $\hat{F}_n$ とは異なる。

6

任意の p.d.f p と $\hat{p}$ の Hellinger の距離を

$$h(p, \hat{p}) = \left\{ \frac{1}{2} \int_{\mathbb{R}} \left\{ \sqrt{p(x)} - \sqrt{\hat{p}(x)} \right\}^2 dm(x) \right\}^{1/2}$$

ただし m は $\mathbb{R}$ 上の Lebesgue 測度

$$p_{\hat{F}^{mec}}(x) = \int k(x|y)\, d\hat{F}^{mec}(x)$$

Stieltjes 積分の意味.

補題 6.49

$$h(p_{\hat{F}^{mec}}, p_{F_0}) \xrightarrow{a.s.} 0 \quad (n \to \infty).$$

証明

関数族

$$\mathcal{G} := \left\{ \frac{2p_F}{p_F + p_{F_0}} \; ; \; F \in \mathcal{F} \right\}$$

すなわち、母数空間 $\mathcal{F}$（分布関数全体）で添え字付けされている関数空間

補題6.47 $\Theta \to \mathcal{F}$ とみる.

$\mathcal{F}$ の距離関数 d をどう取るかは自明ではない！
（ノルム）

$\mathcal{P} \longleftrightarrow \mathcal{M}$ (確率分布全体)

$\{P_n\}_{n=0}^{\infty} \subset \mathcal{M}$ に対して

$$P_n \rightsquigarrow P_0 \ (n\to\infty) \iff d(P_n, P_0) \xrightarrow{n\to\infty} 0$$

となる $\mathcal{M}$ 上の距離関数 d が存在することが知られている。

さらに、$(\mathcal{P}, d)$ はコンパクトであることも知られている。 ⇒ 補題6.42の条件

以後は、記号を乱用して、d を $\mathcal{P}$ 上の距離関数とも思う。

9

$\{F_n\}_{n=0}^{\infty}$ に対して

$$F_n \rightsquigarrow F_0 \ (n \to \infty) \iff d(F_n, F_0) \xrightarrow{n\to\infty} 0$$

さらに、各 x に対して

$$\mathbb{R} \ni \theta \mapsto \theta(x \mid y) \in \mathbb{R}$$

は連続で

$$\lim_{y \to \pm\infty} \theta(x \mid y) = 0$$

より

$\theta(x \mid \cdot)$ は有界連続

である。よって Portmanteau の定理より

各 $x\in\mathbb{R}$ に対して

$$F_n \rightsquigarrow F_0 \ (n\to\infty).$$

$$\Longrightarrow \int_{\mathbb{R}} k(\cdot\,|\,y)\,dF_n(y) \xrightarrow{n\to\infty} \int_{\mathbb{R}} k(\cdot\,|\,y)\,dF_0(y)$$

となる。以上より、

$$\mathcal{F} \ni F \longmapsto p_F(x) = \int_{\mathbb{R}} k(x\,|\,y)\,dF(y)$$

は連続写像である。

関数族 $\mathcal{G} = \left\{ \dfrac{2p_F}{p_F + p_{F_0}} \,;\, F\in\mathcal{F} \right\}$ の包絡関数をとり

$$G(x) := \sup_{F\in\mathcal{F}} \left| \frac{2p_F(x)}{p_F(x) + p_{F_0}(x)} \right| \le 2.$$

よって,

$$G \in L_1(P).$$ ⇒ 補題6.47の条件

さらに, 各 $x \in \mathbb{R}$ に対して, 写像

$$\mathcal{F} \ni F \longmapsto \frac{2P_F(x)}{P_F(x)+P_{F_0}(x)}$$

は連続 ⇒ 補題6.47の条件

また, $\mathcal{F}$ はコンパクトである. ⇒ 補題6.47の条件

補題6.47より, $^\forall \delta > 0$ に対して

$$H_B(\delta, \mathcal{G}, \|\cdot\|_{L_1(P)}) < \infty$$

がわかる. ⇒ $\mathcal{G}$ は P-GC族 (∵) 定理6.4(5)

よって

$$\left| (\hat{P}_n - P) \left(\frac{2 p_{\hat{F}_n^{mec}}}{p_{\hat{F}^{mec}} + F_{F_0}} \right) \right|$$

$$\leq \sup_{g \in \mathcal{G}} \left| (\hat{P}_n - P) g \right| \xrightarrow{a.s.} 0 \quad (n \to \infty)$$

ただし

$$\mathcal{G} = \left\{ \frac{2 p_F}{p_F + p_{F_0}} \,;\, F \in \mathcal{F} \right\}.$$

一方

$$1 - P \left(\frac{2 p_{\hat{F}_n^{mec}}}{p_{\hat{F}_n^{mec}} + p_{F_0}} \right) \leq \left| (\hat{P}_n - P) \left(\frac{2 p_{\hat{F}_n^{mec}}}{p_{\hat{F}_n^{mec}} + p_{F_0}} \right) \right|$$

I 7:

$$h^2\left(p_{\hat{F}_n^{mee}}, p_{F_0}\right) \le 1 - P\left(\frac{2\, p_{\hat{F}^{mee}}}{p_{\hat{F}_n^{mee}} + p_{F_0}}\right)$$

$$\le \sup_{g \in \mathcal{G}} \left| (\hat{P}_n - P)\, g \right| \xrightarrow{a.s.} 0 \quad (n \to \infty).$$

$\square$

8.1 距離空間における弱収束

$(\mathbb{D}, d)$: 距離空間

$\Leftrightarrow {}^\forall x, y, z \in \mathbb{D}$ に対し、① $d(x,y) \geq 0$, ② $d(x,y)=0 \Leftrightarrow x=y$

③ $d(x,y)=d(y,x)$, ④ $d(x,z) \leq d(x,y)+d(y,z)$.

$(\Omega, \mathcal{A}, P_r)$: 確率空間

$X_n : \Omega \to \mathbb{D}$ $(n=1,2,\dots)$ は確率要素

$\Leftrightarrow B \in \mathcal{B}(\mathbb{D})$ に対し、$X_n^{-1}(B) \in \mathcal{A}$.

<u>注意</u> 可測でない場合は気にしないことにする.

$E^* \to E$ をすべてかく.

定義 8.1 $\{X_n\}, X$ は確率要素列

$X_n \rightsquigarrow X \quad (n \to \infty)$

$\Leftrightarrow {}^\forall h \in C_b(\mathbb{D})$ に対し

$$E[h(X_n)] \xrightarrow{n\to\infty} E[h(X)].$$

ただし、$C_b(\mathbb{D})$ は $\mathbb{D}$ 上の $\mathbb{R}$ 値有界連続関数の全体

定義 8.2

$X_n \xrightarrow{P} X \quad (n \to \infty)$

$\Leftrightarrow {}^\forall \varepsilon > 0$ に対し、 $\Pr(d(X_n, X) > \varepsilon) \xrightarrow{n\to\infty} 0.$

注意

$X_n \xrightarrow{d} X \ (n \to \infty)$

$\iff \liminf \Pr(X_n \in G) \geq \Pr(X \in G) \quad (\forall G : \text{open})$

$\iff \limsup \Pr(X_n \in F) \leq \Pr(X \in F) \quad (\forall F : \text{closed})$

$\iff \lim \Pr(X_n \in B) = \Pr(X \in B) \quad (\forall B : \text{Borel s.t. } \Pr(\partial B) = 0)$

$\iff$ 任意の有界 Lipshitz に対して

$$E[h(X_n)] \xrightarrow{n \to \infty} E[h(X)].$$

記号

$\mathcal{X}$: 標本空間

P: $\mathcal{X}$ 上の確率分布

$\mathcal{G} \subset L_2(P)$: $\mathcal{X}$ 上の実数値関数の族

$X_1, \ldots, X_n$: X の独立 copies

$$\hat{P}_n(B) = \frac{1}{n} \# \{ X_j \in B ; j = 1, \ldots, n \} \quad (\forall B \in \mathcal{B}(\mathcal{X}))$$

$$\nu_n(g) := \sqrt{n} \int g \, d(\hat{P}_n - P)$$

$$= \sqrt{n} \left\{ \frac{1}{n} \sum_{j=1}^{n} g(X_j) - Pg \right\}$$

$$= \sqrt{n} \left\{ \frac{1}{n} \sum_{j=1}^{n} g(X_j) - E[g(X)] \right\} \quad (\forall g \in \mathcal{G})$$

$\{\nu_n(g) : g \in \mathcal{G}\}$: 経験過程 線型ではないけど

$\ell^\infty(\mathcal{G})$ $\mathcal{G}$ 上の 有界 汎関数 全体 の集合.

$$\|f\|_\infty := \sup_{g \in \mathcal{G}} |f(g)| \qquad (\forall f \in \ell^\infty(\mathcal{G}))$$

$\nu_n \in \ell^\infty(\mathcal{G})$ をみたすと仮定

$$\|g\|_{L_2(P)} := \sqrt{\int g^2 dP} \qquad (g \in \mathcal{G}).$$

$\{\nu : g \in \mathcal{G}\}$: 中心化された Gauss過程で

$$\mathrm{Cov}(\nu(g_1), \nu(g_2)) = \int g_1 g_2 dP - \int g_1 dP \int g_2 dP$$

$$(g_1, g_2 \in \mathcal{G}). \qquad (A)$$

なので、

$$\begin{bmatrix} \nu(g_1) \\ \nu(g_2) \end{bmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \mathrm{Var}[\nu(g_1)] & \mathrm{Cov}[\nu(g_1), \nu(g_2)] \\ * & \mathrm{Var}[\nu(g_2)] \end{pmatrix} \right)$$

定義 6.15 $\{\nu_n : g \in \mathcal{G}\}_{n=1}^{\infty}$, $\{\nu : g \in \mathcal{G}\}$ は Gauss過程で (A) をみたすとする.

$\forall f : \ell^{\infty}(\mathcal{G}) \longrightarrow \mathbb{R}$ に対して

$$E[f(\nu_n)] \xrightarrow{n \to \infty} E[f(\nu)]$$

のとき、$\mathcal{G}$ は P-Donsker 族という

定理8.19 $\mathcal{X}$を標本空間，Pを$\mathcal{X}$上の確率測度

$\mathcal{G}$を$\mathcal{X}$上の実数値関数のある族:

$\{\nu_n : g \in \mathcal{G}\}$: 関数族$\mathcal{G}$で添え字付けられた

確率過程

$\nu_n : \Omega \times \mathcal{G} \longrightarrow \nu_n(\omega, g)$ と書いて、

$$\Pr\left(\omega \in \Omega ;\ \underbrace{\nu_n(\omega, \cdot)}_{\nu_n\text{の見本路}} \in \underbrace{\ell^\infty(\mathcal{G})}_{\mathcal{G}\text{上の有界な実数値関数の全体}}\right) = 1$$

とし，以下の2つの条件をみたす.

- 関数族$\mathcal{G} \subset L_2(P)$は全有界

・$\forall \eta > 0$ に対して，ある $\delta > 0$ が存在し

$$\limsup_{n\to\infty} \Pr\left(\sup_{\substack{g_1, g_2 \in \mathcal{G} \\ \sqrt{P|g_1 - g_2|^2} \le \delta}} |\nu_n(g_1) - \nu_n(g_2)| > \eta \right) < \eta.$$

このとき，関数族 $\mathcal{G}$ は P-Donsker 族．

証明の概略

関数族 $\mathcal{G}$ は全有界なので $^\forall\delta>0$ に対して

有限部分集合 $\mathcal{G}_\delta \subset \mathcal{G}$ が存在して,

$^\forall g \in \mathcal{G}$ に対して, $^\exists g_\delta \in \mathcal{G}_\delta$ s.t.

$$\|g - g_\delta\|_{L_2(P)} = \sqrt{P|g - g_\delta|^2} < \delta$$

とできる.

$k = \#(\mathcal{G}_\delta)$ とおく.

$$\pi_\delta : \mathcal{G} \ni g \longrightarrow g_\delta \in \mathcal{G}_\delta$$

とする.

<u>証明の目標</u> 任意の有界Lipschitz $f : \ell_\infty(\mathcal{G}) \to \mathbb{R}$ に対して

$$E[f(X_n)] \xrightarrow{n\to\infty} E[f(X)]$$

を示せばよい

① $\mathcal{G}_\delta = \{g_1, g_2, \dots, g_k\}$ ($k = \#(\mathcal{G}_\delta)$)

と書いておくと

$$\{\nu_n \circ \pi_\delta(g) : g \in \mathcal{G}\} \subset \{\nu_n(g_1), \dots, \nu_n(g_n)\}$$

$$\{\nu \circ \pi_\delta(g) : g \in \mathcal{G}\} \subset \{\nu(g), \dots, \nu(g)\}$$

となる. よって, CLTより,

$$E[h(\nu_n \circ \pi_\delta)] \xrightarrow{n\to\infty} E[h(\nu \circ \pi_\delta)]$$ 24

② $\tilde{\nu}$ を ν の連続修正過程とする.

$E[h(\tilde{\nu})] = E[h(\nu)]$ なので

$$E[h(\nu_n)] - E[h(\nu)]$$

$$= E[h(\nu_n)] - E[h(\tilde{\nu})]$$

$$= E[h(\nu_n)] - E[h(\nu_n \circ \pi_\delta)]$$

$$+ E[h(\nu_n \circ \pi_\delta)] - E[h(\tilde{\nu} \circ \pi_\delta)]$$

$$+ E[h(\tilde{\nu} \circ \pi_\delta)] - E[h(\tilde{\nu})]$$

$=: A + B + C$

25

項Aの評価

$$E := \left\{ \sup_{\substack{g_1, g_2 \in \mathcal{G} \\ \|g_1 - g_2\|_{L_2(P)} \leq \delta}} |\nu_n(g_1) - \nu_n(g_2)| \leq \eta \right\}$$

とおく　すると

$$\left| E[h(\nu_n)] - E[h(\nu_n \circ \pi_\delta)] \right|$$

$$= \left| \mathbb{E}\left[\{ h(\nu_n) - h(\nu_n \circ \pi_\varepsilon) \} \mathbb{1}_E \right] \right. \qquad \text{26}$$

$$+ \left. \mathbb{E}\left[\{ h(\nu_n) - h(\nu_n \circ \pi_\varepsilon) \} \mathbb{1}_{E^c} \right] \right|$$

$$\leq \mathbb{E}\left[| h(\nu_n) - h(\nu_n \circ \pi_\delta) | \mathbb{1}_E \right]$$

$\leq \lambda \eta$ $\quad \eta$ is Lipschitz constant

$E \; \text{I} \; \geq n$

$|\nu_n(q_1) - \nu_n(q_2)| \leq \eta$

provided $\rho|q_1 - q_2| \lesssim \delta$

$\rho|\nu_n(q) - \nu \circ \pi_\delta(q)| \lesssim \delta$

$$+ 2 \| h \|_\infty \mathbb{E}\left[\mathbb{1}_{E^c} \right]$$

$$\leq \lambda \eta \, \mathbb{E}[\mathbb{1}_E] + 2 \|h\|_\infty \underbrace{\mathbb{E}[\mathbb{1}_{E^c}]}_{\downarrow_{n \to \infty} 0}$$

$$\xrightarrow{n \to \infty} \lambda \eta$$

Cの評価も同様にできる.

27 //

定理 6.23 $\mathcal{X}$を標本空間, Pを$\mathcal{X}$上の確率測度, $\mathcal{G}$を$\mathcal{X}$上の実数値関数のある族で $\mathcal{G} \subset L_2(P)$.

$$G(x) := \sup_{g \in \mathcal{G}} |g(x)|$$

とする. さらに

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon \|G\|_{L_2(Q)}, \mathcal{G}, \|\cdot\|_{L_2(Q)})}\, d\varepsilon$$

$$< \infty$$

をみたすとする。ただし、Qは $\|G\|_{L_2(Q)} < \infty > 0$

なる任意の有限離散確率測度である。

このとき、$PG^2 < \infty$ と適当な可測性条件のもと、

$\mathcal{G}$ は P-Donsker 族

証明の方針

I　ν_n は漸近同等連続

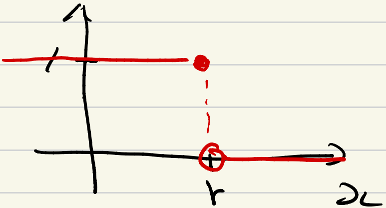
II　$\mathcal{G} \subset L_2(P)$ は全有界

で示せばよい。 //

例 8.24

$$\mathcal{G} = \{ \mathbb{1}_{(-\infty, r]}(x) ; r \in \mathbb{R} \}$$

とおく



$\mathbb{R}$上の任意の離散有限確率測度 Q と

$1 \ge \varepsilon > 0$ に対して

$$N(\varepsilon, \mathcal{G}, L_2(Q)) \le N_B(\varepsilon^2, \mathcal{G}, L_1(Q)) \le \frac{2}{\varepsilon^2}$$

31

$\delta > 2$, $G = 1 \rightsquigarrow \|G\|_{L_2(Q)} = 1 \le a$?.

$$\int_0^1 \sqrt{\log N(\varepsilon, S, L_2(Q))}\, d\varepsilon$$

$$\le \int_0^1 \sqrt{\log\left(\frac{2}{\varepsilon^2}\right)}\, d\varepsilon \qquad \log u \le u - 1$$

$$\le \int_0^1 \sqrt{2\log\left(\frac{2}{\varepsilon}\right)}\, d\varepsilon$$

$$\le \sqrt{2}\int_0^1 \log\left(\frac{2}{\varepsilon}\right) d\varepsilon < \infty \qquad //$$

Cauchy-Schwarz