統計数学特別講義第二　(2025/06/07)

本日の講義内容

- 経験過程の収束の導入
- sup の可測性についての注意
- 経験過程の例
- Kullback-Leibler 偏差 (KL 偏差)
- 計量エントロピー, 被覆数と ε 網

6.1 導入

$X, X_1, X_2, \ldots, X_n$: i.i.d. 確率変数列

仮定: $E[X_1^2] < \infty$.

$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ に対して, SLLN と CLT を適用すると

- $\overline{X}_n \xrightarrow{a.s} E[X] \quad (n \to \infty)$

- $\sqrt{n}\,(\overline{X}_n - E[X]) \rightsquigarrow N(0, \sigma^2) \quad (n \to \infty)$

ただし, $\sigma^2 = Var[X]$.

$g:\mathbb{R}\to\mathbb{R}$ を可測関数で，$0 < E[g^2(X)] < \infty$ と[3]

したとき

・ $$\frac{1}{n}\sum_{j=1}^{n} g(X_j) \xrightarrow{a.s.} E[g(X)] \quad (n\to\infty)$$

・ $$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} \left\{ g(X_j) - E[g(X)] \right\} \rightsquigarrow N(0, \sigma_g^2) \quad (n\to\infty)$$

ただし，$0 < \sigma_g^2 = Var[g(X)] < \infty$.

統計的推測理論の文脈では，関数 g 自身がランダムなものとなる．すなわち，

g の推定量 $\hat{g}(\cdot) = \hat{g}(\cdot \mid X_1, \ldots, X_n)$ になする

経験過程

$$\frac{1}{n}\sum_{j=1}^{n} \hat{g}(X_j)$$

を考察することになる。このために、$\mathcal{G}$ をある関数族とし、

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

の一様な評価が必要になる。

6.1.1 「sup」の可測性についての注意

$\mathcal{G}$ をある関数族としたとき、

$$\sup_{g \in \mathcal{G}} g(X)$$

の挙動を調べたいとする。すると

$$E\left[\sup_{g \in \mathcal{G}} g(X)\right]$$

の数学的な厳密な意味が問題になる。

$\mathcal{G}$ が複雑だと $\sup_{g \in \mathcal{G}} g(x)$ の可測性が保障されない。たとえば、$\mathcal{G}$ の有界関数の全体。

2つのアプローチ

(1) $E^*\left[\sup_{g\in\mathcal{G}} g(X)\right] := \sup_{\tilde{\mathcal{G}}\subset\mathcal{G}} \left\{ E\left[\sup_{g\in\tilde{\mathcal{G}}} g(X)\right] ; \right.$

$\left. \tilde{\mathcal{G}}\text{は有限集合} \right\}.$

(2) $E^*\left[\sup_{g\in\mathcal{G}} g(x)\right] := \inf \left\{ E[U] ; \ U:\Omega\to\mathbb{R}\text{は} \right.$

$\left. U \geq \sup_{g\in\mathcal{G}} g(x)\text{なる確率変数} \right\}.$

以降では，可測性は気にせずに議論を進める.

6.2.1 教育と雇用

標本： $X_j = (Y_j, Z_j)$ $(j=1,2,\dots,n)$

ただし、$Y_j \in \{0,1\}$, $Z_j \in \mathbb{R}$ （本当は 非負）

$\uparrow$ 雇用の有無 $\quad\quad$ $\uparrow$ 教育年限

知りたいこと

$$g^*(z) = E[\, Y=1 \mid Z = z \,] \longleftarrow$$

標本 $\{(Y_j, Z_j)\}_{j=1}^{n}$ から推定

ただし、Y と Z は generic な r.v. とした。

統計的モデル

$$\Delta_1 := \{ g : \mathbb{R} \to (0,1),\ g \text{は非減少} \}$$

関数 g^* の自然な推定量（最尤推定量）

$$\hat{g}_n \in \underset{g \in \Delta_1}{\arg\min} \left[\sum_{j=1}^{n} \{ Y_j \log g(Z_j) + (1-Y) \log (1 - g(Z_j)) \} \right]$$

$\because)$ $\Pr(Y=y \mid Z=z) = \{ g^*(z) \}^y \{ 1 - g^*(z) \}^{1-y}$.

$(y = 0, 1)$

QをZの分布とする。すなわち

$$Q(B) = P_r(Z \in B) \quad ({}^\forall B \in \mathcal{B}(\mathbb{R}))$$

としたとき、推定量$\hat{g}_n$の精度の評価尺度として

$$\|\hat{g}_n - g^*\|_{L_2(Q)} = \left(\int_{\mathbb{R}} \{\hat{g}_n(z) - g^*(z)\}^2 \, dQ(z) \right)^{1/2}$$

を考えると

$$\|\hat{g}_n - g^*\|_{L_2(Q)} = O_p\left(\frac{1}{n^{1/3}}\right)$$

となることが知られている。

別の統計的モデル

$$\Lambda_2 := \left\{ g: \mathbb{R} \to (0,1) \;;\; 0 \leq \frac{dg}{dz} \leq M,\ g \text{ は concave} \right\}$$

ただし、Mは定数

$$\| \hat{g}_n - g^* \|_{L_2(Q)} = O_p\left(\frac{1}{n^{2/5}}\right)$$

$$\Lambda_3 := \left\{ g: \mathbb{R} \to (0,1) \;;\; g(z) = g^*(\theta z),\ \theta \in \mathbb{R} \right.$$

$$\left. g^*(z) = \frac{e^z}{1+e^z} \right\}.$$

$$\| \hat{g}_n - g_* \|_{L_2(Q)} = O_p\left(\frac{1}{\sqrt{n}}\right).$$

統計学では，伝統的に，

Λ_1, Λ_2: nonparametric models（無次元母数空間）

Λ_3: parametric model（有限次元母数空間）

用語の注意（雑な説明）

P^*: 真の分布（データ生成分布）

$\mathcal{P}$: 統計的モデル（確率分布の集まり）

$\mathcal{P}$ の要素を添え字付ける集合を母数空間

$\mathcal{P} \ni P \longleftrightarrow \theta \in \Theta$

6.2.2 Kullback-Leibler 偏差

$\mathcal{X}\ (\subset \mathbb{R})$：標本空間

m： $\mathbb{R}$ 上の Lebesgue 測度

$\{P_\theta : \theta \in \Theta\}$：統計的モデル.

ただし、P_θ は Lebesgue 測度 m に関する p.d.f. で

Θ を母数空間

P_{θ^*}：真の分布，　$\theta^* \in \Theta$

Hellinger 距離

$$h(p, q) = \left(\frac{1}{2} \int_{\mathcal{X}} \{ \sqrt{p} - \sqrt{q} \}^2 \, dm \right)^{1/2}$$

ただし. p, q は $\mathcal{X}$ 上の m に関する p.d.f.

Kullback-Leibler 偏差

$$KL(p, q) = \int_{\mathcal{X}} \log\left(\frac{p(x)}{q(x)} \right) p(x) \, dm(x)$$

ただし、簡単のために. $p(x) > 0$, $q(x) > 0$ $(\forall x \in \mathcal{X})$ と仮定する.

命題6.2

(1) $KL(p, q) \geq 0$.

(2) $h^2(p, q) \leq \frac{1}{2} KL(p, q)$.

証明　$v > 0$に対して、不等式

$$\log v \leq v - 1 ; \quad \frac{1}{2}\log v \leq \sqrt{v} - 1$$

に注意する。

(1) の証明

$$KL(p, q) = \int \log\left(\frac{p(x)}{q(x)}\right) p(x)\, dm(x)$$

$$= -\int \log\left(\frac{q(x)}{p(x)}\right) p(x)\, dm(x)$$

$$\geq -\int \left\{\frac{q(x)}{p(x)} - 1\right\} p(x)\, dm(x)$$

$$= -\int q(x)\, dm(x) + \int p(x)\, dm(x) = 0 \quad //$$

(2)の証明

$$\frac{1}{2}KL(p,q) = \int \frac{1}{2}\log\left(\frac{p(x)}{q(x)}\right) p(x)\,dm(x)$$

$$= -\int \frac{1}{2}\log\left(\frac{q(x)}{p(x)}\right) p(x)\,dm(x)$$

$$\geq -\int \left\{ \sqrt{\frac{q(x)}{p(x)}} - 1 \right\} p(x)\,dm(x)$$

$$= \underbrace{\int p(x)\,dm(x)}_{=1} - \int \sqrt{p(x)q(x)}\,dm(x)$$

$$= \frac{1}{2}\left\{ \underbrace{\int p(x)\,dm(x)}_{=1} + \underbrace{\int q(x)\,dm(x)}_{=1} - \int 2\sqrt{p(x)q(x)}\,dm(x) \right.$$

$$= \frac{1}{2}\int \left\{ \sqrt{p(x)} - \sqrt{q(x)} \right\}^2 dm(x)$$

$$= h^2(p, q).$$

//

$\theta^* \in \Theta$ を真の母数とし

$$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} P_{\theta^*}$$

とし、P_{θ^*} の最尤推定量(MLE)は

$$\hat{p}_{\hat{\theta}_n} \in \underset{\theta \in \Theta}{\arg\min} \left\{ \sum_{j=1}^{n} \log \left(\frac{p_{\theta^*}(x_j)}{p_{\theta}(x_j)} \right) \right\}$$

$$\Longleftrightarrow$$

$$\hat{p}_{\hat{\theta}_n} \in \underset{\theta \in \Theta}{\arg\max} \left\{ \sum_{j=1}^{n} \log p_{\theta}(x_j) \right\}.$$

Kullback-Leibler の経験過程

MLEの定義

$$0 \geq \frac{1}{n}\sum_{j=1}^{n} \log P_{\theta^*}(x_j) - \frac{1}{n}\sum_{j=1}^{n} \log P_{\hat{\theta}_n}(x_j)$$

$$= \frac{1}{n}\sum_{j=1}^{n} \log\left(\frac{P_{\theta^*}(x_j)}{P_{\hat{\theta}_n}(x_j)}\right)$$

$$= \frac{1}{n}\sum_{j=1}^{n} \log\left(\frac{P_{\theta^*}(x_j)}{P_{\hat{\theta}_n}(x_j)}\right) - KL(P_{\theta^*}, P_{\hat{\theta}_n}) + KL(P_{\theta^*}, P_{\hat{\theta}_n})$$

と書き直せる. よって

$$a - b + c \leq 0 \ , \quad c \geq 0$$

$$\curvearrowright \ a - b \leq -c \leq 0$$

$$|a-b| \geq c$$

$$KL(P_{\theta^*}, P_{\hat{\theta}_n}) \leq \left| \frac{1}{n}\sum_{j=1}^{n} \log\left(\frac{P_{\theta^*}(X_j)}{P_{\hat{\theta}_n}(X_j)}\right) - KL(P_{\theta^*}, P_{\hat{\theta}_n}) \right|$$

$$\leq \sup_{\theta \in \Theta} \left| \frac{1}{n}\sum_{j=1}^{n} \log\left(\frac{P_{\theta^*}(X_j)}{P_{\theta}(X_j)}\right) - KL(P_{\theta^*}, P_{\theta}) \right|$$

を得る.

任意の固定した $\theta \in \Theta$ に対して

$$\frac{1}{n}\sum_{j=1}^{n} \log\left(\frac{P_{\theta^*}(X_j)}{P_{\theta}(X_j)}\right) - KL(\theta^*, \theta) = O_p\left(\frac{1}{\sqrt{n}}\right).$$

だが, 一様な評価が必要.

6.3 計量エントロピー, 被覆数 と ε 網

$(\mathbb{D}, d)$: 擬距離空間. すなわち

① $d(x, y) \geq 0$; ② $d(x, y) = d(y, x)$; ③ $d(x, z) \leq d(x, y) + d(y, z)$

$({}^\forall x, y, z \in \mathbb{D})$

注意 「$d(x, y) = 0 \Leftrightarrow x = y$」は成立しなくともよい.

${}^\forall x \in \mathbb{D}$ と $\varepsilon > 0$ に対して

$$B_d(x, \varepsilon) = \{ y \in \mathbb{D} ; d(x, y) < \varepsilon \}$$

と定める. これを ε 球とよぶ.

定義 6.4 $(\mathbb{D}, d)$を擬距離空間とし，$A \subset \mathbb{D}$を空でない部分集合とする．

- Aの半径εの被覆とは，ε球の有限集合$\mathcal{C}$で

$$\bigcup_{C \in \mathcal{C}} C \supset A$$

となるもの．

- Aのε被覆集合の全体を$\mathrm{Cover}(A, \varepsilon)$と記す．
- ε被覆$\mathcal{C}$の中心を$\mathrm{Centers}(\mathcal{C}, \varepsilon)$と記す．

(1) Aに対する被覆数 $N(\varepsilon, A, d)$ を

$$N(\varepsilon, A, d) = \min_{\mathcal{C} \in \mathrm{Cover}(A, \varepsilon)} \#\left(\mathrm{Centers}(\mathcal{C}, \varepsilon)\right).$$

ただし、$\#(\cdot)$ は $(\cdot)$ の要素の個数

(2) $H(\varepsilon, A, d) = \log N(\varepsilon, A, d)$ を A の ε エントロピーという

(3) Aは全有界 $\iff$ $H(\varepsilon, A, d) < \infty$.

6.3.1 関数族のエントロピー

Q: $\mathbb{R}$ 上の確率測度. すなわち, $Q(\mathbb{R})=1$.

$\mathcal{G}$: $\mathbb{R}$ 上の実数値可測関数の族

$\mathcal{G}$ 上の距離 d を

$$d(f,g)=\|f-g\|_{L_p(Q)}=\left(\int_{\mathbb{R}}|f-g|^p\,dQ\right)^{1/p}\quad (f,g\in\mathcal{G}).$$

で定める

ε エントロピー $H(\varepsilon,\mathcal{G},d)$ を $H_p(\varepsilon,\mathcal{G},Q)$ または

$H(\varepsilon,\mathcal{G},\|\cdot\|_{L_p(Q)})$ とも書くことにする.

通常

$$\mathcal{G} \subset L_p(Q) := \{ g : \mathbb{R} \to \mathbb{R} ; \int |g|^p \, dQ < \infty \}$$

を考える

$\|g\|_\infty := \sup_{x \in \mathbb{R}} |g(x)|$ に関するエントロピーを $H_\infty(\varepsilon, \mathcal{G})$

と書く.

定義6.6 $\varepsilon > 0$ に対して，$N_{p,B}(\varepsilon, \mathcal{G}, Q)$ を次の条件をみたす関数の組 $\{(g_j^L, g_j^R)\}_{j=1}^N$ の最小の個数とする．

- $^\forall j \in \{1, \ldots, N\}$ に対して

$$\| g_j^R - g_j^L \|_{L^p(Q)} \leq \varepsilon.$$

- $^\forall g \in \mathcal{G}$ に対して，$^\exists j_0 \in \{1, \cdots, N\}$ s.t.

$$g_j^L(x) \leq g(x) \leq g_j^R(x) \qquad (x \in \mathbb{R}).$$

$H_{p,B}(\varepsilon, \mathcal{G}, Q) = \log N_{p,B}(\varepsilon, \mathcal{G}, Q)$ と記す．

命題6.7

(1) $H_p(\varepsilon, \mathcal{G}, Q) \leq H_{p,B}(\varepsilon, \mathcal{G}, Q)$.

(2) $H_{p,B}(\varepsilon, \mathcal{G}, Q) \leq H_\infty\left(\frac{\varepsilon}{2}, \mathcal{G}\right)$

(3) $d(x, y) \leq d'(x, y) \ (\forall x, y \in D)$

$\Rightarrow H(\varepsilon, \mathcal{G}, d) \leq H(\varepsilon, A, d')$.

6.3.2 ε網

定義6.8 (D, d)を擬距離空間とし、$\|x\| = \sqrt{d(x,x)}$ $(x \in D)$ と記す。$\varepsilon > 0$とする。空でない部分集合$A\ (\subset D)$のε網とは Aの有限部分集合$\{c_1, c_2, \ldots, c_N\}$で以下をみたすもの である

- $\forall j, k \in \{1, \ldots, N\}$, $j \neq k$ に対して、$\|c_j - c_k\| \geq \varepsilon$.
- 集合$\{c_1, \ldots, c_N\}$は包含関係による順序に関して 最大

命題6.9 部分集合 $A \subset \mathbb{D}$ の ε 網 $\{c_1, \ldots, c_N\}$ は

A の被覆 $\mathcal{C}$ の中心とすることができる.

証明 $j \in \{1, \ldots, N\}$ とし

$$B_j := \{x \in \mathbb{D} ; \|x - c_j\| < \varepsilon\}$$

とする. すると

$$\bigcup_{j=1}^{N} B_j \supset A$$

となることを背理法を用いて示そう.

そのために. ある $c_0 \in A$ が存在して, $c_0 \notin \bigcup_{j=1}^{N} B_j$ と仮定する.

すると $\{C_1, \dots, C_N, C_0\}$ も部分集合 A の ε 網となり、

$\# \{C_1, \dots, C_N, C_0\} = N+1$ となるので、$\{C_1, \dots, C_N\}$

が A の ε 網の仮定に反する。 $\square$

補題 6.10 d を $\mathbb{R}^d$ の Euclid 距離とし、$O_d \in \mathbb{R}^d$ を原点とし、$R > 0$ を取る。$A = B_d(O_d, R) \subset \mathbb{R}^d$ に対して

$$N(\varepsilon, A, d) \leq \left(\frac{2R + \varepsilon}{\varepsilon}\right)^d$$

が成立する。

証明 $\{C_j\}_{j=1}^N$ を A の ε 組とすると、これを中心にして A の被覆ができるので

$$N(\varepsilon, A, d) \leq N$$

である．また

$$\bigcup_{j=1}^{N} B_d(c_j, \varepsilon) \supset A$$

である．また

$$\bigcup_{j=1}^{N} B_d\left(c_j, \frac{\varepsilon}{2}\right) \subset B_d\left(0_d, R+\frac{\varepsilon}{2}\right) \qquad (*)$$

である．$j \neq k$ に対して，$\|c_j - c_k\| \geq \varepsilon$ なので．

$$B_d\left(c_j, \frac{\varepsilon}{2}\right) \cap B_d\left(c_k, \frac{\varepsilon}{2}\right) = \phi$$

となる．$(*)$ より

$$N \,\mathrm{vol}(B_d(0_d, 1))\left(\frac{\varepsilon}{2}\right)^d \leq \mathrm{vol}(B_d(0_d, 1))\left(R+\frac{\varepsilon}{2}\right)^d$$

がわかる. ただし vol(・)は(・)の体積である.

両辺 $\mathrm{vol}(B_d(0_d, 1))\left(\frac{\varepsilon}{2}\right)^d$ でわると

$$N \leq \left(\frac{2R+\varepsilon}{\varepsilon}\right)^d$$

//.

例 6.12

$$\mathcal{G} := \left\{ g : [0,1] \to [0,1] \text{ s.t. } \sup_{s \in [0,1]} |\dot{g}(\tau)| \le 1 \right\}$$

すると

$$H_\infty(\varepsilon, \mathcal{G}) \lesssim \frac{1}{\varepsilon}$$

例 6.14

$$\mathcal{G} := \{ g : [0,1] \to \mathbb{R} \text{ は Lipschitz } \}$$

$\Leftrightarrow$ $\exists L > 0$ s.t. $|g(x) - g(y)| \le L \cdot |x - y|$ $(\forall x, y \in [0,1])$

すると

$$H_\infty(\varepsilon, \mathcal{G}) \lesssim \log\left(1 + \frac{1}{\varepsilon}\right) + \frac{1}{\varepsilon}$$

これが消す。

6.4 括弧付きエントロピーによる Glivenko-Cantelli 族の定理

定義 6.15 P を $\mathbb{R}$ 上の確率測度とし、

$$X, X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} P$$

とする。関数族 $\mathcal{G} = \{g : \mathbb{R} \to \mathbb{R}\}$ は P-Glivenko-Cantelli 族 (P-GC 族) であるとは

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

をみたすときをいう。

定理 6.16 $^\forall \varepsilon > 0$ に対して

$$H_{1,B}(\varepsilon, \mathcal{G}, P) < \infty \Longrightarrow \mathcal{G} \text{ は } P\text{-GC族}$$

証明 $\varepsilon > 0$ を固定。$N := N_{1,B}(\varepsilon, \mathcal{G}, P) < \infty$ とおく。

すると関数の組 $\{(g_j^L, g_j^R)\}_{j=1}^N$ が存在して

$$\max_{j \in \{1,\dots,N\}} \sup_{x \in \mathbb{R}} |g_j^R(x) - g_j^L(x)| \leq \varepsilon.$$

さらに、$^\forall g \in \mathcal{G}$ に対して、$^\exists j_0 \in \{1, \dots, N\}$ s.t.

$$g_{j_0}^L(x) \leq g(x) \leq g_{j_0}^R(x) \qquad (x \in \mathbb{R}).$$

したがって

$$\int g \, d(\hat{P}_n - P) = \int g \, d\hat{P}_n - \int g \, dP$$

$$\leq \int g_{j_0}^R \, d\hat{P}_n - \int g \, dP$$

$$= \int g_{j_0}^R \, d(\hat{P}_n - P) + \int \underbrace{(g_{j_0}^R - g)}_{-\varepsilon \leq (\cdots) \leq \varepsilon} dP$$

$$\leq \int g_{j_0}^R \, d(\hat{P}_n - P) + \varepsilon$$

同様に

$$\int g \, d(\hat{P}_n - P) \geq \int g_{j_0}^L \, d(\hat{P}_n - P) - \varepsilon$$

となる.

$\{(g_j^L, g_j^R)\}_{j=1}^N$ は有限個なので、SLLN から

$$\max_{j\in\{1,\dots,N\}} \left| \int g_j^L \, d(\hat{P}_n - P) \right| \xrightarrow{a.s.} 0 \quad (n\to\infty)$$

$$\max_{j\in\{1,\dots,N\}} \left| \int g_j^R \, d(\hat{P}_n - P) \right| \xrightarrow{a.s.} 0 \quad (n\to\infty)$$

よって、

$$\lim_{n\to\infty} \Pr\left(\sup_{g\in\mathcal{G}} \left| \int g \, d(\hat{P}_n - P) \right| \le \varepsilon \right) = 1$$

//

定義 6.17 $G(x) := \sup_{x \in \mathbb{R}} |g(x)|$ $(x \in \mathbb{R})$ を訂算関数とする。

補題 6.18 $\forall \varepsilon > 0$ に対して, $H_{1,B}(\varepsilon, \mathcal{G}, P) < \infty$

$$\Rightarrow G \in L_1(P) \text{ i.e., } \int G \, dP < \infty$$

証明 命題 6.7 (1) より,

$$H_{1,B}(\varepsilon, \mathcal{G}, P) < \infty \Rightarrow H_1(\varepsilon, \mathcal{G}, P) < \infty.$$

よって, $(\mathcal{G}, \|\cdot\|_{L_1(P)})$ は全有界。すなわち, $N := H_{1,B}(\varepsilon, \mathcal{G}, P)$

と書いたとき, $\exists \{g_1, \ldots, g_N\} \subset \mathcal{G}$ s.t.

$$\mathcal{G} \subset \bigcup_{j=1}^{N} B(g_j, \varepsilon).$$

$L_1(P)$は完備なので，$\mathcal{G}$の閉包 $cl(\mathcal{G})$も完備．

よって，$cl(\mathcal{G})$は全有界かつ閉なので，コンパクト．

よって，$^{\exists}R>0$ s.t.

$$\{\|g\|_{L_1(P)} : g \in \mathbb{R}\} \subset [-R, R] \Rightarrow \sup_{g \in \mathcal{G}} \|g\|_{L_1(P)} \leq R.$$

いま，$\{(g_j^L, g_j^R)\}_{j=1}^N \subset \mathcal{G}$をうまく取ると，$^{\forall}g \in \mathcal{G}$に対して，

$^{\exists}j_0 \in \{1, \dots, N\}$ s.t.

$$g_j^L(x) \leq g(x) \leq g_j^R(x) \quad (x \in \mathbb{R}) \text{ かつ } \|g_j^R - g_j^L\|_{L_1(P)} \leq \varepsilon.$$

$$\Rightarrow |g(x)| \leq |g_j^L(x)| + |g_j^R(x) - g_j^L(x)|.$$

よって

$$\int G(\omega)\,dP = \int \sup_{g\in S} |g(\omega)|\,dP$$

$$\leq \int \max_{j\in\{1,\dots,N\}} \left\{ |g_j^L(\omega)| + |g_j^R(\omega) - g_j^L(\omega)| \right\} dP$$

$$\leq \sum_{j=1}^{N} \left\{ \int |g_j^L(\omega)|\,dP + \int |g_j^R(\omega) - g_j^L(\omega)|\,dP \right\}$$

$$\leq N(R+\varepsilon).$$

//