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

$(\mathbb{M}, d)$ を距離空間とし, $C_b(\mathbb{M})$ を $\mathbb{M}$ 上で定義さらた連続有界函数全体の集まりとする.

$(\Omega, \mathcal{A}, \mathsf{Pr})$ を確率空間とする. $n = 1, 2, \ldots$ に対し, $X_n : \Omega \to \mathbb{M}$ 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, $X : \Omega \to \mathbb{M}$ を任意の写像としたとき, 任意の $h \in C_b(\mathbb{M})$ に対して,

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

とする.

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

$$\mathsf{E}^*[h(X_n)] \overset{n\to\infty}{\longrightarrow} \mathsf{E}[h(X)]$$

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

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

$$\mathsf{Pr}^*(d(X_n, X) > \epsilon) \overset{n\to\infty}{\longrightarrow} 0$$

をみたすときをいう. これを $X_n \overset{\mathrm{P}}{\to} X\,(n \to \infty)$ と記す. ここで

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

である.

補題 8.3. 任意の写像 $X_n : \Omega \longrightarrow \mathbb{M}\,(n = 1, 2, \ldots)$ と確率要素 $X : \Omega \longrightarrow \mathbb{M}$ に対して, 次は同値である.
(i) 任意の $h \in C_b(\mathbb{M})$ に対して

$$\mathsf{E}^*[h(X_n)] \overset{n\to\infty}{\longrightarrow} \mathsf{E}[h(X)]$$

となる.
(ii) 任意の有界リプシッツ函数 h に対して

$$\mathsf{E}^*[h(X_n)] \stackrel{n\to\infty}{\longrightarrow} \mathsf{E}[h(X)]$$

となる.
(iii) 任意の開集合 G に対して

$$\liminf_{n\to\infty} \mathsf{Pr}_*(X_n \in G) \geq \mathsf{Pr}(X \in G)$$

となる. ただし, $\mathsf{Pr}_*(A) = \sup\{\mathsf{Pr}(B);\ B \subset A,\ B \in \mathcal{F}_n\}$ である.
(iv) 任意の閉集合 F に対して

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

となる. ただし, $\mathsf{Pr}^*(A) = \inf\{\mathsf{Pr}(B);\ B \supset A,\ B \in \mathcal{F}_n\}$ である.
(v) 任意のボレル集合 B で, $\mathsf{Pr}(X \in B \setminus B^\circ) = 0$ なるものに対して

$$\mathsf{Pr}^*(X_n \in B) \stackrel{n\to\infty}{\longrightarrow} \mathsf{Pr}(X \in B)$$

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 $X : \Omega \longrightarrow \mathbb{M}$ がタイトまたは緊密であるとは, 任意の $\epsilon > 0$ に対して, $\mathbb{M}$ のコンパクト集合 K が存在して

$$\mathsf{Pr}(X \notin K) < \epsilon$$

とできるときをいう.

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

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

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

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

である.

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

$$\mathsf{E}^*[h(X_n)] - \mathsf{E}_*[h(X_n)] \stackrel{n\to\infty}{\longrightarrow} 0$$

をみたすときをいう. ここで

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

である.

定理 8.7. $X_n : \Omega \longrightarrow \mathbb{M}$ を任意の写像列とする.
(i) あるタイトな確率要素 $X : \Omega \longrightarrow \mathbb{M}$ が存在して, $X_n \rightsquigarrow X\ (n \to \infty)$ 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 $X : \Omega \longrightarrow \mathbb{M}$ が存在して

$$X_{n(k)} \rightsquigarrow X \qquad (k \to \infty)$$

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

$$\mathsf{Pr}\left\{\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \leq z\right\} \stackrel{n\to\infty}{\longrightarrow} \Phi(z) \qquad (\forall z \in \mathbb{R})$$

が成立する. ただし, Φ は標準正規分布の分布函数, 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. □

<u>記号</u>

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

または

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

などと記す.

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

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

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

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.

$$\sqrt{n}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}) \rightsquigarrow \mathsf{N}_d(\boldsymbol{0}_d,\, \boldsymbol{\Sigma})$$

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

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

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

$\mathbb{M} = \mathbb{R}$ とし, 確率変数 X は分布関数 F を持つ, すなわち,

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

$X_1,\, X_2,\, \ldots,\, X_n$ は X の独立複製としたとき, 経験分布関数を

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

で定義する.

これらを用いて

$$W_n(x) := \sqrt{n}(\widehat{\mathsf{F}}_n(x) - \mathsf{F}(x)) \qquad (x \in \mathbb{R})$$

と定める.

中心極限定理（定理 8.8）より, 各 x に対して

$$W_n(x) \xrightarrow{d} \mathsf{N}(0,\, \mathsf{F}(x)(1 - \mathsf{F}(x)))$$

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

$$\begin{bmatrix} W_n(x) \\ W_n(y) \end{bmatrix} \rightsquigarrow \mathsf{N}_2(0,\, \boldsymbol{\Sigma}(x,\, y))$$
$$\boldsymbol{\Sigma}(x,\, y) = \begin{bmatrix} \mathsf{F}(x)(1-\mathsf{F}(x)) & \mathsf{F}(x)(1-\mathsf{F}(y)) \\ \mathsf{F}(x)(1-\mathsf{F}(y)) & \mathsf{F}(y)(1-\mathsf{F}(y)) \end{bmatrix}$$

である.

確率過程

$$\widetilde{\nu}_n = \{W_n(x);\, x \in \mathbb{R}\}$$

を定義する.

定義 8.10. $\mathcal{K}_0$ を $[0,\, 1]$ 上の有界な実数値函数の集まりとする. 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 $t_1,\, t_2,\, \ldots,\, t_d \in (0,\, 1)$ において

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

である.
(iii) $0 \leq s \leq t \leq 1$ に対して

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

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 $\widetilde{\nu}_{\mathsf{F}}$ を

$$\widetilde{\nu}_{\mathsf{F}}(t) := \nu(\mathsf{F}(t)), \qquad (t \in \mathbb{R})$$

で定義する.

したがって

$$\widetilde{\nu}_{\mathsf{F}} = \nu \circ \mathsf{F}$$

である.

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

$$\begin{aligned} \widetilde{\nu}_n(t) &= \sqrt{n}(\widehat{\mathsf{F}}_n(t) - \mathsf{F}(t)), \\ \widetilde{\nu}_{\mathsf{F}}(t) &= \nu(\mathsf{F}(t)) \qquad (t \in \mathbb{R}) \end{aligned}$$

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

$$\mathsf{E}^*[h(\widetilde{\nu}_n)] \longrightarrow \mathsf{E}[h(\widetilde{\nu}_{\mathsf{F}})] \qquad (n \to \infty)$$

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, $\widetilde{\nu}_{\mathsf{F}} = \nu \circ T$ もほとんど確実に連続である. したがって, ある意味で, $\widetilde{\nu}_n$ は連続的に近似されなけらばならない. 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,

$$|\widetilde{\nu}_n(t_n) - \widetilde{\nu}_{\mathsf{F}}(t)| \xrightarrow{\mathrm{P}} 0, \qquad (n \to \infty)$$

である. これを漸近連続という.

8.4 Donsker 族

$\mathbb{M}$ を距離空間または準距離空間とし, $X : \Omega \longrightarrow \mathbb{M}$ は分布 P をもつ確率要素とし, $X_1, X_2, \ldots, X_n$ を X の独立複製とする. 関数 $g : \mathbb{M} \longrightarrow \mathbb{R}$ の族 $\mathcal{G}$ を考える.

関数 g の期待値は

$$\mathsf{P}g := \mathsf{E}[g(X)] \qquad (\mathsf{P}|g| < \infty)$$

であり, $X_1, X_2, \ldots, X_n$ に基づく, g の（経験的）期待値は

$$\widehat{\mathsf{P}}_n g := \frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

である. $\widehat{\mathsf{P}}_n$ を経験測度という.

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

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

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

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

で記す.

$\sigma^2(g) < \infty$ のとき,

$$\nu_n(g) \rightsquigarrow \mathsf{N}(0,\, \sigma^2(g)) \qquad (n \to \infty)$$

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, $g_k,\, g_\ell \in \mathcal{G}\,(k,\, \ell = 1,\, 2,\, \ldots,\, d\, k \neq \ell)$ は異なる関数とし, $g_k(X)$ と $g_\ell(X)$ の共分散を

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

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

$$\begin{pmatrix} \nu_n(g_1) \\ \nu_n(g_2) \\ \vdots \\ \nu_n(g_d) \end{pmatrix} \rightsquigarrow \mathsf{N}_d(\mathbf{0},\, \Sigma_{g_1,\, g_2,\, \ldots,\, g_r}) \qquad (n \to \infty)$$

となる. ただし, $\boldsymbol{\Sigma}_{g_1,\, g_2,\, \ldots,\, g_r}$ は対称行列で

$$\boldsymbol{\Sigma}_{g_1,\, g_2,\, \ldots,\, g_r} = \begin{pmatrix} \sigma^2(g_1) & \sigma(g_1,\, g_2) & \cdots & \sigma(g_1,\, g_d) \\ \sigma(g_2,\, g_1) & \sigma^2(g_2) & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(g_d,\, g_1) & \cdots & \cdots & \sigma^2(g_d) \end{pmatrix} \tag{8.2}$$

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 $\{g_1,\, g_2,\, \ldots,\, g_d\} \subset \mathcal{G}$ 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 $\mathsf{N}_d(\mathbf{0},\, \boldsymbol{\Sigma}_{g_1,\, g_2,\, \ldots,\, g_r})$ を持つとする. ただし, $\boldsymbol{\Sigma}_{g_1,\, g_2,\, \ldots,\, g_r}$ 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 $\mathcal{G}$ で添字付けれた P-Brown 橋という.

定義 8.15. $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{P}$ 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 **??**) 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, $\nu_n, \nu \in \mathcal{K} = \{k : \mathcal{G} \longrightarrow \mathbb{R}; k \text{ は有界}\}$ である. 関数族 $\mathcal{G}$ は P-Donsker 族であるとは

$$\nu_n \rightsquigarrow \nu \qquad (n \to \infty)$$

をみたすときをいう. すなわち, $\mathcal{K}$ 上の任意の有界連続実数値関数[1] h に対して

$$\mathsf{E}^*[h(\nu_n)] \overset{n\to\infty}{\longrightarrow} \mathsf{E}[h(\nu)]$$

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 $\sigma(g_n - g_0) \overset{n\to\infty}{\longrightarrow} 0$ なるものに対して

$$|\nu_n(g_n) - \nu_n(g_0)| \overset{\mathrm{P}}{\to} 0 \qquad (n \to \infty)$$

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.

<u>記法</u>: $g \in L_2(\mathsf{P})$ に対して

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

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

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

定義 8.18. 族 $\mathcal{G}$ は全有界であるとは, $\|\cdot\|_{L_2(\mathsf{P})}$ 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 $H_2(\cdot, \mathcal{G}, \|\cdot\|_{L_2(\mathsf{P})})$ は有界のときである.

定理 8.19. $\mathbb{M}$ を距離空間とし, $\mathbb{M}$ 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 は漸近連続であることである.

Proof. 補題 8.20 として証明する. □

---

[1] 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.

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 $\mathcal{G}$ で添え字づけれた確率過程とし, ν_n に見本路はほとんど確実に $\ell_\infty(\mathcal{G})$ に属すとする. ただし

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

である. 以下のふたつの条件

- 関数族 $\mathcal{G} \subset L_2(\mathsf{P})$ 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

$$\limsup_{n\to\infty} \mathsf{Pr}\left(\sup_{g_1,\, g_2\in\mathcal{G};\ \sqrt{\mathsf{P}|g_1-g_2|^2}\le\delta} \big|\nu_n(g_1) - \nu_n(g_2)\big| > \eta \right) \le \eta \tag{8.3}$$

  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

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

とできる. さらに, $k = \#(\mathcal{G}_\delta)$ と [2] おく. いま, 関数 $h : \mathbb{M} \longrightarrow \mathbb{R}$ は Lipschitz [3] とする.

I. chaining argument のための射影の定義. $\mathcal{G}$ から $\mathcal{G}_\delta$ 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 $\mathcal{G}_\delta$ のある元 g_δ で

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

をみたすもの [4] 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, $\big\{g_1, g_2, \dots, g_k\big\} := \mathcal{G}_\delta$ と書いたとき, $\nu_n \circ \Pi_\delta$ の値域は $\big\{\nu_n(g_1), \nu_n(g_2), \dots, \nu_n(g_k)\big\}$ となる. このとき, Lipschitz 関数 $\widetilde{h} : \mathbb{R}^k \longrightarrow \mathbb{R}$ が存在して, $h \in C_b(\mathbb{M})$ に対して

$$h(\nu_n \circ \Pi_\delta) = \widetilde{h}(\nu_n \circ \Pi_\delta) \tag{8.4}$$

[2]すなわち, 有限部分集合 $\mathcal{G}_\delta$ の要素の個数である.

[3]すなわち, ある $\exists\lambda > 0$ が存在して, 任意の $\forall x, y \in \mathbb{M}$ に対して, $\big|h(x) - h(y)\big| \le \lambda d(x, y)$ となる.

[4]$\mathcal{G}_\delta$ は $\mathcal{G}$ の稠密な有限部分集合なので, そのような $\mathcal{G}_\delta$ 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.

と書ける. 古典的な CLT から

$$\nu_n \Pi_\delta \rightsquigarrow \mathsf{N}_k\big(\mathbf{0}_k,\, \mathbf{\Sigma}\big) \qquad (n \to \infty)$$

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

$$\mathsf{Cov}\big(\nu(g_1),\, \nu(g_2)\big) = \mathsf{P}g_1 g_2 - \mathsf{P}g_1 \mathsf{P}g_2$$

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

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

となり

$$\nu_n \circ \Pi_\delta \rightsquigarrow \nu \circ \Pi_\delta \qquad (n \to \infty) \tag{8.5}$$

となる. すなわち

$$\mathsf{E}\big[h(\nu_n \circ \Pi_\delta)\big] = \mathsf{E}\big[\widetilde{h}(\nu_n \circ \Pi_\delta\big] \overset{n\to\infty}{\longrightarrow} \mathsf{E}\big[\widetilde{h}(\nu \circ \Pi_\delta\big] = \mathsf{E}\big[h(\nu \circ \Pi_\delta)\big]$$

となる. ただし, h に対応した $\widetilde{h}$ は Lipschitz 関数 ((10.3) を参考のこと) である.

次に, (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, $\widetilde{\mathcal{G}} \subset \mathcal{G}$ を有限部分集合とする. このとき, Partmanteau の定理から

$$\begin{aligned}
&\mathsf{Pr}\Bigg(\sup_{g_1,\, g_2,\, \in \widetilde{\mathcal{G}};\, \|g_1 - g_2\|_{L_2(\mathsf{P})} \le \delta} \big|\nu(g_1) - \nu(g_2)\big| > \eta\Bigg) \\
&\quad \le \liminf \mathsf{Pr}\Bigg(\sup_{g_1,\, g_2,\, \in \widetilde{\mathcal{G}};\, \|g_1 - g_2\|_{L_2(\mathsf{P})} \le \delta} \big|\nu_n(g_1) - \nu_n(g_2)\big| > \eta\Bigg) \\
&\quad \le \liminf \mathsf{Pr}\Bigg(\sup_{g_1,\, g_2,\, \in \mathcal{G};\, \|g_1 - g_2\|_{L_2(\mathsf{P})} \le \delta} \big|\nu_n(g_1) - \nu_n(g_2)\big| > \eta\Bigg) \\
&\quad \le \limsup \mathsf{Pr}\Bigg(\sup_{g_1,\, g_2,\, \in \mathcal{G};\, \|g_1 - g_2\|_{L_2(\mathsf{P})} \le \delta} \big|\nu_n(g_1) - \nu_n(g_2)\big| > \eta\Bigg) \le \eta
\end{aligned}$$

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, $\widetilde{\mathcal{G}}$ 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 $\widetilde{\mathcal{G}}$ を取ることができる. したがって, $\mathcal{G}$ 上の確率過程 $\widetilde{\nu}$ 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, $g \in \widetilde{\mathcal{G}}$ のときは

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

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

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

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

$$\mathsf{Pr}\Biggl(\sup_{g_1,\, g_2,\, \in\widetilde{\mathcal{G}};\, \|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta} \bigl|\widetilde{\nu}(g_1)-\widetilde{\nu}(g_2)\bigr| > \eta\Biggr)$$

が成立している.

II. $\nu_n \rightsquigarrow \widetilde{\nu}\,(n\to\infty)$ の証明. まず

$$\begin{aligned}
&\mathsf{E}\bigl[h(\nu_n)\bigr]-\mathsf{E}\bigl[h(\widetilde{\nu})\bigr]\\
&\quad=\bigl\{\mathsf{E}\bigl[h(\nu_n)\bigr]-\mathsf{E}\bigl[h(\nu_n\circ\Pi_\delta)\bigr]\bigr\}+\bigl\{\mathsf{E}\bigl[h(\nu_n\circ\Pi_\delta)\bigr]-\mathsf{E}\bigl[h(\nu\circ\Pi_\delta)\bigr]\bigr\}\\
&\qquad+\bigl\{\mathsf{E}\bigl[h(\widetilde{\nu}\circ\Pi_\delta)\bigr]-\mathsf{E}\bigl[h(\widetilde{\nu})\bigr]\bigr\}
\end{aligned}\tag{8.6}$$

と書ける. しかし

$$\begin{aligned}
&\mathsf{E}\bigl[h(\nu_n)\bigr]-\mathsf{E}\bigl[h(\nu_n\circ\Pi_\delta)\bigr]\\
&\quad=\mathsf{E}\Biggl[\bigl\{h(\nu_n)-h(\nu_n\circ\Pi_\delta)\bigr\}\mathbb{1}\{\sup_{g_1,\, g_2\in\mathcal{G};\, \|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\bigl|\nu_n(g_1)-\nu_n(g_2)\bigr|\le\eta\}\Biggr]\\
&\qquad+\mathsf{E}\Biggl[\bigl\{h(\nu_n)-h(\nu_n\circ\Pi_\delta)\bigr\}\mathbb{1}\{\sup_{g_1,\, g_2\in\mathcal{G};\, \|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\bigl|\nu_n(g_1)-\nu_n(g_2)\bigr|>\eta\}\Biggr]\\
&\quad\le\mathsf{E}\Biggl[\underbrace{\bigl|h(\nu_n)-h(\nu_n\circ\Pi_\delta)\bigr|}_{\le\lambda\eta\because h\text{ は Lipschitz}}\mathbb{1}\{\sup_{g_1,\, g_2\in\mathcal{G};\, \|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\bigl|\nu_n(g_1)-\nu_n(g_2)\bigr|\le\eta\}\Biggr]\\
&\qquad+2\|h\|_\infty\mathsf{E}\Biggl[\mathbb{1}\{\sup_{g_1,\, g_2\in\mathcal{G};\, \|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\bigl|\nu_n(g_1)-\nu_n(g_2)\bigr|>\eta\}\Biggr]\\
&\quad=\lambda\eta+2\|h\|_\infty\mathsf{Pr}\Biggl(\sup_{g_1,\, g_2\in\mathcal{G};\, \|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\bigl|\nu_n(g_1)-\nu_n(g_2)\bigr|>\eta\Biggr)\\
&\quad\stackrel{n\to\infty}{\longrightarrow}\lambda\eta
\end{aligned}$$

がわかる. 最後に, $\eta\to 0$ とすれば

$$\mathsf{E}\bigl[h(\nu_n)\bigr]-\mathsf{E}\bigl[h(\nu_n\circ\Pi_\delta)\bigr]\stackrel{n\to\infty}{\longrightarrow}0\tag{8.7}$$

がわかる. 同様に

$$
\begin{aligned}
&\mathsf{E}\big[h(\widetilde{\nu})\big] - \mathsf{E}\big[h(\widetilde{\nu}\circ\Pi_\delta)\big] \\
&\quad = \mathsf{E}\Bigg[\big\{h(\widetilde{\nu}) - h(\widetilde{\nu}\circ\Pi_\delta)\big\}\mathbb{1}\{\sup_{g_1,\,g_2\in\mathcal{G};\,\|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\big|\nu_n(g_1)-\nu_n(g_2)\big| \le \eta\}\Bigg] \\
&\qquad + \mathsf{E}\Bigg[\big\{h(\nu_n) - h(\nu_n\circ\Pi_\delta)\big\}\mathbb{1}\{\sup_{g_1,\,g_2\in\mathcal{G};\,\|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\big|\widetilde{\nu}(g_1)-\widetilde{\nu}(g_2)\big| > \eta\}\Bigg] \\
&\quad \le \mathsf{E}\Bigg[\underbrace{\big|h(\widetilde{\nu}) - h(\widetilde{\nu}\circ\Pi_\delta)\big|}_{\le\lambda\eta\,\because h\text{ は Lipschitz}}\mathbb{1}\{\sup_{g_1,\,g_2\in\mathcal{G};\,\|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\big|\widetilde{\nu}(g_1)-\widetilde{\nu}(g_2)\big| \le \eta\}\Bigg] \\
&\qquad + 2\|h\|_\infty\mathsf{E}\Bigg[\mathbb{1}\{\sup_{g_1,\,g_2\in\mathcal{G};\,\|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\big|\widetilde{\nu}(g_1)-\widetilde{\nu}(g_2)\big| > \eta\}\Bigg] \\
&\quad = \lambda\eta + 2\|h\|_\infty\mathsf{Pr}\Bigg(\sup_{g_1,\,g_2\in\mathcal{G};\,\|g_1-g_2\|_{L_2(\mathsf{P})}\le\delta}\big|\widetilde{\nu}(g_1)-\widetilde{\nu}(g_2)\big| > \eta\Bigg) \\
&\quad \xrightarrow{\eta\to0;\,\delta\to0}\ 0
\end{aligned}
$$

から

$$
\mathsf{E}\big[h(\widetilde{\nu})\big] - \mathsf{E}\big[h(\widetilde{\nu}\circ\Pi_\delta)\big] \xrightarrow{n\to\infty} 0 \tag{8.8}
$$

がわかる. 最後に, (10.4) から

$$
\mathsf{E}\big[h(\nu_n\circ\Pi_\delta)\big] - \mathsf{E}\big[h(\nu\circ\Pi_\delta)\big] \xrightarrow{n\to\infty} 0 \tag{8.9}
$$

がわかる. (10.5) – (8.9) を合わせると定理は証明できる. □

8.6 P-Donsker の定理

定理 8.21. $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim}$ P とする. ノルム空間 $\big(\mathcal{G}, \|\cdot\|\big) \subset L_2(\mathsf{P})$ は封筒関数 $G \in L_2(\mathsf{P})$ 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

- $\displaystyle\int_0^1 H(\delta)\,\mathrm{d}\delta < \infty$,
- $\displaystyle\lim_{A\to\infty}\limsup_n \mathsf{Pr}\Bigg(\sup_{\delta>0}\frac{H\big(\delta, \mathcal{G}, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big)}{H(\delta)} > A\Bigg) = 0$ 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

$$
\limsup_{n\to\infty}\mathsf{Pr}\Bigg(\sup_{g_1,\,g_2\in\mathcal{G};\,\sqrt{\mathsf{P}|g_1-g_2|^2}\le\delta}\big|\nu_n(g_1)-\nu_n(g_2)\big| > \eta\Bigg) \le \eta
$$

となる．ただし, $\nu_n = \sqrt{n}\big(\widehat{\mathsf{P}}_n - \mathsf{P}\big)$ 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．

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

$$E_{n,\,M} := \left\{ \sup_{\delta > 0} \frac{H\big(\delta, \mathcal{G}, \|\cdot\|_{L2(\widehat{\mathsf{P}}_n)}\big)}{H(\delta)} > M \right\}$$

とおく．このとき

$$\begin{aligned}
\mathsf{Pr}(A_\delta) &= \mathsf{E}\big[\mathbb{1}_{A_\delta}\mathbb{1}_{E_{n,\,M}}\big] + \mathsf{E}\big[\mathbb{1}_{A_\delta}\mathbb{1}_{E^c_{n,\,M}}\big]\big] \\
&\le \mathsf{Pr}(E_{n,M}) + \frac{\mathsf{E}\left[\left\{ \sup_{g_1,\, g_2 \in \mathcal{G};\, \|g_1 - g_2\|_{L_2(\mathsf{P})}} \big|\nu_n(g_1) - \nu_2(g_2)\big| \right\} \mathbb{1}_{E^c_{n,\,M}}\right]}{\eta} \\
&\quad \left(\because\ A_\delta \Rightarrow \frac{\sup_{g_1,\, g_2 \in \mathcal{G};\, \|g_1 - g_2\|_{L_2(\mathsf{P})}} \big|\nu_n(g_1) - \nu_2(g_2)\big|}{\eta} > 1 \right) \\
&\le \mathsf{Pr}(E_{n,M}) + \frac{\mathsf{E}\left[\left\{ \sup_{g_1,\, g_2 \in \mathcal{G};\, \|g_1 - g_2\|_{L_2(\mathsf{P})}} \big|\nu_n(g_1 - g_2)\big| \right\} \mathbb{1}_{E^c_{n,\,M}}\right]}{\eta} \\
&\le \mathsf{Pr}(E_{n,M}) + 24\frac{\mathsf{E}\left[\int_0^{\delta_n/2} \sqrt{H\big(\epsilon, \mathcal{G}^\delta, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big)}\,\mathrm{d}\epsilon \mathbb{1}_{E^c_{n,\,M}}\right]}{\eta} \\
&\quad \big(\because (6.31)\big)
\end{aligned}$$

となる． ただし, $\mathcal{G}^\delta := \big\{g_1 - g_2;\, g_1,\, g_2 \in \mathcal{G};\, \|g_1 - g_2\|_{L_2(\mathsf{P})} \le \delta\big\}$ と $\delta_n := \sup_{g \in \mathcal{G}^\delta} \|g\|_{L_2(\widehat{\mathsf{P}}_n)}$ である．まず, $\mathcal{G}$ の ϵ 被覆は $\mathcal{G}^\delta$ 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 $g_1, g_2, \ldots, g_k$ と書く．このとき, $g^{\mathrm{diff}}_{i,j} := g_i - g_j$ とおくと, これを中心とした, 2ϵ 球全体は $\mathcal{G}^\delta$ の被覆となることがわかる．よって, $\#\{g^{\mathrm{diff}}_{i,j};\, i,\, j = 1, 2, \ldots, k\} = k^2$ となるので

$$H\big(2\epsilon, \mathcal{G}^\delta, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big) \le 2H\big(\epsilon, \mathcal{G}, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big) \tag{8.10}$$

がわかる．

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

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

となるので

$$\big\|g_1 - g_2\big\|^2_{L_2(\widehat{\mathsf{P}}_n)} \le 2\big\|G\big\|_{L_2(\widehat{\mathsf{P}}_n)}\big\|g_1 - g_2\big\|_{L_2(\widehat{\mathsf{P}}_n)}$$

がわかる. 古典的な大数の強法則から

$$\|G\|_{L_2(\widehat{\mathsf{P}}_n)} \overset{\text{a.s.}}{\to} \|G\|_{L_2(\mathsf{P})} \qquad (n \to \infty)$$

がわかる. このことから, 十分大きな n に対して

$$\big\|g_1 - g_2\big\|^2_{L_2(\widehat{\mathsf{P}}_n)} \le 4\big\|G\big\|_{L_2(\mathsf{P})}\big\|g_1 - g_2\big\|_{L_2(\widehat{\mathsf{P}}_n)} \tag{8.11}$$

となる. いま

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

とおく. すると (8.10) と (8.11) から, $\forall\epsilon > 0$ に対して

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

となることがわかる. 上の不等式と定理の条件から

$$\begin{aligned} &\frac{1}{n}H\big(\epsilon,\, \mathcal{H},\, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big) \\ &\qquad \le \frac{1}{n}H\left(\frac{\epsilon}{8\|G\|_{L_2(\mathsf{P})}},\, \mathcal{G},\, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\right) \overset{\mathrm{P}}{\to} 0 \qquad (n \to \infty) \end{aligned} \tag{8.12}$$

がわかる. さらに, $G \in L_2(\mathsf{P})$ 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[5] $4G^2(x)$ は $L_1(\mathsf{P})$ となることがわかる. よって, (8.12) と $\mathcal{H}$ の封筒関数が $L_1(\mathsf{P})$ 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

$$\sup_{g_1,\, g_2 \in \mathcal{G}}\left|\big\|g_1 - g_2\big\|_{L_2(\widehat{\mathsf{P}}_n)} - \big\|g_1 - g_2\big\|_{L_2(\mathsf{P})}\right| \overset{\text{a.s.}}{\to} 0 \qquad (n \to \infty)$$

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

$$\frac{1}{2}\big\|g_1 - g_2\big\|_{L_2(\widehat{\mathsf{P}}_n)} \le \big\|g_1 - g_2\big\|_{L_2(\mathsf{P})} \le 2\big\|g_1 - g_2\big\|_{L_2(\widehat{\mathsf{P}}_n)} \tag{8.13}$$

となる. したがって, 十分大きな n に対して

$$H\big(2\epsilon,\, \mathcal{G},\, \|\cdot\|_{L_2(\mathsf{P})}\big) \le H\big(\epsilon,\, \mathcal{G},\, \|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big)$$

[5]すなわち, $\mathcal{H}$ の封筒関数である.

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

$$\delta_n \leq 2\delta$$

となること[6]もわかる. (8.10) から

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

となるので

$$\begin{aligned}\mathsf{Pr}(A_\delta) &\leq \mathsf{Pr}(E_{n,M}) + 48\sqrt{2}\frac{\mathsf{E}\Bigl[\int_0^{\delta/2}\sqrt{H\bigl(\epsilon,\,\mathcal{G},\,\|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\bigr)}\,\mathrm{d}\epsilon\mathbb{1}_{E^c_{n,M}}\Bigr]}{\eta}\\ &\leq \mathsf{Pr}(E_{n,M}) + 48\sqrt{2}\frac{\mathsf{E}\Bigl[\int_0^{\delta/2}\sqrt{H(\epsilon)}\,\mathrm{d}\epsilon\Bigr]}{\eta}\end{aligned}$$

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

定理 8.22. $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{P}$ とし, $\mathcal{G} \subset L_2(\mathsf{P})$ 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.

- $G \in L_2(\mathsf{P})$,
- $\displaystyle\int_0^1 H(\delta)\,\mathrm{d}\delta < \infty$,
- $\displaystyle\lim_{A\to\infty}\limsup_n \mathsf{Pr}\Biggl(\sup_{\delta>0}\frac{H\bigl(\delta,\,\mathcal{G},\,\|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\bigr)}{H(\delta)} > A\Biggr) = 0$

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. □

[6] $\delta_n = \sup_{g\in\mathcal{G}^{\mathrm{delta}}}\|g\|_{L_2(\widehat{\mathsf{P}}_n)}$ だったので, (8.13) から直ちにわかる.

8.7 一様エントロピー条件

定理 8.23. $\bigl(\Omega, \mathcal{A}, \mathsf{Pr}\bigr)$ 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

$$\int_0^\infty \sup_{\mathsf{Q}} \sqrt{\log N\bigl(\epsilon \|G\|_{L_2(\mathsf{Q})}, \mathcal{G}, \|\cdot\|_{L_2(\mathsf{Q})}\bigr)}\,\mathrm{d}\epsilon < \infty \tag{8.14}$$

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 $\|G\|_{L_2(\mathsf{Q})} := \int G^2\,\mathrm{d}\mathsf{Q} > 0$ 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 $\mathcal{G}^\delta = \bigl\{g_1 - g_2;\, g_1, g_2 \in \mathcal{G}, \|g_1 - g_2\|_{L_2(\mathsf{P})} < \delta\bigr\}$ と $\mathcal{G}^2_\infty = \bigl\{g^2;\, g \in \mathcal{G}, \sup_{x\in\mathbb{X}} |g(x)| < \infty\bigr\}$ 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. まず, 補題 8.20 から

I ν_n が漸近連続であること,

II $\mathcal{G}$ が $L_2(\mathsf{P})$ において全有界でこと

確認すればよいことに注意する.
<u>**I の証明**</u>. そのために, $\{\delta_n\}_{n=1}^\infty$ を $\delta_n \downarrow 0\,(n \to \infty)$ なる任意の数列としたとき

$$\mathsf{Pr}\Biggl(\sup_{g\in\mathcal{G}^{\delta_n}} \|\nu_n(g)\| > x\Biggr) \overset{n\to\infty}{\longrightarrow} 0$$

を確認すればよいことに注意する.
<u>① $\mathsf{Pr}\Bigl(\sup_{g\in\mathcal{G}^{\delta_n}} \|\nu_n(g)\| > x\Bigr)$ **の上限の導出**</u>. 任意の $x > 0$ に対して, Markov の不等式から

$$\mathsf{Pr}\Biggl(\sup_{g\in\mathcal{G}^\delta} \|\nu_n(g)\| > x\Biggr) \le \frac{\mathsf{E}\Bigl[\sup_{g\in\mathcal{G}^\delta} |\nu_n(g)|\Bigr]}{x} \tag{8.15}$$

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, $Y_1, Y_2, \ldots, Y_n$ を $X_1, X_2, \ldots, X_n$ の独立複製とし, $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ 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

$$
\begin{aligned}
\mathsf{E}\left[\sup_{g\in\mathcal{G}^\delta}|\nu_n(g)|\right] &= \mathsf{E}\left[\sup_{g\in\mathcal{G}^\delta}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^n\bigl\{g(X_j)-\mathsf{P}g\bigr\}\right|\right] \\
&= \mathsf{E}\left[\sup_{g\in\mathcal{G}^\delta}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^n\bigl\{g(X_j)-\mathsf{E}^Y[g(Y_j)]\bigr\}\right|\right] \\
&\le \mathsf{E}\left[\sup_{g\in\mathcal{G}^\delta}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^n\bigl\{g(X_j)-g(Y_j)\bigr\}\right|\right] \\
&\qquad (\because \text{Jensen の不等式}) \\
&= \mathsf{E}\left[\sup_{g\in\mathcal{G}^\delta}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^n\epsilon_j\bigl\{g(X_j)-g(Y_j)\bigr\}\right|\right] \\
&\le 2\mathsf{E}\left[\sup_{g\in\mathcal{G}^\delta}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^n\epsilon_j g(X_j)\right|\right] \qquad (8.16)
\end{aligned}
$$

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

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

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

$$
\frac{1}{\sqrt{n}}\sum_{j=1}^n\epsilon_j g(X_j)\ \sim\ \mathsf{subG}\bigl(\|g\|_{L_2(\widehat{\mathsf{P}}_n)}\bigr)
$$

となる. 定理 6.27 から

$$
\begin{aligned}
&\mathsf{E}\Big[\sup_{g\in\mathcal{G}^{\delta_n}}\Big|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\epsilon_j g(X_j)\Big|\,\Big|\,\boldsymbol{X}\Big]\\
&\quad\le \mathsf{E}\Big[\sup_{g_1,\,g_2\in\mathcal{G}}\Big|\Big\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\epsilon_j g_1(X_j)-\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\epsilon_j g_2(X_j)\Big\}\Big|\,\Big|\,\boldsymbol{X}\Big]\\
&\quad\le 24\int_0^\infty \sqrt{\log N\big(\epsilon,\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\widehat{\mathsf{P}}_n)}\big)}\,\mathrm{d}\epsilon\\
&\quad\le 24\int_0^\infty \sup_{\mathsf{Q}}\sqrt{\log N\big(\epsilon,\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)}\,\mathrm{d}\epsilon \qquad (8.17)
\end{aligned}
$$

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

② 積分 $\int_0^\infty \sup_{\mathsf{Q}}\sqrt{\log N\big(\epsilon,\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)}\,\mathrm{d}\epsilon$ の評価. 十分大きな ϵ に対して, 関数族 $\mathcal{G}^{\delta}$ 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

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

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

$$
N\big(\epsilon,\,\mathcal{G}^{\mathrm{diff}},\,\|\cdot\|_{L_2(\mathsf{Q})}\big) \le N^2\big(\epsilon/2,\,\mathcal{G},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)
$$

と評価できる. ただし, 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

$$
\begin{aligned}
&\int_0^{\theta_n}\sup_{\mathsf{Q}}\sqrt{\log N\big(\epsilon,\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)}\,\mathrm{d}\epsilon\\
&\quad= \int_0^{\theta_n/\|G\|_{L_2(\widehat{\mathsf{P}}_n)}}\sup_{\mathsf{Q}}\sqrt{\log N\big(\epsilon\|G\|_{L_2(\widehat{\mathsf{P}}_n)},\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)}\,\mathrm{d}\epsilon\times\|G\|_{L_2(\widehat{\mathsf{P}}_n)}\\
&\quad\le \int_0^{\theta_n/\|G\|_{L_2(\widehat{\mathsf{P}}_n)}}\sup_{\mathsf{Q}}\sqrt{\log N\big(\epsilon\|G\|_{L_2(\mathsf{Q})},\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)}\,\mathrm{d}\epsilon\times\|G\|_{L_2(\widehat{\mathsf{P}}_n)}
\end{aligned}
$$

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

$$
\|G\|_{L_2(\widehat{\mathsf{P}}_n)} \overset{\text{a.s.}}{\to} \|G\|_{L_2(\mathsf{P})} \qquad (n\to\infty)
$$

がわかる. Cauchy-Schwarz の不等式から

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

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

$$\mathsf{E}\left[\int_0^{\theta_n/\|G\|_{L_2(\widehat{\mathsf{P}}_n)}}\sup_{\mathsf{Q}}\sqrt{\log N\big(\epsilon\|G\|_{L_2(\mathsf{Q})},\,\mathcal{G}^{\delta_n},\,\|\cdot\|_{L_2(\mathsf{Q})}\big)}\,\mathrm{d}\epsilon\times\|G\|_{L_2(\widehat{\mathsf{P}}_n)}\right]\xrightarrow{n\to\infty}0 \tag{8.18}$$

がわかる. (8.15) – (8.18) を合わせると, $\theta_n\xrightarrow{\mathrm{P}}0\,(n\to\infty)$ がわかれば

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

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 $\theta_n\xrightarrow{\mathrm{P}}0\,(n\to\infty)$ を示せばよい.

③ $\theta_n\xrightarrow{\mathrm{P}}0\,(n\to\infty)$ **の証明**.

$$\sup\big\{\mathsf{P}g^2;\,g\in\mathcal{G}^{\delta_n}\big\}\xrightarrow{n\to\infty}0$$

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

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

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

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

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

となるので

$$N\Bigl(4\epsilon\left\|G\right\|^2_{L_2(\widehat{\mathsf{P}}_n)},\, \bigl\{\mathcal{G}^{\mathrm{diff}}\bigr\}^2,\, L_1(\widehat{\mathsf{P}}_n)\Bigr) \le N\Bigl(\epsilon\left\|G\right\|_{L_2(\widehat{\mathsf{P}}_n)},\, \mathcal{G}^{\mathrm{diff}},\, L_2(\widehat{\mathsf{P}}_2)\Bigr)$$

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

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

がわかる. 定理 6.20 から関数族 $\bigl\{\mathcal{G}^{\mathrm{diff}}\bigr\}^2$ 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

$$\sup_{g\in\mathcal{G}^{\mathrm{diff}}}\left|\bigl(\widehat{\mathsf{P}}_n - \mathsf{P}\bigr)g^2\right| \xrightarrow{\mathrm{a.s.}} 0 \qquad (n \to \infty)$$

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

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

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

例 8.24. (1). $\mathcal{G} := \bigl\{\mathbb{1}_{(\infty,r]};\, r \in \mathbb{R}\bigr\}$ とする. このとき, $\mathbb{R}$ 上の任意の有限離散確率測度 Q と $\forall 1 \le \epsilon > 0$ に対して

$$N\bigl(\epsilon,\, \mathcal{G},\, L_2(\mathsf{Q})\bigr) \le N_B\bigl(\epsilon^2,\, \mathcal{G},\, L_1(\mathsf{Q})\bigr) \le \frac{2}{\epsilon^2}$$

となる.

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

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