

第8章 一様中心極限定理

8.1 距離空間における弱収束の定義と性質

(\mathbb{D}, d) を距離空間とし, $C_b(\mathbb{D})$ を \mathbb{D} 上で定義された連続有界関数全体の集まりとする.

$(\Omega, \mathcal{A}, \Pr)$ を確率空間とする. $n = 1, 2, \dots$ に対し, $X_n : \Omega \rightarrow \mathbb{D}$ は任意の写像 (Borel 可測性を仮定しなくともよい) とする. さらに, $X : \Omega \rightarrow \mathbb{D}$ を任意の写像としたとき, 任意の $h \in C_b(\mathbb{D})$ に対して

$$E^*[h(X)] := \inf \{ E[U]; U : \Omega \rightarrow \mathbb{R} \text{ はボレル可測で,} \\ U \geq h(X) \text{ で, } E[U] \text{ は存在} \}$$

とする.

定義 8.1. 任意の写像 $X_n : \Omega \rightarrow \mathbb{D}$ ($n = 1, 2, \dots$) は確率要素 $X : \Omega \rightarrow \mathbb{D}$ に分布収束するとは, 任意の $h \in C_b(\mathbb{D})$ に対して

$$E^*[h(X_n)] \xrightarrow{n \rightarrow \infty} E[h(X)]$$

が成立するときをいう. ここで $X : \Omega \rightarrow \mathbb{D}$ は Borel 可測を仮定していることに注意せよ. これを $X_n \rightsquigarrow X$ ($n \rightarrow \infty$) と記す.

定義 8.2. 任意の写像 $X_n : \Omega \rightarrow \mathbb{D}$ が $X : \Omega \rightarrow \mathbb{D}$ に確率収束するとは, 任意の $\epsilon > 0$ に対して

$$\Pr^*(d(X_n, X) > \epsilon) \xrightarrow{n \rightarrow \infty} 0$$

をみたすときをいう. これを $X_n \xrightarrow{\Pr} X$ ($n \rightarrow \infty$) と記す. ここで

$$\Pr^*(A) := \inf \{ \Pr(B); A \subset B, B \in \mathcal{A} \}$$

である.

補題 8.3. 任意の写像 $X_n : \Omega \rightarrow \mathbb{D}$ ($n = 1, 2, \dots$) と確率要素 $X : \Omega \rightarrow \mathbb{D}$ に対して, 次は同値である.

(i) 任意の $h \in C_b(\mathbb{D})$ に対して

$$E^*[h(X_n)] \xrightarrow{n \rightarrow \infty} E[h(X)]$$

となる.

(ii) 任意の有界リプシッツ函数 h に対して

$$E^*[h(X_n)] \xrightarrow{n \rightarrow \infty} E[h(X)]$$

となる.

(iii) 任意の開集合 G に対して

$$\liminf_{n \rightarrow \infty} \Pr_*(X_n \in G) \geq \Pr(X \in G)$$

となる. ただし, $\Pr_*(A) = \sup\{\Pr(B); B \subset A, B \in \mathcal{F}_n\}$ である.

(iv) 任意の閉集合 F に対して

$$\limsup_{n \rightarrow \infty} \Pr^*(X_n \in F) \leq \Pr(X \in F)$$

となる. ただし, $\Pr^*(A) = \inf\{\Pr(B); B \supset A, B \in \mathcal{F}\}$ である.

(v) 任意のボレル集合 B で, $\Pr(X \in B \setminus B^\circ) = 0$ なるものに対して

$$\Pr^*(X_n \in B) \xrightarrow{n \rightarrow \infty} \Pr(X \in B)$$

となる.

Proof. 久保木・鈴木 (2015, ppp.63–64) を参照. □

定義 8.4. Borel 可測確率要素 $X : \Omega \rightarrow \mathbb{D}$ がタイトまたは緊密であるとは, 任意の $\epsilon > 0$ に対して, \mathbb{D} のコンパクト集合 K が存在して

$$\Pr(X \notin K) < \epsilon$$

とできるときをいう.

定義 8.5. 任意の写像列 $X_n : \Omega \rightarrow \mathbb{D}$ は漸近タイトであるとは, 任意の $\epsilon > 0$ に対して, \mathbb{D} のコンパクト集合 K が存在して, すべての $\delta > 0$ に対して

$$\limsup_{n \rightarrow \infty} \Pr^*(X_n \notin K^\delta) < \epsilon$$

をみたすときをいう. ここで, $y \in \mathbb{D}$ に対して

$$\begin{aligned} d(y, K) &:= \inf\{d(y, x); x \in K\}, \\ K^\delta &:= \{y \in \mathbb{D}; d(y, K) < \delta\} \end{aligned}$$

である.

定義 8.6. 任意の写像列 $X_n : \Omega \rightarrow \mathbb{D}$ が漸近可測であるとは、任意の $h \in C_b(\mathbb{D})$ に対して、

$$E^*[h(X_n)] - E_*[h(X_n)] \xrightarrow{n \rightarrow \infty} 0$$

をみたすときをいう。ここで

$$E_*[h(X)] := \sup\{E[L]; h(X_n) \geq L, L : \Omega \rightarrow \mathbb{R} \text{ は可測}\}$$

である。

定理 8.7. $X_n : \Omega \rightarrow \mathbb{D}$ を任意の写像列とする。

(i) あるタイトな確率要素 $X : \Omega \rightarrow \mathbb{D}$ が存在して、 $X_n \rightsquigarrow X$ ($n \rightarrow \infty$) のとき、 $\{X_n\}_{n=1}^\infty$ は漸近タイトかつ漸近可測である。

(ii) X_n が漸近タイトかつ漸近可測ならば、 $\{X_n\}_{n=1}^\infty$ の部分列 $\{X_{n(k)}\}_{k=1}^\infty$ とタイトな確率要素 $X : \Omega \rightarrow \mathbb{D}$ が存在して

$$X_{n(k)} \rightsquigarrow X \quad (k \rightarrow \infty)$$

となる。

Proof. 証明は略。 □

8.2 \mathbb{R}^p 値確率変数列に対する中心極限定理

$\mathbb{X} = \mathbb{R}$ とする。

定理 8.8. X_1, X_2, \dots は独立同一分布に従う確率変数列で、 $E[X_1] = \mu$, $\text{Var}[X_1] = \sigma^2$ ($\sigma > 0$) は存在する。このとき、

$$\Pr\left\{\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq z\right\} \xrightarrow{n \rightarrow \infty} \Phi(z) \quad (\forall z \in \mathbb{R})$$

が成立する。ただし、 Φ は標準正規分布の分布関数、すなわち

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

である。

Proof. 証明は節 4.4.1 を参照のこと。 □

記号

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \rightsquigarrow \mathbf{N}(0, 1)$$

または

$$\sqrt{n}(\bar{X}_n - \mu) \rightsquigarrow N(0, \sigma^2)$$

などと記す.

次に, $\mathbb{D} = \mathbb{R}^d$ ($d \geq 2, d \in \mathbb{N}$) とする.

$\mathbf{X}_1, \mathbf{X}_2, \dots$ は独立同一分布に従う確率変数列で,

$$\begin{aligned}\boldsymbol{\mu} &= E[\mathbf{X}_1], \\ \boldsymbol{\Sigma} &= E[\mathbf{X}_1 \mathbf{X}_1^\top] - \boldsymbol{\mu} \boldsymbol{\mu}^\top\end{aligned}$$

とする. ただし, $\boldsymbol{\mu}^\top$ は縦ベクトル $\boldsymbol{\mu}$ の転置である.

定理 8.9.

$$\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \rightsquigarrow N_d(\mathbf{0}_d, \boldsymbol{\Sigma})$$

である. すなわち, 任意の縦ベクトル $\mathbf{a} \in \mathbb{R}^d$ に対して

$$\sqrt{n}[\mathbf{a}^\top (\bar{\mathbf{X}}_n - \boldsymbol{\mu})] \rightsquigarrow N(0, \mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a})$$

が成立していることである.

Proof. 証明は略. □

8.3 Donsker 型中心極限定理

標本空間を $\mathbb{X} = \mathbb{R}$ とし, 確率変数 X は \mathbb{R} 上の積率分布関数¹ F を持つ. すなわち

$$F(x) = \Pr(X \leq x) \quad (\forall x \in \mathbb{R})$$

である. X_1, X_2, \dots, X_n は X の独立複製としたとき, \mathbb{R} 上の経験分布関数 \hat{F}_n を

$$\begin{aligned}\hat{F}_n(x) &:= \frac{1}{n} \#\{X_i \leq x, i = 1, 2, \dots, n\} \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, x]}(X_i) \quad (x \in \mathbb{R})\end{aligned}$$

で定義する. ただし, $\#(\cdot)$ は (\cdot) の要素の個数を表す.

これらを用いて

$$\nu_n(x) := \sqrt{n}(\hat{F}_n(x) - F(x)) \quad (x \in \mathbb{R})$$

と定める.

¹簡単に分布関数ともいうことにする.

中心極限定理 (定理 8.8) より, 各 x に対して

$$\nu_n(x) \rightsquigarrow \mathbf{N}(0, F(x)(1 - F(x))) \quad (n \rightarrow \infty)$$

となる. また, 定理 8.9 より, $x, y \in \mathbb{R} (x < y)$ に対して

$$\begin{bmatrix} \nu_n(x) \\ \nu_n(y) \end{bmatrix} \rightsquigarrow \mathbf{N}_2(0, \Sigma(x, y)) \quad (n \rightarrow \infty)$$

となる. ただし

$$\Sigma(x, y) = \begin{bmatrix} F(x)(1 - F(x)) & F(x)(1 - F(y)) \\ F(x)(1 - F(y)) & F(y)(1 - F(y)) \end{bmatrix}$$

である.

確率過程

$$\tilde{\nu}_n = \{\nu_n(x); x \in \mathbb{R}\}$$

を定義する.

定義 8.10. \mathcal{K}_0 を $[0, 1]$ 上の有界な実数値関数の集まりとする. 確率過程 $\nu \in \mathcal{K}_0$ は標準ブラウン橋であるとは, 次の条件をみたすものである.

(i) $\nu(0) = \nu(1) = 0$.

(ii) 任意の有限個の点 $t_1, t_2, \dots, t_d \in (0, 1)$ において

$$\begin{pmatrix} \nu(t_1) \\ \nu(t_2) \\ \vdots \\ \nu(t_d) \end{pmatrix} \sim \mathbf{N}_d(\mathbf{0}, \Sigma)$$

である.

(iii) $0 \leq s \leq t \leq 1$ に対して

$$\text{Cov}[\nu(s), \nu(t)] = s(1 - t)$$

である.

(iv) ν の見本路はほとんど確実に連続である.

次に, 確率過程 $\tilde{\nu}_F$ を

$$\tilde{\nu}_F(t) := \nu(F(t)), \quad (t \in \mathbb{R})$$

で定義する.

したがって

$$\tilde{\nu}_F = \nu \circ F$$

である.

定理 8.11. $\mathcal{K} = \{h : \mathbb{R} \rightarrow \mathbb{R}; h \text{ は有界} \}$ とする. \mathcal{K} 上の確率過程

$$\begin{aligned}\tilde{v}_n(t) &= \sqrt{n}(\widehat{F}_n(t) - F(t)), \\ \tilde{v}_F(t) &= \nu(F(t)) \quad (t \in \mathbb{R})\end{aligned}$$

を考える. このとき, 以下が成立する: 任意の有界連続関数 $h : \mathcal{K} \rightarrow \mathbb{R}$ に対して,

$$E^*[h(\tilde{v}_n)] \rightarrow E[h(\tilde{v}_F)] \quad (n \rightarrow \infty)$$

が成立する.

Proof. 証明は後ほどする. □

注意 8.12. F は連続関数とする. このとき, ν はほとんど確実に連続なので, $\tilde{v}_F = \nu \circ T$ もほとんど確実に連続である. したがって, ある意味で, \tilde{v}_n は連続的に近似されなければならない. 実際, 任意の $t \in \mathbb{R}$ と t に収束する点列 $\{t_n\}_{n=1}^\infty$ に対して,

$$|\tilde{v}_n(t_n) - \tilde{v}_F(t)| \xrightarrow{P} 0, \quad (n \rightarrow \infty)$$

である. これを漸近連続という.

8.4 Donsker 族

\mathbb{X} を標本空間²とし, $X : \Omega \rightarrow \mathbb{X}$ は分布 P をもつ確率変数とし, X_1, X_2, \dots, X_n を X の独立複製とする. \mathcal{G} を関数 $g : \mathbb{X} \rightarrow \mathbb{R}$ のある族とする. 以下では, $\mathcal{G} \subset L_2(P)$ を仮定する.

関数 $g \in \mathcal{G}$ の期待値は

$$Pg := E[g(X)]$$

であり, X_1, X_2, \dots, X_n に基づく, 経験確率測度 \widehat{P}_n を

$$\widehat{P}_n(B) = \frac{1}{n} \#\{X_j \in B; j = 1, 2, \dots, n\} \quad (\forall B \in \mathcal{B}(\mathbb{X}))$$

で定める. ただし, $\#(\cdot)$ は (\cdot) の要素の個数を表す. g の (経験的) 期待値は

$$\widehat{P}_n g := \frac{1}{n} \sum_{j=1}^n g(X_j)$$

である.

² \mathbb{R} と考えてよい.

定義 8.13. 関数族 \mathcal{G} で添え字付られた経験過程を

$$\begin{aligned}\nu_n(g) &:= \sqrt{n}(\widehat{P}_n - P)g \\ &= \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n g(X_i) - E[g(X)] \right\} \quad (g \in \mathcal{G})\end{aligned}\quad (8.1)$$

で定義する.

まず, 固定した g に対する中心極限定理を思い出そう. $g(X)$ の分散を

$$\sigma^2(g) := \text{Var}[g(X)] = P g^2 - (P g)^2$$

で記す.

$\sigma^2(g) < \infty$ のとき,

$$\nu_n(g) \rightsquigarrow \mathbf{N}(0, \sigma^2(g)) \quad (n \rightarrow \infty)$$

となる. 中心極限定理は有限個の g に対しても同時に成立する. $r \in \mathbb{N}$ に対し, $g_k, g_\ell \in \mathcal{G}$ ($k, \ell = 1, 2, \dots, d$, $k \neq \ell$) は異なる関数とし, $g_k(X)$ と $g_\ell(X)$ の共分散を

$$\sigma(g_k, g_\ell) := \text{Cov}[g_k(X), g_\ell(X)] = P(g_k g_\ell) - (P g_k)(P g_\ell)$$

と記す. $k = 1, 2, \dots, d$ に対して, $\sigma^2(g_k) < \infty$ のとき

$$\begin{pmatrix} \nu_n(g_1) \\ \nu_n(g_2) \\ \vdots \\ \nu_n(g_d) \end{pmatrix} \rightsquigarrow \mathbf{N}_d(\mathbf{0}, \Sigma_{g_1, g_2, \dots, g_r}) \quad (n \rightarrow \infty)$$

となる. ただし, $\Sigma_{g_1, g_2, \dots, g_r}$ は対称行列で

$$\Sigma_{g_1, g_2, \dots, g_r} = \begin{pmatrix} \sigma^2(g_1) & \sigma(g_1, g_2) & \cdots & \sigma(g_1, g_d) \\ \sigma(g_2, g_1) & \sigma^2(g_2) & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(g_d, g_1) & \cdots & \cdots & \sigma^2(g_d) \end{pmatrix} \quad (8.2)$$

と定めた.

定義 8.14. ν を \mathcal{G} で添え字付られた Gauss 過程とする. 各 $r \in \mathbb{N}$ と各有限限の部分集合 $\{g_1, g_2, \dots, g_d\} \subset \mathcal{G}$ に対して, d 次元ベクトル

$$\begin{pmatrix} \nu(g_1) \\ \nu(g_2) \\ \vdots \\ \nu(g_d) \end{pmatrix}$$

が多変量正規分布 $\mathbf{N}_d(\mathbf{0}, \Sigma_{g_1, g_2, \dots, g_r})$ を持つとする. ただし, $\Sigma_{g_1, g_2, \dots, g_r}$ は (10.1) で与えられる分散共分散行列である. このとき, ν を \mathcal{G} で添字付けられた P-Brown 橋という.

定義 8.15. $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} P$ とし, \mathcal{G} を \mathbb{X} 上の実数値関数の族とする. ν_n を (8.1) で与えられた経験過程とし, ν を定義 8.15 で与えられた P-Brown 橋とする. さらに, ν_n と ν は \mathcal{G} 上の有界関数とする. すなわち, $\nu_n, \nu \in \ell_\infty(\mathcal{G}) = \{k : \mathcal{G} \rightarrow \mathbb{R}; k \text{ は有界}\}$ である. 関数族 \mathcal{G} は P-Donsker 族であるとは

$$\nu_n \rightsquigarrow \nu \quad (n \rightarrow \infty)$$

をみたすときをいう. すなわち, $\ell_\infty(\mathcal{G})$ 上の任意の有界連続実数値関数³ h に対して

$$E^*[h(\nu_n)] \xrightarrow{n \rightarrow \infty} E[h(\nu)]$$

をみたすことである.

定義 8.16. 定義 8.15 の記号を踏襲する. \mathcal{G} 上の ν_n が $g_0 \in \mathcal{G}$ において, 漸近同等連続であるとは, 任意の列 $\{g_n\}_{n=1}^\infty \subset \mathcal{G}$ で $\sigma(g_n - g_0) \xrightarrow{n \rightarrow \infty} 0$ なるものに対して

$$|\nu_n(g_n) - \nu_n(g_0)| \xrightarrow{P} 0 \quad (n \rightarrow \infty)$$

が成り立つときをいう. これがすべての $g_0 \in \mathcal{G}$ に対して成り立つとき, ν_n は \mathcal{G} 上で漸近同等連続であるという.

記法: $g \in L_2(P)$ に対して

$$\|g\|_{L_2(P)}^2 := Pg^2$$

と記す. すなわち, $\|\cdot\|_{L_2(P)}$ は $L_2(P)$ ノルムである.

注意 8.17. $g \in L_2(P)$ に対して, $\sigma(g) = \sqrt{Pg^2 - (Pg)^2} \leq \|g\|_{L_2(P)}$ である.

定義 8.18. 族 \mathcal{G} は全有界であるとは, $\|\cdot\|_{L_2(P)}$ によって誘導される距離に関して全有界であるときをいう. すなわち, \mathcal{G} の計量エントロピー $H_2(\cdot, \mathcal{G}, \|\cdot\|_{L_2(P)})$ は有界のときである.

³任意の $\epsilon > 0$ に対して, $\delta > 0$ が存在して, \mathcal{K} の任意の要素 \tilde{k} が

$$\sup_{x \in \mathcal{X}} |\tilde{k}(x) - k(x)| < \delta$$

ならば,

$$|f(\tilde{k}) - f(k)| < \epsilon$$

をみたすときをいう.

定理 8.19. \mathbb{X} を標本空間とする. \mathbb{X} 上の実数値関数のなす族 \mathcal{G} は全有界とする. このとき, \mathcal{G} が Donsker 族であるための必要十分条件は, \mathcal{G} 上の過程として ν_n は漸近同等連続であることである.

Proof. 補題 8.20 として証明する. □

8.5 P-Donsker 族のための基本的な補題

補題 8.20. P を \mathbb{X} 上の確率測度とし, ν_n を関数族 \mathcal{G} で添え字づけられた確率過程とし, ν_n の見本路はほとんど確実に $\ell_\infty(\mathcal{G})$ に属すとする. ただし

$$\ell_\infty(\mathcal{G}) := \left\{ k : \mathcal{G} \rightarrow \mathbb{R}; \sup_{g \in \mathcal{G}} |k(g)| < \infty \right\}$$

である. 以下のふたつの条件

- 関数族 $\mathcal{G} \subset L_2(P)$ は全有界であること,
- 任意の $\eta > 0$ に対して, ある $\delta > 0$ が存在して

$$\limsup_{n \rightarrow \infty} \Pr \left(\sup_{g_1, g_2 \in \mathcal{G}; \sqrt{P|g_1 - g_2|^2} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta \right) \leq \eta \quad (8.3)$$

であること

が成立するとする. このとき, 関数族 \mathcal{G} は P-Donsker 族である.

Proof. $\delta > 0$ を固定する. 関数族 \mathcal{G} は全有界なので, 有限部分集合 $\mathcal{G}_\delta \subset \mathcal{G}$ をうまく見つけると, $\forall g \in \mathcal{G}$ に対して, ある $\exists g_\delta \in \mathcal{G}_\delta$ があって

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

とできる. さらに, $k = \#\mathcal{G}_\delta$ と⁴おく. いま, 関数 $h : \ell_\infty(\mathcal{G}) \rightarrow \mathbb{R}$ を任意の有界 Lipschitz⁵ とする. すると補題 8.3 から

$$\nu_n \rightsquigarrow \nu (n \rightarrow \infty) \Leftrightarrow E[h(\nu_n)] \xrightarrow{n \rightarrow \infty} E[h(\nu)] \quad (8.4)$$

となるので

$$\mathcal{G} \text{ は全有界で (8.3) をみたせば, } E[h(\nu_n)] \xrightarrow{n \rightarrow \infty} E[h(\nu)]$$

となることを示せばよいことがわかる.

⁴すなわち, 有限部分集合 \mathcal{G}_δ の要素の個数である.

⁵すなわち, ある $\exists \lambda > 0$ が存在して, 任意の $\forall x, y \in \mathbb{D}$ に対して, $|h(x) - h(y)| \leq \lambda d(x, y)$ となる.

① 射影の定義. \mathcal{G} から \mathcal{G}_δ への射影 Π_δ を以下のように定める. 任意の $g \in \mathcal{G}$ に対して, $\Pi_\delta g$ を \mathcal{G}_δ のある元 g_δ で

$$\|g - g_\delta\|_{L_2(\mathbb{P})} < \delta$$

をみたすもの⁶に対応させる. すると $k = \#(\mathcal{G}_\delta)$ なので, \mathcal{G} で添え字付けられている確率過程 $\{\nu_n \circ \Pi_\delta; g \in \mathcal{G}\}$ は長さ k の確率ベクトルに他ならない. 実際, $\{g_1, g_2, \dots, g_k\} := \mathcal{G}_\delta$ と書いたとき, $\nu_n \circ \Pi_\delta$ の値域は $\{\nu_n(g_1), \nu_n(g_2), \dots, \nu_n(g_k)\}$ となる. このとき, Lipschitz 関数 $\tilde{h}: \mathbb{R}^k \rightarrow \mathbb{R}$ が存在して, 任意の有界 Lipschitz $h: \ell_\infty(\mathcal{G}) \rightarrow \mathbb{R}$ に対して

$$h(\nu_n \circ \Pi_\delta) = \tilde{h}(\nu_n \circ \Pi_\delta) \quad (8.5)$$

と書ける.

② 射影 $\Pi_\delta \nu_n$ の収束先. 古典的な CLT から

$$\nu_n \Pi_\delta \rightsquigarrow \mathbf{N}_k(\mathbf{0}_k, \Sigma) \quad (n \rightarrow \infty)$$

となる. たたし, $\Sigma = (\mathbb{P}(g_i g_j) - \mathbb{P}g_i \mathbb{P}g_j)_{i,j=1,2,\dots,k}$ である.

いま, 中心化した Gauss 過程 ν で, $\forall g_1, g_2 \in \mathcal{G}$ に対して

$$\text{Cov}(\nu(g_1), \nu(g_2)) = \mathbb{P}g_1 g_2 - \mathbb{P}g_1 \mathbb{P}g_2$$

をみたすものを考える. このとき, $\nu \circ \Pi_\delta$ は確率ベクトルで

$$\nu \circ \Pi_\delta \sim \mathbf{N}_k(\mathbf{0}_k, \Sigma)$$

となり

$$\nu_n \circ \Pi_\delta \rightsquigarrow \nu \circ \Pi_\delta \quad (n \rightarrow \infty) \quad (8.6)$$

となる. すなわち

$$\mathbb{E}[h(\nu_n \circ \Pi_\delta)] = \mathbb{E}[\tilde{h}(\nu_n \circ \Pi_\delta)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[\tilde{h}(\nu \circ \Pi_\delta)] = \mathbb{E}[h(\nu \circ \Pi_\delta)]$$

となる. ただし, h に対応した \tilde{h} は Lipschitz 関数 ((8.5) を参考のこと) である.

③ ν の修正連続過程 $\tilde{\nu}$ について. (8.3) において, ν_n を ν に置き換えた式も成立することを確認する. そのために, $\tilde{\mathcal{G}} \subset \mathcal{G}$ を有限部分集合とす

⁶ \mathcal{G}_δ は \mathcal{G} の稠密な有限部分集合なので, そのような \mathcal{G}_δ の元 g_δ は必ず存在する. また, 複数のそのような要素があったときには, 適当に 1 つを選べばよい.

る. このとき, Partmanteau の定理から

$$\begin{aligned} & \Pr\left(\sup_{g_1, g_2, \in \tilde{\mathcal{G}}; \|g_1 - g_2\|_{L_2(\mathcal{P})} \leq \delta} |\nu(g_1) - \nu(g_2)| > \eta\right) \\ & \leq \liminf \Pr\left(\sup_{g_1, g_2, \in \tilde{\mathcal{G}}; \|g_1 - g_2\|_{L_2(\mathcal{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta\right) \\ & \leq \liminf \Pr\left(\sup_{g_1, g_2, \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathcal{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta\right) \\ & \leq \limsup \Pr\left(\sup_{g_1, g_2, \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathcal{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta\right) \leq \eta \end{aligned}$$

となることから確認できた. また, 任意の有限部分集合に対して, 最後の不等号は成立するので, $\tilde{\mathcal{G}}$ を高々加算集合としても大丈夫なことがわかる. \mathcal{G} は全有界なので, \mathcal{G} の稠密な可算部分集合 $\tilde{\mathcal{G}}$ を取ることができる. したがって, \mathcal{G} 上の確率過程 $\tilde{\nu}$ を以下のように定めることができる. 任意の $g \in \mathcal{G}$ に対して, $g \in \tilde{\mathcal{G}}$ のときは

$$\tilde{\nu}(g) = \nu(g)$$

とし, $g \notin \tilde{\mathcal{G}}$ のとき, $\{g_n\}_{n=1}^{\infty} \subset \mathcal{G}$ で $g_n \xrightarrow{n \rightarrow \infty} g$ となるものに対して

$$\nu(g) = \lim_{n \rightarrow \infty} \nu(g_n)$$

と定める. この作り方から, $\tilde{\nu}$ はほとんど確実に連続な確率過程で ν の修正で

$$\Pr\left(\sup_{g_1, g_2, \in \tilde{\mathcal{G}}; \|g_1 - g_2\|_{L_2(\mathcal{P})} \leq \delta} |\tilde{\nu}(g_1) - \tilde{\nu}(g_2)| > \eta\right) \leq \eta$$

が成立している.

④ $\nu_n \rightsquigarrow \tilde{\nu} (n \rightarrow \infty)$ の証明 (ただし, $\tilde{\nu}$ は ν の修正連続過程). まず

$$\begin{aligned} & \mathbb{E}[h(\nu_n)] - \mathbb{E}[h(\tilde{\nu})] \\ & = \underbrace{\{\mathbb{E}[h(\nu_n)] - \mathbb{E}[h(\nu_n \circ \Pi_\delta)]\}}_{=:A} + \underbrace{\{\mathbb{E}[h(\nu_n \circ \Pi_\delta)] - \mathbb{E}[h(\nu \circ \Pi_\delta)]\}}_{=:B} \\ & \quad + \underbrace{\{\mathbb{E}[h(\tilde{\nu} \circ \Pi_\delta)] - \mathbb{E}[h(\tilde{\nu})]\}}_{=:C} \end{aligned} \tag{8.7}$$

と書ける. 右辺の各行を評価していく.

項 A の評価.

$$\begin{aligned}
& \left| \mathbb{E}[h(\nu_n)] - \mathbb{E}[h(\nu_n \circ \Pi_\delta)] \right| \\
& \leq \left| \mathbb{E} \left[\{h(\nu_n) - h(\nu_n \circ \Pi_\delta)\} \mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| \leq \eta \right\} \right] \right| \\
& \quad + \left| \mathbb{E} \left[\{h(\nu_n) - h(\nu_n \circ \Pi_\delta)\} \mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta \right\} \right] \right| \\
& \leq \mathbb{E} \left[\underbrace{|h(\nu_n) - h(\nu_n \circ \Pi_\delta)|}_{\leq \lambda\eta: h \text{ は Lipschitz}} \mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| \leq \eta \right\} \right] \\
& \quad + 2\|h\|_\infty \mathbb{E} \left[\mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta \right\} \right] \\
& = \lambda\eta + 2\|h\|_\infty \Pr \left(\sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta \right) \\
& \xrightarrow{n \rightarrow \infty} \lambda\eta
\end{aligned}$$

がわかる. 最後に, $\eta \rightarrow 0$ とすれば

$$\mathbb{E}[h(\nu_n)] - \mathbb{E}[h(\nu_n \circ \Pi_\delta)] \xrightarrow{n \rightarrow \infty} 0 \quad (8.8)$$

がわかる.

項 B の評価. 同様に

$$\begin{aligned}
& \left| \mathbb{E}[h(\tilde{\nu})] - \mathbb{E}[h(\tilde{\nu} \circ \Pi_\delta)] \right| \\
& \leq \left| \mathbb{E} \left[\{h(\tilde{\nu}) - h(\tilde{\nu} \circ \Pi_\delta)\} \mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| \leq \eta \right\} \right] \right| \\
& \quad + \left| \mathbb{E} \left[\{h(\nu_n) - h(\nu_n \circ \Pi_\delta)\} \mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\tilde{\nu}(g_1) - \tilde{\nu}(g_2)| > \eta \right\} \right] \right| \\
& \leq \mathbb{E} \left[\underbrace{|h(\tilde{\nu}) - h(\tilde{\nu} \circ \Pi_\delta)|}_{\leq \lambda\eta: h \text{ は Lipschitz}} \mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\tilde{\nu}(g_1) - \tilde{\nu}(g_2)| \leq \eta \right\} \right] \\
& \quad + 2\|h\|_\infty \mathbb{E} \left[\mathbb{1} \left\{ \sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\tilde{\nu}(g_1) - \tilde{\nu}(g_2)| > \eta \right\} \right] \\
& = \lambda\eta + 2\|h\|_\infty \Pr \left(\sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta} |\tilde{\nu}(g_1) - \tilde{\nu}(g_2)| > \eta \right) \\
& \xrightarrow{\eta \rightarrow 0; \delta \rightarrow 0} 0
\end{aligned}$$

から

$$\mathbb{E}[h(\tilde{\nu})] - \mathbb{E}[h(\tilde{\nu} \circ \Pi_\delta)] \xrightarrow{n \rightarrow \infty} 0 \quad (8.9)$$

がわかる.

項 C の評価. 最後に, (8.6) から

$$\mathbb{E}[h(\nu_n \circ \Pi_\delta)] - \mathbb{E}[h(\nu \circ \Pi_\delta)] \xrightarrow{n \rightarrow \infty} 0 \quad (8.10)$$

がわかる.

(8.7) – (8.10) を合わせると定理は証明できる. \square

8.6 P-Donsker の定理

定理 8.21. $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathbb{P}$ とする. ノルム空間 $(\mathcal{G}, \|\cdot\|) \subset L_2(\mathbb{P})$ は封筒関数 $G \in L_2(\mathbb{P})$ を持つとする. さらに, 閉区間 $[0, 1]$ 上で定義された非減少実数値関数 H が存在して

- $\int_0^1 H(\delta) \, d\delta < \infty,$
- $\lim_{A \rightarrow \infty} \limsup_n \Pr\left(\sup_{\delta > 0} \frac{H(\delta, \mathcal{G}, \|\cdot\|_{L_2(\hat{\mathbb{P}}_n)})}{H(\delta)} > A\right) = 0$ をみたすとす
る.

このとき, 任意の $\eta > 0$ に対して, ある $\exists \delta > 0$ が存在して

$$\limsup_{n \rightarrow \infty} \Pr\left(\sup_{g_1, g_2 \in \mathcal{G}; \sqrt{\mathbb{P}|g_1 - g_2|^2} \leq \delta} |\nu_n(g_1) - \nu_n(g_2)| > \eta\right) \leq \eta$$

となる. ただし, $\nu_n = \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P})$ である. したがって, \mathcal{G} は P-Donsker 族である.

Proof. まず, $\delta > 0, \eta > 0, M > 0$ に対して, 以下の事象を定める.

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

$$E_{n, M} := \left\{ \sup_{\delta > 0} \frac{H(\delta, \mathcal{G}, \|\cdot\|_{L_2(\hat{\mathbb{P}}_n)})}{H(\delta)} > M \right\}$$

とおく. このとき

$$\begin{aligned}
\Pr(A_\delta) &= \mathbb{E}[\mathbf{1}_{A_\delta} \mathbf{1}_{E_{n,M}}] + \mathbb{E}[\mathbf{1}_{A_\delta} \mathbf{1}_{E_{n,M}^c}] \\
&\leq \Pr(E_{n,M}) + \frac{\mathbb{E}\left[\left\{\sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})}} |\nu_n(g_1) - \nu_2(g_2)|\right\} \mathbf{1}_{E_{n,M}^c}\right]}{\eta} \\
&\quad \left(\because A_\delta \Rightarrow \frac{\sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})}} |\nu_n(g_1) - \nu_2(g_2)|}{\eta} > 1\right) \\
&\leq \Pr(E_{n,M}) + \frac{\mathbb{E}\left[\left\{\sup_{g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})}} |\nu_n(g_1 - g_2)|\right\} \mathbf{1}_{E_{n,M}^c}\right]}{\eta} \\
&\leq \Pr(E_{n,M}) + 24 \frac{\mathbb{E}\left[\int_0^{\delta_n/2} \sqrt{H(\epsilon, \mathcal{G}^\delta, \|\cdot\|_{L_2(\hat{\mathbb{P}}_n)})} d\epsilon \mathbf{1}_{E_{n,M}^c}\right]}{\eta} \\
&\quad (\because (6.34))
\end{aligned}$$

となる. ただし, $\mathcal{G}^\delta := \{g_1 - g_2; g_1, g_2 \in \mathcal{G}; \|g_1 - g_2\|_{L_2(\mathbb{P})} \leq \delta\}$ と $\delta_n := \sup_{g \in \mathcal{G}^\delta} \|g\|_{L_2(\hat{\mathbb{P}}_n)}$ である. まず, \mathcal{G} の ϵ 被覆は \mathcal{G}^δ の 2ϵ 被覆を与えることに注意する. 実際, \mathcal{G} の ϵ 被覆の中心を g_1, g_2, \dots, g_k と書く. このとき, $g_{i,j}^{\text{diff}} := g_i - g_j$ とおくと, これを中心とした, 2ϵ 球全体は \mathcal{G}^δ の被覆となることがわかる. よって, $\#\{g_{i,j}^{\text{diff}}; i, j = 1, 2, \dots, k\} = k^2$ となるので

$$H(2\epsilon, \mathcal{G}^\delta, \|\cdot\|_{L_2(\hat{\mathbb{P}}_n)}) \leq 2H(\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\hat{\mathbb{P}}_n)}) \quad (8.11)$$

がわかる.

定理 6.20 の経験過程 $\{\|g_1 - g_1\|_{L_2(\hat{\mathbb{P}}_n)}^2; g_1, g_2 \in \mathcal{G}\}$ への適用. 任意の $x \in \mathbb{X}$ に対して

$$\begin{aligned}
(g_1 - g_2)^2(x) &= |g_1(x) - g_2(x)| \times |g_1(x) - g_2(x)| \\
&\leq \{|g_1(x)| + |g_2(x)|\} |g_1(x) - g_2(x)| \\
&\leq 2G(x) |g_1(x) - g_2(x)| \\
&\leq \sqrt{4G^2(x)} \sqrt{\{g_1(x) - g_2(x)\}^2}
\end{aligned}$$

となるので

$$\|g_1 - g_2\|_{L_2(\hat{\mathbb{P}}_n)}^2 \leq 2\|G\|_{L_2(\hat{\mathbb{P}}_n)} \|g_1 - g_2\|_{L_2(\hat{\mathbb{P}}_n)}$$

がわかる. 古典的な大数の強法則から

$$\|G\|_{L_2(\hat{\mathbb{P}}_n)} \xrightarrow{\text{a.s.}} \|G\|_{L_2(\mathbb{P})} \quad (n \rightarrow \infty)$$

がわかる. このことから, 十分大きな n に対して

$$\|g_1 - g_2\|_{L_2(\hat{\mathcal{P}}_n)}^2 \leq 4\|G\|_{L_2(\mathcal{P})}\|g_1 - g_2\|_{L_2(\hat{\mathcal{P}}_n)} \quad (8.12)$$

となる. いま

$$\mathcal{H} := \{(g_1 - g_2)^2 : g_1, g_2 \in \mathcal{G}\}$$

とおく. すると (8.11) と (8.12) から, $\forall \epsilon > 0$ に対して

$$\begin{aligned} H(\epsilon, \mathcal{H}, \|\cdot\|_{L_2(\hat{\mathcal{P}}_n)}) &\leq H\left(\frac{\epsilon}{4\|G\|_{L_2(\mathcal{P})}}, \mathcal{G}^\delta, \|\cdot\|_{L_2(\hat{\mathcal{P}}_n)}\right) \\ &\leq H\left(\frac{\epsilon}{8\|G\|_{L_2(\mathcal{P})}}, \mathcal{G}, \|\cdot\|_{L_2(\hat{\mathcal{P}}_n)}\right) \end{aligned}$$

となることがわかる. 上の不等式と定理の条件から

$$\begin{aligned} &\frac{1}{n}H(\epsilon, \mathcal{H}, \|\cdot\|_{L_2(\hat{\mathcal{P}}_n)}) \\ &\leq \frac{1}{n}H\left(\frac{\epsilon}{8\|G\|_{L_2(\mathcal{P})}}, \mathcal{G}, \|\cdot\|_{L_2(\hat{\mathcal{P}}_n)}\right) \xrightarrow{\mathbb{P}} 0 \quad (n \rightarrow \infty) \quad (8.13) \end{aligned}$$

がわかる. さらに, $G \in L_2(\mathcal{P})$ なので, 関数 $\mathbb{X} \ni x \mapsto \sup_{g_1, g_2 \in \mathcal{G}} (g_1 - g_2)^2(x)$ の封筒関数⁷ $4G^2(x)$ は $L_1(\mathcal{P})$ となることがわかる. よって, (8.13) と \mathcal{H} の封筒関数が $L_1(\mathcal{P})$ であることが確認できた. truncate argument を用いた定理 6.20 の証明の II. の部分の議論と同じことをすると

$$\sup_{g_1, g_2 \in \mathcal{G}} \left| \|g_1 - g_2\|_{L_2(\hat{\mathcal{P}}_n)} - \|g_1 - g_2\|_{L_2(\mathcal{P})} \right| \xrightarrow{\text{a.s.}} 0 \quad (n \rightarrow \infty)$$

がわかる. したがって, 十分大きな n と任意の $g_1, g_2 \in \mathcal{G}$ に対して

$$\frac{1}{2}\|g_1 - g_2\|_{L_2(\hat{\mathcal{P}}_n)} \leq \|g_1 - g_2\|_{L_2(\mathcal{P})} \leq 2\|g_1 - g_2\|_{L_2(\hat{\mathcal{P}}_n)} \quad (8.14)$$

となる. したがって, 十分大きな n に対して

$$H(2\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\mathcal{P})}) \leq H(\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\hat{\mathcal{P}}_n)})$$

となるので, \mathcal{G} は全有界であることがわかる. 同様に, 十分大きな n に対して

$$\delta_n \leq 2\delta$$

⁷すなわち, \mathcal{H} の封筒関数である.

となること⁸もわかる. (8.11) から

$$\begin{aligned} \int_0^{\delta_n/2} \sqrt{H(\epsilon, \mathcal{G}^\delta, \|\cdot\|_{L_2(\hat{P}_n)})} d\epsilon &\leq \int_0^{\delta_n/2} \sqrt{2H(\epsilon/2, \mathcal{G}, \|\cdot\|_{L_2(\hat{P}_n)})} d\epsilon \\ &\leq \sqrt{2} \int_0^\delta \sqrt{H(\epsilon/2, \mathcal{G}, \|\cdot\|_{L_2(\hat{P}_n)})} d\epsilon \\ &\leq 2\sqrt{2} \int_0^{\delta/2} \sqrt{H(\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\hat{P}_n)})} d\epsilon \end{aligned}$$

となるので

$$\begin{aligned} \Pr(A_\delta) &\leq \Pr(E_{n,M}) + 48\sqrt{2} \frac{\mathbb{E} \left[\int_0^{\delta/2} \sqrt{H(\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\hat{P}_n)})} d\epsilon \mathbf{1}_{E_{n,M}^c} \right]}{\eta} \\ &\leq \Pr(E_{n,M}) + 48\sqrt{2} \frac{\mathbb{E} \left[\int_0^{\delta/2} \sqrt{H(\epsilon)} d\epsilon \right]}{\eta} \end{aligned}$$

を得る. $n \rightarrow \infty, M \rightarrow \infty, \delta \rightarrow 0$ とすると定理の主張が証明できる. \square

定理 8.22. $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} P$ とし, $\mathcal{G} \subset L_2(P)$ を関数族とし, G を \mathcal{G} の封筒関数とする. 以下を仮定する.

- $G \in L_2(P)$,
- $\int_0^1 H(\delta) d\delta < \infty$,
- $\lim_{A \rightarrow \infty} \limsup_n \Pr \left(\sup_{\delta > 0} \frac{H(\delta, \mathcal{G}, \|\cdot\|_{L_2(\hat{P}_n)})}{H(\delta)} > A \right) = 0$

をみたすとする. このとき, \mathcal{G} は P -Donsker 族である.

Proof. 定理 8.21 と補題 8.20 から直ちにわかる. \square

8.7 一様エントロピー条件

定理 8.23. $(\Omega, \mathcal{A}, \Pr)$ を基礎の確率空間とし, P を \mathbb{X} 上の誘導された確率分布とする. \mathcal{G} を \mathbb{X} 上の可測関数の族で

$$\int_0^\infty \sup_Q \sqrt{\log N(\epsilon \|G\|_{L_2(Q)}, \mathcal{G}, \|\cdot\|_{L_2(Q)})} d\epsilon < \infty \quad (8.15)$$

⁸ $\delta_n = \sup_{g \in \mathcal{G}^{\text{delta}}} \|g\|_{L_2(\hat{P}_n)}$ だったので, (8.14) から直ちにわかる.

を満たすとする. ただし, G は関数族 \mathcal{G} の封筒関数, \sup は \mathbb{X} 上のすべての有限離散確率測度で $\|G\|_{L_2(Q)} := \int G^2 dQ > 0$ なるものに関して取ったものである. さらに, すべての $\delta > 0$ に対して, 関数族 $\mathcal{G}^\delta = \{g_1 - g_2; g_1, g_2 \in \mathcal{G}, \|g_1 - g_2\|_{L_2(P)} < \delta\}$ と $\mathcal{G}_\infty^2 = \{g^2; g \in \mathcal{G}, \sup_{x \in \mathbb{X}} |g(x)| < \infty\}$ は \Pr 可測とする. このとき, $PG^2 < \infty$ ならば, \mathcal{G} は P -Donsker 族である.

Proof. 補題 8.20 から, 関数族 \mathcal{G} は P -Donsker であることを証明するために

I ν_n が漸近同等連続であること,

II \mathcal{G} が $L_2(P)$ において全有界でこと

を確認すればよいことに注意する.

I の証明. そのために, $\{\delta_n\}_{n=1}^\infty$ を $\delta_n \downarrow 0 (n \rightarrow \infty)$ なる任意の数列としたとき

$$\Pr\left(\sup_{g \in \mathcal{G}^{\delta_n}} \|\nu_n(g)\| > x\right) \xrightarrow{n \rightarrow \infty} 0$$

を確認すればよいことに注意する.

① $\Pr\left(\sup_{g \in \mathcal{G}^{\delta_n}} \|\nu_n(g)\| > x\right)$ の上限の導出. 任意の $x > 0$ に対して, Markov の不等式から

$$\Pr\left(\sup_{g \in \mathcal{G}^\delta} \|\nu_n(g)\| > x\right) \leq \frac{\mathbb{E}\left[\sup_{g \in \mathcal{G}^\delta} |\nu_n(g)|\right]}{x} \quad (8.16)$$

となる. 次に, 対称化トリックを用いて, 右辺の分子の期待値を評価する. そのために, Y_1, Y_2, \dots, Y_n を X_1, X_2, \dots, X_n の独立複製とし, $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ を $\{X_j\}_{j=1}^n$ と $\{Y_j\}_{j=1}^n$ とは独立な Rademacher 列とする.

すると

$$\begin{aligned}
\mathbb{E} \left[\sup_{g \in \mathcal{G}^\delta} |\nu_n(g)| \right] &= \mathbb{E} \left[\sup_{g \in \mathcal{G}^\delta} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \{g(X_j) - \mathbb{P}g\} \right| \right] \\
&= \mathbb{E} \left[\sup_{g \in \mathcal{G}^\delta} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \{g(X_j) - \mathbb{E}^Y [g(Y_j)]\} \right| \right] \\
&\leq \mathbb{E} \left[\sup_{g \in \mathcal{G}^\delta} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \{g(X_j) - g(Y_j)\} \right| \right] \\
&\quad (\because \text{Jensen の不等式}) \\
&= \mathbb{E} \left[\sup_{g \in \mathcal{G}^\delta} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j \{g(X_j) - g(Y_j)\} \right| \right] \\
&\leq 2\mathbb{E} \left[\sup_{g \in \mathcal{G}^\delta} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g(X_j) \right| \right] \tag{8.17}
\end{aligned}$$

を得る. $\mathbf{X} = (X_1, X_2, \dots, X_n)$ とおき, 補題 5.13(Hoeffding の補題) から, 任意の $\lambda > 0$ に対して

$$\begin{aligned}
\mathbb{E} \left[\exp \left(\frac{\lambda}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g(X_j) \right) \middle| \mathbf{X} \right] &= \mathbb{E} \left[\prod_{j=1}^n \exp \left(\frac{\lambda}{\sqrt{n}} \epsilon_j g(X_j) \right) \middle| \mathbf{X} \right] \\
&= \prod_{j=1}^n \mathbb{E} \left[\exp \left(\frac{\lambda}{\sqrt{n}} \epsilon_j g(X_j) \right) \middle| \mathbf{X} \right] \\
&\leq \prod_{j=1}^n \exp \left(\frac{\lambda^2}{8} \left\{ \frac{g(X_j)}{\sqrt{n}} - \left(-\frac{g(X_j)}{\sqrt{n}} \right) \right\}^2 \right) \\
&\leq \prod_{j=1}^n \exp \left(\frac{\lambda^2}{2n} g^2(X_j) \right) \\
&= \exp \left(\frac{\lambda^2}{2n} \sum_{j=1}^n g^2(X_j) \right) \\
&= \exp \left(\frac{\lambda^2 \|g\|_{L_2(\hat{\mathbb{P}}_n)}^2}{2} \right)
\end{aligned}$$

と評価できる. したがって, X_1, X_2, \dots, X_n を与えたとき

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g(X_j) \sim \text{subG}(\|g\|_{L_2(\hat{\mathbb{P}}_n)})$$

となる. 定理 6.29 から

$$\begin{aligned}
& \mathbb{E} \left[\sup_{g \in \mathcal{G}^{\delta_n}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g(X_j) \right| \middle| \mathbf{X} \right] \\
& \leq \mathbb{E} \left[\sup_{g_1, g_2 \in \mathcal{G}} \left| \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g_1(X_j) - \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g_2(X_j) \right\} \right| \middle| \mathbf{X} \right] \\
& \leq 24 \int_0^\infty \sqrt{\log N(\epsilon, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\hat{\mathbb{P}}_n)})} d\epsilon \\
& \leq 24 \int_0^\infty \sup_{\mathbb{Q}} \sqrt{\log N(\epsilon, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathbb{Q})})} d\epsilon \tag{8.18}
\end{aligned}$$

を得る. ただし, \sup は任意の有限離散確率分布 \mathbb{Q} に関して取ったものである.

② 積分 $\int_0^\infty \sup_{\mathbb{Q}} \sqrt{\log N(\epsilon, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathbb{Q})})} d\epsilon$ の評価. 十分大きな ϵ に対して, 関数族 \mathcal{G}^δ は一つの球に含まれるので, 上の積分の被積分関数は 0 となる. これは ϵ が

$$\theta_n := \sup_{g \in \mathcal{G}^{\delta_n}} \|g\|_{L_2(\hat{\mathbb{P}}_n)}^2 = \sup_{g \in \mathcal{G}^{\delta_n}} \left| \frac{1}{n} \sum_{j=1}^n g^2(X_j) \right|$$

より大きい場合も同じである. 関数族 \mathcal{G}^δ の被覆数は関数族 $\mathcal{G}^{\text{diff}} := \{g_1 - g_2; g_1, g_2 \in \mathcal{G}\}$ の被覆数で上から評価できる. さらに

$$N(\epsilon, \mathcal{G}^{\text{diff}}, \|\cdot\|_{L_2(\mathbb{Q})}) \leq N^2(\epsilon/2, \mathcal{G}, \|\cdot\|_{L_2(\mathbb{Q})})$$

と評価できる. ただし, \mathbb{Q} は \mathbb{X} 上の任意の確率測度である. 積分を $(0, \theta_n)$ に制限し, 変数変換をすると

$$\begin{aligned}
& \int_0^{\theta_n} \sup_{\mathbb{Q}} \sqrt{\log N(\epsilon, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathbb{Q})})} d\epsilon \\
& = \int_0^{\theta_n / \|G\|_{L_2(\hat{\mathbb{P}}_n)}} \sup_{\mathbb{Q}} \sqrt{\log N(\epsilon \|G\|_{L_2(\hat{\mathbb{P}}_n)}, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathbb{Q})})} d\epsilon \times \|G\|_{L_2(\hat{\mathbb{P}}_n)} \\
& \leq \int_0^{\theta_n / \|G\|_{L_2(\hat{\mathbb{P}}_n)}} \sup_{\mathbb{Q}} \sqrt{\log N(\epsilon \|G\|_{L_2(\mathbb{Q})}, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathbb{Q})})} d\epsilon \times \|G\|_{L_2(\hat{\mathbb{P}}_n)}
\end{aligned}$$

と上から評価できる. 定理の仮定から, 上の積分は可積である. 大数の強法則から

$$\|G\|_{L_2(\hat{\mathbb{P}}_n)} \xrightarrow{\text{a.s.}} \|G\|_{L_2(\mathbb{P})} \quad (n \rightarrow \infty)$$

がわかる. Cauchy-Schwarz の不等式から

$$\begin{aligned} & \left| \mathbb{E} \left[\int_0^{\theta_n / \|G\|_{L_2(\hat{P}_n)}} \sup_{\mathcal{Q}} \sqrt{\log N(\epsilon \|G\|_{L_2(\mathcal{Q})}, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathcal{Q})})} d\epsilon \times \|G\|_{L_2(\hat{P}_n)} \right] \right| \\ & \leq \sqrt{\mathbb{E} \left[\left\{ \int_0^{\theta_n / \|G\|_{L_2(\hat{P}_n)}} \sup_{\mathcal{Q}} \sqrt{\log N(\epsilon \|G\|_{L_2(\mathcal{Q})}, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathcal{Q})})} d\epsilon \right\}^2 \right]} \\ & \quad \times \sqrt{\mathbb{E} [\|G\|_{L_2(\hat{P}_n)}^2]} \end{aligned}$$

となるので, 優収束定理から, $\theta_n \xrightarrow{P} 0 (n \rightarrow \infty)$ がわかれば

$$\mathbb{E} \left[\int_0^{\theta_n / \|G\|_{L_2(\hat{P}_n)}} \sup_{\mathcal{Q}} \sqrt{\log N(\epsilon \|G\|_{L_2(\mathcal{Q})}, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathcal{Q})})} d\epsilon \times \|G\|_{L_2(\hat{P}_n)} \right] \xrightarrow{n \rightarrow \infty} 0 \quad (8.19)$$

がわかる. (8.16) – (8.19) を合わせると, $\theta_n \xrightarrow{P} 0 (n \rightarrow \infty)$ がわかれば

$$\begin{aligned} \Pr \left(\sup_{g \in \mathcal{G}^{\delta_n}} \|\nu_n(g)\| > x \right) & \leq \frac{\mathbb{E} \left[\sup_{g \in \mathcal{G}^{\delta_n}} |\nu_n(g)| \right]}{x} \\ & \leq \frac{2 \mathbb{E} \left[\sup_{g \in \mathcal{G}^{\delta_n}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j g(X_j) \right| \right]}{x} \\ & \leq \frac{24 \mathbb{E} \left[\int_0^\infty \sup_{\mathcal{Q}} \sqrt{\log N(\epsilon, \mathcal{G}^{\delta_n}, \|\cdot\|_{L_2(\mathcal{Q})})} d\epsilon \right]}{x} \\ & \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

がわかる. したがって, ν_n の漸近同等連続性を証明するためには, あとは $\theta_n \xrightarrow{P} 0 (n \rightarrow \infty)$ を示せばよい.

③ $\theta_n \xrightarrow{P} 0 (n \rightarrow \infty)$ の証明.

$$\sup \{ P g^2; g \in \mathcal{G}^{\delta_n} \} \xrightarrow{n \rightarrow \infty} 0$$

であり, $\mathcal{G}^{\delta_n} \subset \mathcal{G}^{\text{diff}}$ であることに注意すると

$$\sup_{g \in \mathcal{G}^{\text{diff}}} |\hat{P}_n g^2 - P g^2| \xrightarrow{\text{a.s.}} 0 \quad (n \rightarrow \infty)$$

を示せば十分であることがわかる. すなわち, 関数族 $\{\mathcal{G}^{\text{diff}}\}^2$ は P-GC 族であることがわかればよい. まず, 関数族 $\{\mathcal{G}^{\text{diff}}\}^2$ の封筒関数は $(2G)^2$ である. これは, 定理の仮定から $4G^2 \in L_1(P)$ であることがわかる. さら

に, $g_1, g_2 \in \mathcal{G}^{\text{diff}}$ に対して

$$\begin{aligned} \widehat{P}_n |g_1^2 - g_2^2|^2 &= \widehat{P}_n \left(|g_1 - g_2| \times |g_1 + g_2| \right) \\ &\leq 4\widehat{P}_n \left(|g_1 - g_2| G \right) \\ &\leq 4\sqrt{\|g_1 - g_2\|_{L_2(\widehat{P}_n)}} \sqrt{\|G\|_{L_2(\widehat{P}_n)}} \end{aligned}$$

となるので

$$N\left(4\epsilon \|G\|_{L_2(\widehat{P}_n)}^2, \{\mathcal{G}^{\text{diff}}\}^2, L_1(\widehat{P}_n)\right) \leq N\left(\epsilon \|G\|_{L_2(\widehat{P}_n)}, \mathcal{G}^{\text{diff}}, L_2(\widehat{P}_n)\right)$$

となる. 定理の仮定から, 上の不等式の左辺は有界となる. よって

$$\frac{1}{n} \log N\left(4\epsilon \|G\|_{L_2(\widehat{P}_n)}^2, \{\mathcal{G}^{\text{diff}}\}^2, L_1(\widehat{P}_n)\right) \xrightarrow{P} 0$$

がわかる. 定理 6.20 から関数族 $\{\mathcal{G}^{\text{diff}}\}^2$ は P-GC 族であることがわかる.

II の証明. まず, I の証明の最後の部分から

$$\sup_{g \in \mathcal{G}^{\text{diff}}} \left| (\widehat{P}_n - P)g^2 \right| \xrightarrow{\text{a.s.}} 0 \quad (n \rightarrow \infty)$$

であった. 任意の $\epsilon > 0$ に対して, n を十分大きくとると

$$\sup_{g \in \mathcal{G}^{\text{diff}}} \left| (\widehat{P}_n - P)g^2 \right| < \epsilon^2$$

となる. 定理の仮定から $N(\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\widehat{P}_n)})$ は有限である. よって, \mathcal{G} の $\|\cdot\|_{L_2(\widehat{P}_n)}$ に関する ϵ 網は $\|\cdot\|_{L_2(P)}$ における $\sqrt{2}\epsilon$ 網となることがわかる. よって, \mathcal{G} は $L_2(P)$ の全有界集合であることがわかる. \square

例 8.24. (1). $\mathcal{G} := \{\mathbb{1}_{(-\infty, r]}; r \in \mathbb{R}\}$ とする. 明らかに \mathcal{G} は全有界である. このとき, \mathbb{R} 上の任意の有限離散確率測度 Q と $\forall 1 \geq \epsilon > 0$ に対して

$$N(\epsilon, \mathcal{G}, L_2(Q)) \leq N_B(\epsilon^2, \mathcal{G}, L_1(Q)) \leq \frac{2}{\epsilon^2} \quad (8.20)$$

となる. $\epsilon > 1$ のとき, $\log N(\epsilon \|G\|_{L_2(Q)}, \mathcal{G}, \|\cdot\|_{L_2(Q)}) = 1$ とあり, $G(x) = \sup_{r \in \mathbb{R}} |\mathbb{1}_{(-\infty, r]}(x)| = 1$ であることに注意する. さらに

$$\int_0^1 \log\left(\frac{1}{\epsilon}\right) d\epsilon < \infty$$

なので

$$\begin{aligned}
 & \int \sup_{\mathcal{Q}} \sqrt{\log N(\epsilon \|G\|_{L_2(\mathcal{Q})}, \mathcal{G}, \|\cdot\|_{L_2(\mathcal{Q})})} d\epsilon \\
 & \leq 2 \int_0^1 \sqrt{\log\left(\frac{\sqrt{2}}{\epsilon}\right)} d\epsilon \\
 & \leq 2 \sqrt{\int_0^1 \log\left(\frac{\sqrt{2}}{\epsilon}\right) d\epsilon} \quad (\because \text{Cauchy-Schwarz の不等式}) \\
 & < \infty
 \end{aligned}$$

なので, 関数族 $\{\mathbf{1}_{(-\infty, r_j]}; r_j \in \mathbb{R}\}$ は P-Donsker 族となる.

よって, F を \mathbb{R} 上の分布関数とし, $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} F$ とし

$$\hat{F}_n(x) = \frac{1}{n} \#\{X_j \leq x; j = 1, 2, \dots, n\}$$

としたとき

$$\sqrt{n}\{\hat{F}_n - F\} \rightsquigarrow \tilde{\nu}_F \quad (n \rightarrow \infty)$$

となることがわかる. ただし, $\tilde{\nu}_F$ は定義 8.10 で与えられる Gauss 過程である.

(2). $\mathbf{r} = (r_1, r_2, \dots, r_d)^\top \in \mathbb{R}^d$ に対して, $[-\infty, \mathbf{r}] := (-\infty, r_1] \times (-\infty, r_2] \times \dots \times (-\infty, r_d]$ と書くことにする. 関数族 $\{\mathbf{1}_{(-\infty, \mathbf{r}]}; \mathbf{r} \in \mathbb{R}^d\}$ の被覆数は $\left(\frac{1}{\epsilon}\right)^d$ と上から評価できる. よって, 関数族 $\{\mathbf{1}_{(-\infty, \mathbf{r}]}; \mathbf{r} \in \mathbb{R}^d\}$ も P-Donsker となることがわかる. \square