第4章　確率変数列と分布列の収束

確率論の最も重要な側面のひとつは確率変数列の挙動に関することである．確率論のこの部分のことを「大標本論」,「極限論」,「漸近論」と数理統計学では呼んでいる．大標本論の基本的な問いは次である．確率変数列 $X_1, X_2, \ldots$ の極限の振る舞いについて言えることは何だろうか? 統計学はデータの収集にかかわる学問である．したがってデータを集めれば集めるほど何が起こるかを調べることは重要である．

すこし実解析の話題を復習する．実数列 $\{x_n\}$ が点 x に収束するとは, 任意の $\epsilon > 0$ に対してある正の整数 $N \in \mathbb{N}$ が存在して

$$\forall n > N \Rightarrow |x_n - x| < \epsilon$$

が成り立つことである．このことを $\lim_{n\to\infty} x_n = x$ と書いた．たとえば $x_n = x\,(\forall n \in \mathbb{N})$ ならば, 明らかに $\lim_{n\to\infty} x_n = x$ となる．これと同じことを確率空間 $\big(\Omega, \mathcal{A}, \mathsf{Pr}\big)$ 上の確率変数列について考えてみる．

$$X_1, X_2, \ldots \overset{\text{i.i.d.}}{\sim} \mathsf{N}(0, 1)$$

とする．すなわちこの確率変数列の任意の有限個の確率変数列は独立同一に $\mathsf{N}(0, 1)$ に従う．さらに別の確率変数 X も $\mathsf{N}(0, 1)$ に従うとする．このとき X_n は X に「収束」するとしたい．しかし

$$\mathsf{Pr}\big(X_n \neq X\big) = 1 \qquad (\forall n \in \mathbb{N})$$

である．

別の例をあげる．$X_1, X_2, \ldots$ は独立で

$$X_n \sim \mathsf{N}\left(0, \frac{1}{n}\right) \qquad (n = 1, 2, \ldots)$$

とする．直観的には, 十分大きな n に対して X_n は 0 の近辺に集中すると予想するだろう．しかし X_n は連続型確率変数なので, すべての n に対して

$$\mathsf{Pr}\big(X_n = 0\big) = 0$$

である．これらの例から, 確率変数列の収束については, 実数列の収束とは異なる道具立てが必要になることがわかる．この章では確率変数列の収束の定義を述べ, これに関わる基本的な事項をまとめたうえで次の重要な事項を説明する．

(1) **大数の法則**. $X_1, X_2, \ldots$ を i.i.d. 確率変数列とし

$$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j, \qquad \mu = \mathsf{E}[X_1]\,(-\infty < \mu < \infty)$$

とする．このとき $\overline{X}_n$ は高い確率で μ の近くにいることを保証する定理である．

(2) **中心極限定理**. $\mathsf{Var}[X_1] = \sigma^2\,(0 < \sigma < \infty)$ とする．n が十分大きいとき $\sqrt{n}\big(\overline{X}_n - \mu\big)$ の分布は正規分布で近似できることを保証する定理である．

以上の事項について証明なしで主張まずを紹介する．証明については補遺 G にまとめることにする．

4.1 確率変数列の収束のタイプ

定義 4.1. 確率空間 $\big(\Omega, \mathcal{A}, \mathsf{Pr}\big)$ 上の $X, X_1, X_2, \ldots$ を確率変数列とする．さらに, 各 $X_n\,(n = 1, 2, \ldots)$ の c.d.f. を F_n, X の c.d.f. を F とする．

(1) 確率変数列 $\{X_n\}$ は確率変数 X に**確率収束**するとは, $\forall\epsilon > 0$ に対して

$$\lim_{n\to\infty} \mathsf{Pr}\big(\big|X_n - X\big| \geq \epsilon\big) = 0$$

が成り立つときをいう．このことを $X_n \xrightarrow{\mathrm{P}} X$ と書く．

(2) 確率変数列 $\{X_n\}$ は確率変数 X に**概収束**するとは

$$\mathsf{Pr}\left(\omega \in \Omega;\, \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 0$$

が成り立つときをいう．このことを $X_n \xrightarrow{\mathrm{a.s.}} X$ と書く．

(3) 確率変数列 $\{X_n\}$ は確率変数 X に**分布収束**するとは, F のすべての連続点 x において

$$\lim_{n\to\infty} \mathsf{F}_n(x) = \mathsf{F}(x)$$

が成り立つときをいう．このことを $X_n \rightsquigarrow X$ と書く．また, $X_n \rightsquigarrow X$ かつ $X \sim \mathsf{N}(0, 1)$ のときには $X_n \rightsquigarrow \mathsf{N}(0, 1)$ とも書くことがある．

注意 4.2. (1) c を定数とする．$\mathsf{Pr}(X = c) = 1$ かつ $X_n \xrightarrow{\mathrm{P}} X$ とき $X_n \xrightarrow{\mathrm{P}} c$ と記す.
(2) $\mathsf{Pr}(X = c) = 1$ かつ $X_n \rightsquigarrow X$ とき $X_n \rightsquigarrow c$ と記す. □

定義 4.3. $X_1, X_2, \ldots$ を確率変数列とし X を別の確率変数とする．さらに $\mathsf{E}\bigl[X_n^2\bigr] < \infty$, $\mathsf{E}[X^2] < \infty\,(n = 1, 2, \ldots)$ とする．確率変数列 $\{X_n\}$ が確率変数 X に**平均 2 乗の意味で収束**するとは

$$\lim_{n\to\infty} \mathsf{E}\bigl[(X_n - X)^2\bigr] = 0$$

が成り立つときをいう．このことを $X_n \stackrel{\text{q.m.}}{\to} X$ と記す

注意 4.4. 確率収束と概収束の違いを説明する．$X, X_1, X_2, \ldots$ を同じ確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上で定義された確率変数列とし

$$M_n(\omega) := \sup_{k\geq n} |X_k(\omega) - X(\omega)| \qquad (\omega \in \Omega)$$

とおく．すると

$$X_n \stackrel{\text{a.s.}}{\to} X \Leftrightarrow M_n \stackrel{\mathrm{P}}{\to} 0$$

となることにが以下の議論からわかる.

まず, $\forall \ell \in \mathbb{N}$ に対して, ある十分大きな $n_0 \in \mathbb{N}$ があって

$$\forall n \geq n_0 \Rightarrow M_n(\omega) < \frac{1}{\ell}$$

ならば

$$M_n(\omega) \xrightarrow{n\to\infty} 0$$

となることを示す．そのために, $\ell \in \mathbb{N}$ に対して

$$B_\ell := \bigcup_{n=n_0}^{\infty} \left\{ \omega \in \Omega;\, M_n(\omega) < \frac{1}{\ell} \right\}$$

と定めると

$$\left\{ \omega \in \Omega;\, \lim_{n\to\infty} X_n(\omega) = X(\omega) \right\} = \bigcap_{\ell=1}^{\infty} B_\ell$$

と書けるので

$$X_n \stackrel{\text{a.s.}}{\to} X \Leftrightarrow \mathsf{Pr}\left(\bigcap_{\ell=1}^{\infty} B_\ell \right) = 1$$

がわかる.

次に, $\ell_1 < \ell_2\,(\ell_1,\,\ell_2 \in \mathbb{N})$ に対し

$$B_{\ell_1} = \bigcup_{n=1}^{\infty}\left\{\omega \in \Omega;\ M_n(\omega) < \frac{1}{\ell_1}\right\} \supset \bigcup_{n=1}^{\infty}\left\{\omega \in \Omega;\ M_n(\omega) < \frac{1}{\ell_2}\right\} = B_{\ell_2}$$

なので

$$B_1 \supset B_2 \supset \cdots \supset B_\ell \supset \cdots \qquad \Bigl(\Rightarrow \mathsf{Pr}(B_1) \geq \mathsf{Pr}(B_2) \geq \cdots \geq \mathsf{Pr}(B_\ell) \geq \cdots\Bigr)$$

となる. このことと補題 1.7(7) から

$$\mathsf{Pr}\left(\bigcap_{\ell=1}^{\infty} B_\ell\right) = \lim_{\ell\to\infty}\mathsf{Pr}(B_\ell) \leq \mathsf{Pr}(B_\ell) \leq 1\,(\forall \ell \in \mathbb{N})$$

となる. よって

$$1 = \mathsf{Pr}\left(\bigcap_{\ell=1}^{\infty} B_\ell\right) \Leftrightarrow \mathsf{Pr}(B_\ell) = 1\,(\forall \ell \in \mathbb{N})$$

がわかる. さらに, $n_1 < n_2\,(n_1,\,n_2 \in \mathbb{N})$ に対して

$$M_{n_1}(\omega) := \sup_{k \geq n_1}|X_k(\omega) - X(\omega)| \geq \sup_{k \geq n_2}|X_k(\omega) - X(\omega)|$$

から

$$\left\{\omega \in \Omega;\ M_{n_1}(\omega) < \frac{1}{\ell}\right\} \subset \left\{\omega \in \Omega;\ M_{n_2}(\omega) < \frac{1}{\ell}\right\}$$

がわかる. よって, $\left\{\omega \in \Omega;\ M_n(\omega) < \frac{1}{\ell}\right\}_{n=1}^{\infty}$ は増大列となるで, 補題 1.7(6) から

$$\begin{aligned}\mathsf{Pr}(B_\ell) &= \mathsf{Pr}\left(\bigcup_{n=1}^{\infty}\left\{\omega \in \Omega;\ M_n(\omega) < \frac{1}{\ell}\right\}\right)\\ &= \lim_{n\to\infty}\mathsf{Pr}\left(\left\{\omega \in \Omega;\ M_n(\omega) < \frac{1}{\ell}\right\}\right)\end{aligned}$$

となる. よって

$$\begin{aligned}1 = \mathsf{Pr}\left(\bigcap_{\ell=1}^{\infty} B_\ell\right) &\Leftrightarrow \mathsf{Pr}(B_\ell) = 1\,(\forall \ell \in \mathbb{N})\\ &\Leftrightarrow \lim_{n\to\infty}\mathsf{Pr}\left(\left\{\omega \in \Omega;\ M_n(\omega) < \frac{1}{\ell}\right\}\right) = 1\,(\forall \ell \in \mathbb{N})\\ &\Leftrightarrow \lim_{n\to\infty}\mathsf{Pr}\left(\left\{\omega \in \Omega;\ M_n(\omega) \geq \frac{1}{\ell}\right\}\right) = 0\,(\forall \ell \in \mathbb{N})\end{aligned}$$

がわかる．したがって

$$X_n \xrightarrow{\text{a.s.}} X \;\Leftrightarrow\; M_n \xrightarrow{\text{P}} 0 \tag{4.1}$$

となる．以上の議論から, 概収束は n ステップ後の $|X_k - X|\,(k \geq n)$ が最大 (極大) になるものが 0 に確率収束することと同値である． □

例 4.5. $X_1, X_2, \ldots$ は同じ確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上で定義された独立な確率変数列で

$$X_n \sim \mathsf{Ber}(p_n)\,(n \in \mathbb{N})$$

とする．ただし, $0 < p_n < 1$ とする．すると

$$X_n \xrightarrow{\text{P}} 0 \;\Leftrightarrow\; p_n \xrightarrow{n\to\infty} 0$$

である．実際, 十分小さな任意の $\epsilon > 0$ に対して, $\mathsf{Pr}(X_n \leq -\epsilon) = 0$ と $X_n(\omega) \geq \epsilon \;\Leftrightarrow\; X_n(\omega) = 1\,(\omega \in \Omega)$ であることに注意すると

$$\begin{aligned}\mathsf{Pr}(|X_n| \geq \epsilon) &= \mathsf{Pr}\big(\{X_n \leq -\epsilon\} \cup \{X_n \geq \epsilon\}\big) = \mathsf{Pr}(X_n \geq \epsilon) = \mathsf{Pr}(X_n = 1)\\ &= p_n\end{aligned}$$

がわかる．

次に

$$M_n = \sup_{k \geq n} |X_k - 0|$$

とし, $\pi_n = \mathsf{Pr}(M_n = 1)$ とおく．すると $M_n(\omega) \in \{0, 1\}\,(\omega \in \Omega)$ なので

$$\begin{aligned}1 - \pi_n = \mathsf{Pr}(M_n = 0) &= \mathsf{Pr}\left(\bigcap_{k=n}^{\infty} \{X_k = 0\}\right)\\ &= \prod_{k=n}^{\infty} \mathsf{Pr}(X_k = 0) \qquad (\because \{X_n\}_{n=1}^{\infty} \text{ は独立})\\ &= \prod_{k=n}^{\infty} (1 - p_k)\end{aligned}$$

を得る．すると (4.1) と $X_n \xrightarrow{\text{a.s.}} 0$ から $\pi_n \xrightarrow{n\to\infty} 0$ がわかる．ここで

$$\prod_{k=n}^{\infty} (1 - p_k) \to 1\,(n \to \infty) \Leftrightarrow \sum_{k=n}^{\infty} p_k < \infty\,(n \to \infty)$$

であること[1]に注意すると

$$X_n \xrightarrow{\text{a.s.}} 0 \;\Leftrightarrow\; \sum_{k=n}^{\infty} p_k < \infty\,(n \to \infty)$$

[1] たとえば, [28, pp.91 − 92] を参照のこと．

がわかる.

たとえば, $p_n = \dfrac{1}{n}\,(n \in \mathbb{N})$ とすると

$$X_n \xrightarrow{\mathrm{P}} 0$$

だが

$$\mathsf{Pr}\left(\left\{\omega \in \Omega;\ \lim_{n\to\infty} X_n(\omega) = 0\right\}\right) < 1 \ \Leftrightarrow\ X_n \overset{\text{a.s.}}{\to} 0 \text{ は偽}$$

がわかる. □

注意 4.6. $n = 1, 2, \ldots$ に対して $X_n \sim \mathsf{N}(0, 1/n)$ とする. $\{X_n\}_{n=1}^{\infty}$ は 0 に収束することが期待される.

まず分布収束について確認する. そのために次のような c.d.f. を考える.

$$\mathsf{F}(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

と定義する. すなわち $\mathsf{Pr}(X = 0) = 1$ をみたす確率変数 X の c.d.f. である. $\mathsf{F}(x)$ は $x = 0$ で不連続であることに注意する. したがって $x = 0$ 以外での各点収束を言えばよい.

そのために $\sqrt{n}X_n \sim \mathsf{N}(0, 1)\,(n = 1, 2, \ldots)$ であることに注意する. ただし, Z は標準正規分布 $\mathsf{N}(0, 1)$ に従う確率変数とする. $x < 0$ のとき

$$\begin{aligned}\mathsf{F}_n(x) &= \mathsf{Pr}\big(X_n \leq x\big) = \mathsf{Pr}\big(\sqrt{n}X_n \leq \sqrt{n}x\big) = \mathsf{Pr}(Z \leq \sqrt{n}x) \\ &= \int_{-\infty}^{\sqrt{n}x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,\mathrm{d}z \to 0 \quad (n \to \infty)\end{aligned}$$

となる. 最後の極限は $\sqrt{n}x \overset{n\to\infty}{\longrightarrow} -\infty$ より $\mathsf{F}(x) = 0\,(x < 0)$ がわかる.

つぎに $x > 0$ のとき

$$\begin{aligned}\mathsf{F}_n(x) &= \mathsf{Pr}\big(X_n \leq x\big) = \mathsf{Pr}\big(\sqrt{n}X_n \leq \sqrt{n}x\big) = \mathsf{Pr}(Z \leq \sqrt{n}x) \\ &= \int_{-\infty}^{\sqrt{n}x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,\mathrm{d}z \to 1 \quad (n \to \infty)\end{aligned}$$

となる. 最後の極限は $\sqrt{n}x \overset{n\to\infty}{\longrightarrow} \infty$ よりわかる. よって, $\mathsf{F}(x) = 1\,(x > 0)$ である.

以上の議論から $x \neq 0$ のとき

$$\lim_{n\to\infty} \mathsf{F}_n(x) = \mathsf{F}(x)$$

がわかる.

しかし, $x = 0$ は $\mathsf{F}(x)$ の不連続点なので, $\mathsf{F}_n(0) = 1/2 \neq 1 = F(0)$ であることは問題ない.

最後に, 確率収束を示そう. Chebyshev の不等式 (系 3.3) より, $\forall\epsilon > 0$ に対して

$$\mathsf{Pr}\big(\big|X_n\big| \geq \epsilon\big) = \mathsf{Pr}\big(X_n^2 \geq \epsilon^2\big) \leq \frac{\mathsf{E}\big[X_n^2\big]}{\epsilon^2} = \frac{1}{n\epsilon^2} \xrightarrow{n\to\infty} 0$$

となる. 最後の等号は, $X_n \sim \mathsf{N}(0,\, 1/n)$ なので, $\mathsf{E}[X_n^2] = \dfrac{1}{n}$ となることを用いた. よって $X_n \xrightarrow{\mathrm{P}} 0$ である. □

次に収束のタイプ間の関係について述べる.

定理 4.7. 次に関係が成立する.

(1) $X_n \xrightarrow{\mathrm{q.m.}} X \;\Rightarrow\; X_n \xrightarrow{\mathrm{P}} X$ である.

(2) $X_n \xrightarrow{\mathrm{P}} X \;\Rightarrow\; X_n \rightsquigarrow X$ である.

(3) $X_n \rightsquigarrow X$ かつある定数 c があって $\mathsf{Pr}(X = c) = 1$ のとき $X_n \xrightarrow{\mathrm{P}} c$ である.

Proof. 節 4.**??** で示す. □

注意 4.8. 定理 4.7(1)(2) の逆は一般に成立しない.
(1) **定理** 4.7(1) **の逆の反例**. $U \sim \mathsf{Unif}(0,\, 1)$, $X_n = \sqrt{n}\mathbb{1}_{(0,\,1/n)}(U)\; (n = 1,\, 2,\, \ldots)$, $X = 0$ とおく. このとき $\forall\epsilon > 0$ に対して

$$\begin{aligned}\mathsf{Pr}\big(\big|X_n\big| \geq \epsilon\big) &= \mathsf{Pr}\big(\sqrt{n}\mathbb{1}_{(0,\,1/n)}(U) \geq \epsilon\big) = \mathsf{Pr}\left(0 < U < \frac{1}{n}\right) \\ &= \mathsf{Pr}\left(0 \leq U \leq \frac{1}{n}\right) = \frac{1}{n} \xrightarrow{n\to\infty} 0\end{aligned}$$

となる. よって $X_n \xrightarrow{\mathrm{P}} X$ である. しかし

$$\mathsf{E}\big[X_n^2\big] = n\int_0^{1/n} \mathrm{d}u = 1$$

なので, $X_n \xrightarrow{\mathrm{q.m.}} X$ は成立しない.
(2) **定理** 4.7(2) **の逆の反例**. $X \sim \mathsf{N}(0,\, 1)$ とし, $X_n = -X\; (n = 1,\, 2,\, \ldots)$ とする. したがって, $X_n \sim \mathsf{N}(0,\, 1)$ である. 明らかに

$$\lim_{n\to\infty} \mathsf{F}_n(x) = \mathsf{F}(x)$$

である．よって $X_n \rightsquigarrow X$ となる．しかし $\forall\epsilon > 0$ に対して

$$\mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big) = \mathsf{Pr}(|2X| > \epsilon) = \mathsf{Pr}\left(|X| > \frac{\epsilon}{2}\right) \neq 0$$

である．よって $X_n \xrightarrow{\mathrm{P}} X$ は成立しない． □

定理 4.9 (Portmanteau の補題)**.** 確率変数列 $X_n\,(n = 1, 2, \ldots)$, X に対し, X_n と X の分布関数を F_n と F とそれぞれ書く．このとき, 以下の (1) ～ (7) は同値である．
(1) $X_n \rightsquigarrow X$ である．
(2) $\mathbb{R}$ 上の任意の有界連続関数 g に対して

$$\lim_{n\to\infty} \mathsf{E}[g(X_n)] = \mathsf{E}[g(X)]$$

である．
(3) $\mathbb{R}$ 上の任意の有界 Lipshitz 連続関数 g に対して

$$\lim_{n\to\infty} \mathsf{E}[g(X_n)] = \mathsf{E}[g(X)]$$

である．
(4) $\mathbb{R}$ 上の任意の非負値連続関数 g に対して

$$\liminf_{n\to\infty} \mathsf{E}[g(X_n)] \geq \mathsf{E}[g(X)]$$

である．
(5) $\mathbb{R}$ の任意の開集合 O に対して

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

である．
(6) $\mathbb{R}$ の任意の閉集合 C に対して

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

である．
(7) $\mathbb{R}$ の任意の Borel 集合 B が $\mathsf{Pr}(X \in \partial B) = 0$ ならば

$$\lim_{n\to\infty} \mathsf{Pr}(X_n \in B) = \mathsf{Pr}(X \in B)$$

である．ただし, ∂B は B の境界である．

Proof. 節 4.4.5.4 で示す．ただし, 証明には積分の収束定理の知識が必要である． □

注意 4.10. 定理 4.9 を有限次元の確率ベクトルに拡張することができる. 記号が煩雑にはなるが, 証明は本質的に同じである. □

定理 4.11. X_n, X, Y_n, Y $(n = 1, 2, \ldots)$ は確率変数列とする. $g : \mathbb{R} \to \mathbb{R}$ を連続関数とし, c を定数とする.

(1) $X_n \xrightarrow{\mathrm{P}} X$ かつ $Y_n \xrightarrow{\mathrm{P}} Y \;\Rightarrow\; X_n + Y_n \xrightarrow{\mathrm{P}} X + Y$ である.

(2) $X_n \xrightarrow{\mathrm{q.m.}} X$ かつ $Y_n \xrightarrow{\mathrm{q.m.}} Y \;\Rightarrow\; X_n + Y_n \xrightarrow{\mathrm{q.m.}} X + Y$ である.

(3) $X_n \rightsquigarrow X$ かつ $Y_n \xrightarrow{\mathrm{P}} c \Rightarrow X_n + Y_n \rightsquigarrow X + c$ である.

(4) $X_n \xrightarrow{\mathrm{P}} X$ かつ $Y_n \xrightarrow{\mathrm{P}} Y \;\Rightarrow\; X_n Y_n \xrightarrow{\mathrm{P}} XY$ である.

(5) $X_n \rightsquigarrow X$ かつ $Y_n \xrightarrow{\mathrm{P}} c \;\Rightarrow\; X_n Y_n \rightsquigarrow cX$ である.

(6) $X_n \xrightarrow{\mathrm{P}} X \;\Rightarrow\; g(X_n) \xrightarrow{\mathrm{P}} g(X)$ である.

(7) $X_n \rightsquigarrow X \;\Rightarrow\; g(X_n) \rightsquigarrow g(X)$ である.

Proof. 節 4.4.5.5 で示す. □

定義 4.12. o_P と O_P の記号をここで導入すること (2024/05/29 記).

4.2 大数の法則

定理 4.13. (大数の弱法則) $X_1, X_2, \ldots$ は確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の i.i.d. 確率変数列とする. $\mathsf{E}\big[|X_1|\big] < \infty$ のとき

$$\overline{X}_n := \frac{1}{n}\big(X_1 + X_2 + \cdots + X_n\big) \xrightarrow{\mathrm{P}} \mu, \qquad \mu = \mathsf{E}[X_1]$$

が成立する.

Proof. より強い条件 $\mathsf{E}\big[X_1^2\big] < \infty$ のもとで定理の主張を証明する. 定理の仮定のもとの証明は節 4.5.1 で与える. $\mathsf{Var}[X_1] = \sigma^2$ とおく. Chebyshev の不等式 (系 3.3) より

$$\mathsf{Pr}\big(|\overline{X}_n - \mu| \geq \epsilon\big) \leq \frac{\mathsf{Var}\big[\overline{X}_n\big]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \xrightarrow{n\to\infty} 0$$

よりわかる. □

大数の強法則を述べる前に, 定理の証明に必要な補題を述べる.

補題 4.14. X_1, X_2, ... は非負値確率変数列で, 条件

$$\sum_{n=1}^{\infty} \mathsf{E}[X_n] < \infty \tag{4.2}$$

をみたすとする. このとき

$$\mathsf{Pr}\left(\omega \in \Omega;\ \sum_{n=1}^{\infty} X_n(\omega) < \infty\right) = 1$$

が成り立つ.

Proof. 証明は節 4.5.2 で行う. □

定理 4.15. (大数の強法則) X_1, X_2, ... は確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の i.i.d. 確率変数列とし, $\mathsf{E}\left[|X_1|^4\right] < \infty$ とする. このとき

$$\overline{X}_n := \frac{1}{n}\left(X_1 + X_2 + \cdots + X_n\right) \overset{\text{a.s.}}{\to} \mu, \qquad \mu = \mathsf{E}[X_1] \tag{4.3}$$

が成立する.

Proof. まず

$$\mathsf{E}\left[|X_1|^4\right] =: K < \infty, \quad T_n = X_1 + X_2 + \cdots + X_n, \quad \overline{X}_n = \frac{T_n}{n}$$

とおく.

最初に, $\mu = 0$ として (4.3) を示す. 以下の事象の包含関係に注意する. すなわち

$$\left\{\omega \in \Omega;\ \sum_{n=1}^{\infty} \left(\overline{X}_n(\omega)\right)^4 < \infty\right\} \subset \left\{\omega \in \Omega;\ \lim_{n\to\infty} \overline{X}_n(\omega) = 0\right\}$$

が成立する. このことより

$$\mathsf{Pr}\left(\sum_{n=1}^{\infty} \overline{X}_n^4 < \infty\right) = 1 \tag{4.4}$$

がわかると

$$1 = \mathsf{Pr}\left(\sum_{n=1}^{\infty} \overline{X}_n^4 < \infty\right) \leq \mathsf{Pr}\left(\lim_{n\to\infty} \overline{X}_n = 0\right)$$

となり, $\mu = 0$ のときに (4.3) がわかる.

以下では $\mu=0$ として (4.4) を示す．そのために多項定理を用いて, T_n を展開する．すると

$$\bigl(X_1+\cdots+X_n\bigr)^4=\sum_{\substack{\ell_1+\cdots+\ell_n=4\\ \ell_j\geq 0\,(j=1,\ldots,n)}}\frac{4!}{\ell_1!\times\ell_2!\times\cdots\times\ell_n!}X_1^{\ell_1}\times X_2^{\ell_2}\times\cdots\times X_n^{\ell_n}$$

となる．ただし $\ell_1,\,\ell_2,\,\ldots,\,\ell_n$ は 0 以上の整数である．ここで $X_1^{\ell_1},\,X_2^{\ell_2},\,\ldots,\,X_n^{\ell_n}$ は独立あることに注意すると

$$\mathsf{E}\bigl[T_n^4\bigr]=\sum_{\substack{\ell_1+\ell_2+\cdots+\ell_n=4\\ \ell_j\geq 0\,(j=1,2,\ldots,n)}}\frac{4!}{\ell_1!\ell_2!\times\cdots\times\ell_n!}\mathsf{E}\bigl[X_1^{\ell_1}\bigr]\times\mathsf{E}\bigl[X_2^{\ell_2}\bigr]\times\cdots\times\mathsf{E}\bigl[X_n^{\ell_n}\bigr]$$

がわかる．さらに, $\mathsf{E}\bigl[X_k\bigr]=\mu=0\,(k=1,\,2,\,\ldots,\,n)$ と $\dfrac{4!}{2!2!}=6$ を用いると

$$\mathsf{E}\bigl[T_n^4\bigr]=\sum_{k=1}^{n}\mathsf{E}\bigl[X_k^4\bigr]+6\sum_{1\leq k<\ell\leq n}\mathsf{E}\bigl[X_k^2\bigr]\,\mathsf{E}\bigl[X_\ell^2\bigr]$$

を得る．ここで $\mathsf{E}\bigl[X_k^4\bigr]=K\,(k=1,\,2,\ldots,\,n)$ と Cauchy-Schwarz の不等式から

$$\mathsf{E}\bigl[X_k^2\bigr]\leq\sqrt{\mathsf{E}\bigl[X_k^4\bigr]}=\sqrt{K}$$

となることに注意すると

$$\mathsf{E}\bigl[T_n^4\bigr]\leq nK+6\sum_{1\leq k<\ell\leq n}\sqrt{K}\,\sqrt{K}=nK+3n(n-1)K\leq 3Kn^2$$

を得る．よって

$$\sum_{n=1}^{\infty}\mathsf{E}\bigl[\overline{X}_n^4\bigr]=\sum_{n=1}^{\infty}\frac{1}{n^4}\mathsf{E}\bigl[T_n^4\bigr]\leq 3K\sum_{n=1}^{\infty}\frac{1}{n^2}<\infty$$

となる．したがって 補題 4.14 から (4.4) が成立することがわかる．

つぎに $\mu\neq 0$ の場合を示す．$Y_k=X_k-\mu\,(k=1,\,2,\,\ldots,\,n)$ とおくと

$$\mathsf{E}\bigl[Y_k^4\bigr]\leq\mathsf{E}\bigl[\bigl\{|X_k|+|\mu|\bigr\}^4\bigr]\leq 8\mathsf{E}\bigl[|X_k|^4+|\mu|^4\bigr]\leq 8\bigl(K+\mu^4\bigr)<\infty$$

が成り立つ．したがってこの定理の証明の前半部分で得られた結果から

$$\mathsf{Pr}\left(\lim_{n\to\infty}\frac{1}{n}\bigl(Y_1+Y_2+\cdots+Y_n\bigr)=0\right)=1$$

となる．最後に, $\forall\omega\in\Omega$ に対して

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\bigl(Y_1(\omega)+Y_2(\omega)+\cdots+Y_n(\omega)\bigr)=0\\ &\quad\Leftrightarrow\ \lim_{n\to\infty}\frac{1}{n}\bigl(X_1(\omega)+X_2(\omega)+\cdots+X_n(\omega)\bigr)=\mu\end{aligned}$$

に注意すればよい． □

注意 4.16. 大数の法則は $\mathsf{E}\big[|X_1|\big] < \infty$ で成立する．この条件下での大数の強法則の証明は定理 G.4 で記した． □

系 4.17 (Weierstrass の近似定理)**.** 閉区間 [0, 1] 上の任意の連続関数 f は多項式の極限として表すことができる．特に, $f(0) = f(1) = 0$ のとき[2], $n \in \mathbb{N}$ に対して

$$B_n(x) = \sum_{j=1}^{n} f\left(\frac{j}{n}\right)\left(\begin{array}{c} n \\ j \end{array}\right) x^j (1-x)^{n-j}$$

とおいたとき

$$f(x) = \lim_{n\to\infty} B_n(x) \tag{4.5}$$

が成立する．

Proof. $x = 0,\ 1$ のとき (4.5) は明らかなので, $0 < x < 1$ に対して, (4.5) が成立すること示す．$X_j\ (j = 1, 2, \ldots)$ は i.i.d. 確率変数列で

$$\mathsf{Pr}(X_j = 1) = x = 1 - \mathsf{Pr}(X_j = 0)$$

をみたすとする．いま

$$\mathsf{E}\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = x$$

に注意して, Chebyshev の不等式 (系 3.3) を用いると, 任意の $\epsilon > 0$ に対して

$$\begin{aligned}
\mathsf{Pr}&\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - x\right| \geq \epsilon\right) \\
&\leq \frac{\mathsf{Var}\left(\dfrac{X_1 + X_2 + \cdots + X_n}{n}\right)}{\epsilon^2} \\
&= \frac{\dfrac{1}{n^2}\big(\mathsf{Var}(X_1) + \mathsf{Var}(X_2) + \cdots + \mathsf{Var}(X_n)\big)}{\epsilon^2} \\
&= \frac{x(1-x)}{n\epsilon^2}
\end{aligned}$$

となる．したがって

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{P} x$$

[2] $f(0) = a,\ f(1) = b$ に対して, $\widetilde{f}(x) = f(x) + \big\{f(b) - f(a)\big\}x + f(a)$ とおくと $\widetilde{f}$ は [0, 1] 上の連続関数で $\widetilde{f}(0) = \widetilde{f}(1) = 0$ となるので, 左記の設定は一般性を失わない仮定である．

がわかる. さらに, 定理 4.7(2) より

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \rightsquigarrow x$$

を得る. f は有界連続関数なので, 定理 4.9(2) から

$$\lim_{n\to\infty} \mathsf{E}\left[f\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)\right] = f(x)$$

となる. しかし, $S := X_1 + X_2 + \cdots + X_n \sim \mathsf{Bino}(n,\, x)$ なので

$$\begin{aligned}\mathsf{E}\left[f\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)\right] &= \mathsf{E}\left[f\left(\frac{S}{n}\right)\right]\\ &= \sum_{j=0}^{n} f\left(\frac{j}{n}\right)\begin{pmatrix} n \\ j \end{pmatrix} x^j(1-x)^{n-j}\\ &= B_n(x)\end{aligned}$$

がわかる. よって, 主張は証明された. □

4.3 中心極限定理

定理 4.18. $X_1, X_2, \ldots$ を i.i.d. 確率変数列とし, $\mathsf{E}[X_1] = \mu$, $\mathsf{Var}[X_1] = \sigma^2\,(0 < \sigma < \infty)$ とする. このとき

$$Z_n := \frac{\overline{X}_n - \mu}{\sqrt{\mathsf{Var}\big[\overline{X}_n\big]}} = \frac{\sqrt{n}\big(\overline{X}_n - \mu\big)}{\sigma} \rightsquigarrow Z$$

が成り立つ. ただし $\overline{X}_n = \dfrac{1}{n}\big(X_1 + X_2 + \cdots + X_n\big)$, $Z \sim \mathsf{N}(0,\, 1)$ である.

Proof. 中心極限定理の証明は第 G 章でする. □

例 4.19. $X_1, X_2, \ldots$ を i.i.d. 確率変数列とし, $\mathsf{E}\big[X_1\big] = \mu$, $\mathsf{Var}\big[X_1\big] = \sigma^2\,(0 < \sigma < \infty)$ とする. このとき中心極限定理 (定理 4.18) より

$$Z_n = \frac{\sqrt{n}\big(\overline{X}_n - \mu\big)}{\sigma} \rightsquigarrow \mathsf{N}(0,\, 1) \tag{4.6}$$

が成立する. ここで

$$S_n^2 := \frac{1}{n-1}\sum_{j=1}^{n}\big(X_j - \overline{X}_n\big)^2, \quad S_n = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\big(X_j - \overline{X}_n\big)^2}$$

とおく．すると

$$S_n^2 = \frac{n}{n-1}\left\{\frac{1}{n}\sum_{j=1}^{n}\left(X_j-\mu\right)^2 - \left(\overline{X}_n-\mu\right)^2\right\}$$

と書き直せる．設定から $\mathsf{E}[(X_1-\mu)^2]=\sigma^2<\infty$ なので, 大数の法則 (定理 4.13) より

$$\frac{1}{n}\sum_{j=1}^{n}\left(X_j-\mu\right)^2 \xrightarrow{\mathrm{P}} \mathsf{E}\left[\left(X_1-\mu\right)^2\right]=\sigma^2, \tag{4.7}$$

$$\overline{X}_n \xrightarrow{\mathrm{P}} \mu \tag{4.8}$$

である．(4.8) に対して, 定理 4.11(6)($g(x)=(x-\mu)^2$) を用いると

$$\left(\overline{X}_n-\mu\right)^2 \xrightarrow{\mathrm{P}} 0 \tag{4.9}$$

となる．さらに, (4.7), (4.9) と定理 4.11(3)(5) より

$$S_n^2 \xrightarrow{\mathrm{P}} \sigma^2$$

である．再度, 定理 4.11(6) を用いると

$$\frac{S_n}{\sigma} \xrightarrow{\mathrm{P}} 1 \tag{4.10}$$

がわかる．最後に (4.6), (4.10) と定理 4.11(5) より

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

を得る．

□

定理 4.20. (Berry-Essèen の不等式) $X_1, X_2, \ldots$ を i.i.d. 確率変数列とする．$\mathsf{E}\left[|X_1|^3\right]<\infty$ のとき

$$\sup_z \left|\mathsf{Pr}\left(Z_n\le z\right)-\Phi(z)\right| \le \frac{33}{4}\frac{\mathsf{E}\left[|X_1-\mu|^3\right]}{\sqrt{n}\sigma^2}$$

が成立する．ただし $Z_n = \frac{\sqrt{n}\left(\overline{X}_n-\mu\right)}{\sigma}$, $\Phi(z)=\int_{-\infty}^{z} e^{-t^2/2}\,dt\,(z\in\mathbb{R})$ である．

Proof. 信じることにする． □

注意 4.21. 証明についてのコメントを書くこと． □

定理 4.22. $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ を i.i.d. 確率ベクトル列とする．ただし $j = 1, 2, \ldots$ に対して

$$\boldsymbol{X}_j = \begin{pmatrix} X_{1j} \\ X_{2j} \\ \vdots \\ X_{dj} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{pmatrix} = \begin{pmatrix} \mathsf{E}[X_{1j}] \\ \mathsf{E}[X_{2j}] \\ \vdots \\ \mathsf{E}[X_{dj}] \end{pmatrix},$$

$$\mathsf{Var}[\boldsymbol{X}_1] = \boldsymbol{\Sigma} = \begin{pmatrix} \mathsf{Var}[X_{11}] & \mathsf{Cov}[X_{11}, X_{21}] & \cdots & \mathsf{Cov}[X_{11}, X_{d1}] \\ \mathsf{Cov}[X_{21}, X_{11}] & \mathsf{Var}[X_{21}] & \cdots & \mathsf{Cov}[X_{21}, X_{d1}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Cov}[X_{d1}, X_{11}] & \mathsf{Cov}[X_{d1}, X_{21}] & \cdots & \mathsf{Var}[X_{d1}] \end{pmatrix}$$

は正定値となる．このとき任意の $\boldsymbol{c} \in \mathbb{R}^d$ に対して

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

が成立する．このことを

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

と書くことにする．

Proof. 証明は後で行う． □

4.4 デルタ法

$\{Y_n\}$ を確率変数列とする．Y_n の極限分布が正規分布のとき滑らかな実数値関数 $g : \mathbb{R} \to \mathbb{R}$ に対して $g(Y_n)$ の極限分布を求めよう．

定理 4.23. (デルタ法) $Y_1, Y_2, \ldots$ を確率変数列とし

$$\frac{\sqrt{n}\big(Y_n - \mu\big)}{\sigma} \rightsquigarrow \mathsf{N}(0, 1),$$

とし, g は $x = \mu$ の近傍で連続微分可能な関数で $\dot{g}(\mu) \neq 0$ とする．ただし, $\mu \in \mathbb{R}$, $0 < \sigma < \infty$, $\dot{g}(t) = \dfrac{\mathrm{d}g}{\mathrm{d}t}(t)$ である．このとき

$$\frac{\sqrt{n}\big(g(Y_n) - g(\mu)\big)}{|\dot{g}(\mu)|\sigma} \rightsquigarrow \mathsf{N}(0, 1)$$

が成立する．すなわち

$$Y_n \approx \mathsf{N}\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow g(Y_n) \approx \mathsf{N}\left(g(\mu), \big(\dot{g}(\mu)\big)^2\frac{\sigma^2}{n}\right)$$

である．ただし「≈」は「分布が近似できる」の意味である．

Proof. 節 4.4.5.5 で証明をする. □

例 4.24. $X_1, X_2, \ldots$ を i.i.d. 確率変数列とし, $\mathsf{E}\big[X_1\big] = \mu$, $\mathsf{Var}\big[X_1\big] = \sigma^2\,(0 < \sigma < \infty)$ とする. このとき中心極限定理 (定理 4.18) より

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

となる. いま

$$W_n = e^{\overline{X}_n}$$

とおく. したがって $g(x) = e^x\,(x \in \mathbb{R})$ とすれば $\overset{\bullet}{g}(x) = e^x$ となる. よってデルタ法 (定理 4.23) より

$$\frac{\sqrt{n}\big(W_n - e^{\mu}\big)}{e^{\mu}\sigma} \rightsquigarrow \mathsf{N}(0,\, 1)$$

となる. よって

$$W_n \approx \mathsf{N}\left(e^{\mu},\, \frac{e^{2\mu}\sigma^2}{n}\right)$$

がわかる. □

定理 4.25. (多次元デルタ法) $\boldsymbol{Y}_n = (Y_{1n}, Y_{2n}, \ldots, Y_{dn})^{\top}\,(n = 1, 2, \ldots)$ を確率ベクトル列とし

$$\sqrt{n}\big(\boldsymbol{Y}_n - \boldsymbol{\mu}\big) \rightsquigarrow \mathsf{N}_d\big(\mathbf{0}_d,\, \boldsymbol{\Sigma}\big)$$

とする. ただし $\boldsymbol{\mu} \in \mathbb{R}^d$ で $\boldsymbol{\Sigma}$ は $d \times d$ の正値対称行列とする. 関数 $g : \mathbb{R}^d \to \mathbb{R}$ は滑らかで

$$\nabla g(\boldsymbol{y}) = \begin{pmatrix} \frac{\partial g}{\partial y_1} \\ \frac{\partial g}{\partial y_2} \\ \vdots \\ \frac{\partial g}{\partial y_d} \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{pmatrix}, \quad \nabla_{\boldsymbol{\mu}} := \nabla g(\boldsymbol{y})\Big|_{\boldsymbol{y}=\boldsymbol{\mu}}$$

とする. このとき

$$\sqrt{n}\big(g\big(\boldsymbol{Y}_n\big) - g\big(\boldsymbol{\mu}\big)\big) \rightsquigarrow \mathsf{N}\big(0,\, \nabla_{\boldsymbol{\mu}}^{\top}\boldsymbol{\Sigma}\nabla_{\boldsymbol{\mu}}\big)$$

となる.

例 4.26.

$$\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}, \begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix}, \ldots, \begin{pmatrix} X_{1n} \\ X_{2n} \end{pmatrix}$$

は i.i.d. 確率ベクトル列で

$$\mathsf{E}\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \boldsymbol{\mu}, \quad \mathsf{Var}\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} = \boldsymbol{\Sigma}$$

とする．ただし $\boldsymbol{\Sigma}$ は 2×2 の正値対称行列である．

$$\overline{X}_1 = \frac{1}{n}\sum_{j=1}^{n} X_{1j}, \quad \overline{X}_2 = \frac{1}{n}\sum_{j=1}^{n} X_{2j}$$

とし

$$Y_n = \overline{X}_1\overline{X}_2, \quad g(s_1,\, s_2) = s_1 s_2 \,(s_1,\, s_2 \in \mathbb{R})$$

とおく．中心極限定理 (定理 4.22) より

$$\sqrt{n}\begin{pmatrix} \overline{X}_1 - \mu_1 \\ \overline{X}_2 - \mu_2 \end{pmatrix} \rightsquigarrow \mathsf{N}_2\bigl(\mathbf{0}_2,\, \boldsymbol{\Sigma}\bigr)$$

となる．いま

$$\nabla g(\boldsymbol{s}) = \begin{pmatrix} \dfrac{\partial g}{\partial s_1} \\ \dfrac{\partial g}{\partial s_2} \end{pmatrix} = \begin{pmatrix} s_2 \\ s_1 \end{pmatrix}, \quad \boldsymbol{s} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}$$

であるので

$$\nabla_{\boldsymbol{\mu}}^{\top}\boldsymbol{\Sigma}\nabla_{\boldsymbol{\mu}} = \mu_2^2\sigma_{11} + 2\mu_1\mu_2\sigma_{12} + \mu_1^2\sigma_{22}$$

である．したがって

$$\sqrt{n}\bigl(\overline{X}_1\overline{X}_2 - \mu_1\mu_2\bigr) \rightsquigarrow \mathsf{N}\bigl(0,\, \mu_2^2\sigma_{11} + 2\mu_1\mu_2\sigma_{12} + \mu_1^2\sigma_{22}\bigr)$$

がわかる． □

4.5 定理 4.13, 補題 4.14, 定理 4.7, 4.9, 4.11, 4.23 の証明

4.5.1 定理 4.13 の証明

任意の $\delta > 0$ と n に対して

$$Y_j := Y_j(n) = \begin{cases} X_j & (|X_j| \le \delta n) \\ 0 & (|X_j| > \delta n) \end{cases}$$

$$Z_j := Z_j(n) = \begin{cases} 0 & (|X_j| \le \delta n) \\ X_j & (|X_j| > \delta n) \end{cases} \quad ; \quad (j = 1,\, 2,\, \ldots,\, n)$$

と定める. 明らかに

$$X_j = Y_j + Z_j, \qquad (j = 1, 2, \ldots, n)$$

である. 以下では, $X_j\,(j = 1, 2, \ldots, n)$ は連続型確率変数とし, 共通の p.d.f. p を持つとして議論を進めていく. このとき

$$\begin{aligned}
\mathsf{Var}[Y_j] &= \mathsf{Var}[Y_1] = \mathsf{E}[Y_1^2] - \bigl(\mathsf{E}[Y_1]\bigr)^2 \le \mathsf{E}[Y_1^2] \\
&= \mathsf{E}\bigl[X_1^2 \mathbb{1}_{[-\delta n,\, \delta n]}(X_1)\bigr] \\
&= \int_{-\infty}^{\infty} x^2 \mathbb{1}_{[-\delta n,\, \delta n]}(x)\mathsf{p}(x)\,\mathrm{d}x = \int_{-\delta n}^{\delta n} x^2 \mathsf{p}(x)\,\mathrm{d}x \\
&\le \delta n \int_{-\delta n}^{\delta n} |x|\mathsf{p}(x)\,\mathrm{d}x \le \delta n \int_{-\infty}^{\infty} |x|\mathsf{p}(x)\,\mathrm{d}x \\
&= \delta n \mathsf{E}[X_1]
\end{aligned}$$

となる. すなわち

$$\mathsf{Var}[Y_j] \le \delta n \mathsf{E}[|X_1|] \qquad (j = 1, 2, \ldots, n) \tag{4.11}$$

である. 次に

$$\mathsf{E}[Y_j] = \mathsf{E}[Y_1] = \mathsf{E}[X_1 \mathbb{1}_{[-\delta n,\, \delta n]}(X_1)] = \int_{-\infty}^{\infty} x \mathbb{1}_{[-\delta n,\, \delta n]}(x)\mathsf{p}(x)\,\mathrm{d}x$$

と書ける. ここで

$$\left| x \mathbb{1}_{[-\delta n,\, \delta n]}(x)\mathsf{p}(x) \right| \le |x|\mathsf{p}(x), \qquad x \mathbb{1}_{[-\delta n,\, \delta n]}(x)\mathsf{p}(x) \overset{n\to\infty}{\longrightarrow} x\mathsf{p}(x)$$

かつ

$$\int_{-\infty}^{\infty} |x|\mathsf{p}(x)\,\mathrm{d}x < \infty$$

であることに注意して, 優収束定理 (定理 C.51) を用いると

$$\mathsf{E}[Y_1] = \int_{-\infty}^{\infty} x \mathbb{1}_{[-\delta n,\, \delta n]}(x)\mathsf{p}(x)\,\mathrm{d}x \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} x\mathsf{p}(x)\,\mathrm{d}x$$

となることがわかる. したがって

$$\mathsf{E}[Y_1] \overset{n\to\infty}{\longrightarrow} \mu \tag{4.12}$$

である．以上のことを踏まえると任意の $\epsilon > 0$ に対して

$$
\begin{aligned}
\mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mathsf{E}[Y_1]\right| \geq \epsilon\right) &= \mathsf{Pr}\left(\left|\sum_{j=1}^{n} Y_j - \mathsf{E}\left[\sum_{j=1}^{n} Y_j\right]\right| \geq n\epsilon\right) \\
&\leq \frac{1}{n^2\epsilon^2}\mathsf{Var}\left[\sum_{j=1}^{n} Y_j\right] \qquad (\because 系\ 3.3) \\
&= \frac{n}{n^2\epsilon^2}\mathsf{Var}[Y_1] \leq \frac{n\delta n\mathsf{E}[|X_1|]}{n^2\epsilon^2} \qquad (\because (4.11)) \\
&= \frac{\delta}{\epsilon^2}\mathsf{E}[|X_1|]
\end{aligned}
$$

となる．したがって

$$
\mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mathsf{E}[Y_1]\right| \geq \epsilon\right) \leq \frac{\delta}{\epsilon^2}\mathsf{E}[|X_1|] \tag{4.13}
$$

となる．よって, (4.12) と (4.13) から, 十分大きな n に対して

$$
\begin{aligned}
\mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mu\right| \geq 2\epsilon\right) &= \mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mathsf{E}[Y_1] + \mathsf{E}[Y_1] - \mu\right| \geq 2\epsilon\right) \\
&\leq \mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mathsf{E}[Y_1]\right| + \left|\mathsf{E}[Y_1] - \mu\right| \geq 2\epsilon\right) \\
&\leq \mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mathsf{E}[Y_1]\right| \geq \epsilon\right) + \mathsf{Pr}\left(\left|\mathsf{E}[Y_1] - \mu\right| \geq \epsilon\right) \\
&\leq \frac{\delta}{\epsilon^2}\mathsf{E}[|X_1|]
\end{aligned}
$$

となる．したがって

$$
\mathsf{Pr}\left(\left|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mu\right| \geq 2\epsilon\right) \leq \frac{\delta}{\epsilon^2}\mathsf{E}[|X_1|] \tag{4.14}
$$

となる．さらに, $\mathsf{E}[|X_1] < \infty$ なので, 任意の $\delta > 0$ と十分大きな n に対して

$$
\int_{|x|>\delta n} |x|\mathsf{p}(x)\,\mathrm{d}x < \delta^2
$$

とできることに注意する．このことより, $j = 1, 2, \ldots, n$ に対して

$$
\begin{aligned}
\mathsf{Pr}(Z_j \neq 0) = \mathsf{Pr}(|Z_j| > \delta n) = \mathsf{Pr}(|X_j| > \delta n) &= \int_{-\infty}^{-\delta n} \mathsf{p}(x)\,\mathrm{d}x + \int_{\delta n}^{\infty} \mathsf{p}(x)\,\mathrm{d}x \\
&= \int_{|x|>\delta n} \mathsf{p}(x)\,\mathrm{d}x < \int_{|x|>\delta n} \frac{|x|}{\delta n}\mathsf{p}(x)\,\mathrm{d}x = \frac{1}{\delta n}\int_{|x|>\delta n} |x|\mathsf{p}(x)\,\mathrm{d}x \\
&< \frac{\delta^2}{\delta n} = \frac{\delta}{n}
\end{aligned} \tag{4.15}
$$

となる．したがって, (4.15) から

$$\mathsf{Pr}\Bigl(\sum_{j=1}^{n} Z_j \neq 0\Bigr) = \mathsf{Pr}\Bigl(\bigcup_{j=1}^{n}\{Z_j \neq 0\}\Bigr) = n\mathsf{Pr}\bigl(Z_1 \neq 0\bigr) \leq \delta \tag{4.16}$$

を得る．よって, (4.14) と (4.16) から, 十分大きな n に対して

$$\begin{aligned}
\mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} X_j - \mu\Biggr| \geq 4\epsilon\Biggr) &= \mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} Y_j + \frac{1}{n}\sum_{j=1}^{n} Z_j - \mu\Biggr| \geq 4\epsilon\Biggr) \\
&\leq \mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mu\Biggr| + \Biggl|\frac{1}{n}\sum_{j=1}^{n} Z_j\Biggr| \geq 4\epsilon\Biggr) \\
&\leq \mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mu\Biggr| \geq 2\epsilon\Biggr) + \mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} Z_j\Biggr| \geq 2\epsilon\Biggr) \\
&\leq \mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} Y_j - \mu\Biggr| \geq 2\epsilon\Biggr) + \mathsf{Pr}\Biggl(\sum_{j=1}^{n} Z_j \neq 0\Biggr) \\
&\leq \frac{\delta}{\epsilon^2}\mathsf{E}[|X_1|] + \delta
\end{aligned}$$

となる．ここで, δ を ϵ^3 とおきなおせば, 十分大きな n に対して

$$\mathsf{Pr}\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^{n} X_j - \mu\Biggr| \geq 4\epsilon\Biggr) \leq \epsilon\mathsf{E}[|X_1|] + \epsilon^3$$

となる．上式は任意の $\epsilon > 0$ で成立したので, 定理の主張は証明された．
□

4.5.2　補題 4.14 の証明

背理法で証明する．そのために

$$\mathsf{Pr}\Biggl(\sum_{k=1}^{\infty} X_k < \infty\Biggr) < 1$$

を仮定する．さらに事象 F と N を

$$F := \Biggl\{\omega \in \Omega;\ \sum_{k=1}^{\infty} X_k(\omega) < \infty\Biggr\}, \qquad N := \Biggl\{\omega \in \Omega;\ \sum_{k=1}^{\infty} X_k(\omega) = \infty\Biggr\}$$

で定める．このとき

$$\Omega = F \cup N \quad \text{かつ} \quad F \cap N = \varnothing$$

が成り立つ. したがって

$$1 = \mathsf{Pr}(\Omega) = \mathsf{Pr}\big(F \cup N\big) = \mathsf{Pr}(F) + \mathsf{Pr}(N) \qquad (\because \mathsf{Pr}\ \text{の加法性})$$

がわかる. この関係式と背理法の仮定 $\mathsf{Pr}(F) < 1$ から $\mathsf{Pr}(N) > 0$ となる. 一方, $\omega \in N$ のとき

$$\sum_{k=1}^{\infty} X_k(\omega) = \infty$$

であり, $\omega \in N^c$ のとき

$$0 \leq \sum_{k=1}^{\infty} X_k(\omega) < \infty$$

である. したがって, 任意の正の実数 $r \geq 0$ に対して

$$r\mathbb{1}_N(\omega) \leq \sum_{k=1}^{\infty} X_k(\omega) \qquad (\forall \omega \in \Omega) \tag{4.17}$$

が成り立つ. (4.17) の両辺の期待値をとる. すると 定義 2.4(4), (4.23) と (4.17) から

$$r \times \mathsf{Pr}(N) = \mathsf{E}[r\mathbb{1}_N] \leq \mathsf{E}\left[\sum_{k=1}^{\infty} X_k\right] = \sum_{k=1}^{\infty} \mathsf{E}\big[X_k\big] < \infty \tag{4.18}$$

を得る. (4.18) の最後の等号は単調収束定理からわかる. ここで $\mathsf{Pr}(N) > 0$ であり, $r \geq 0$ は任意の実数だったので, $r \to \infty$ とすれば, (4.18) の最左辺は $+\infty$ となるので, 矛盾が生じる. したがって, $\mathsf{Pr}(F) = 1$ が成り立つ. □

4.5.3 定理 4.7 の証明

(1) $X_n \overset{qm}{\to} X$ とする. $\epsilon > 0$ を固定する. このとき Markov の不等式 (定理 3.1) より

$$\mathsf{Pr}\big(\big|X_n - X\big| \geq \epsilon\big) = \mathsf{Pr}\big(\big|X_n - X\big|^2 \geq \epsilon^2\big) \leq \frac{\mathsf{E}\big[\big|X_n - X\big|^2\big]}{\epsilon^2} \overset{n\to\infty}{\longrightarrow} 0$$

がわかる.

(2) $X_n \xrightarrow{\mathrm{P}} X$ とする．$\epsilon > 0$ を固定し, x と $x \pm \epsilon$ を X の c.d.f. $\mathsf{F}(x)$ の連続点[3]とする．このとき

$$
\begin{aligned}
\mathsf{F}_n(x) &= \mathsf{Pr}\big(X_n \le x\big) \\
&= \mathsf{Pr}\big(X_n \le x,\, X < x+\epsilon\big) + \mathsf{Pr}\big(X_n \le x,\, X \ge x+\epsilon\big) \\
&\le \mathsf{Pr}\big(X \le x+\epsilon\big) + \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big) \qquad (4.19) \\
&= \mathsf{F}(x+\epsilon) + \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big) \qquad (4.20)
\end{aligned}
$$

となる．同様に

$$
\begin{aligned}
\mathsf{F}(x-\epsilon) &= \mathsf{Pr}\big(X \le x-\epsilon\big) \\
&= \mathsf{Pr}\big(X_n \le x-\epsilon,\, X_n < x\big) + \mathsf{Pr}\big(X \le x-\epsilon,\, X_n \ge x\big) \\
&\le \mathsf{Pr}\big(X_n \le x\big) + \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big) \qquad (4.21) \\
&= \mathsf{F}_n(x) + \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big) \qquad (4.22)
\end{aligned}
$$

となる．(4.20) と (4.22) を合わせると

$$
\mathsf{F}(x-\epsilon) - \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big) \le \mathsf{F}_n(x) \le \mathsf{F}(x+\epsilon) + \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big)
$$

がわかる．上の式の辺々で $n \to \infty$ とすれば

$$
\mathsf{F}(x-\epsilon) \le \liminf_{n\to\infty} \mathsf{F}_n(x) \le \limsup_{n\to\infty} \mathsf{F}_n(x) \le \mathsf{F}(x+\epsilon)
$$

となる．ここで $\epsilon \to 0$ とし x を $\mathsf{F}(x)$ の連続点とすれば

$$
\lim_{n\to\infty} \mathsf{F}_n(x) = \mathsf{F}(x)
$$

がわかる．

(3) $X_n \rightsquigarrow X$ かつある定数 c があって $\mathsf{Pr}(X = c) = 1$ とする．$c \pm \epsilon$ を $\mathsf{F}(x)$ の連続点になるようにして $\epsilon > 0$ を固定する[4]．このとき

$$
\begin{aligned}
\mathsf{Pr}\big(\big|X_n - c\big| \ge \epsilon\big) &= \mathsf{Pr}\big(X_n \le c-\epsilon\big) + \mathsf{Pr}\big(X_n \ge c+\epsilon\big) \\
&= \mathsf{Pr}\big(X_n \le c-\epsilon\big) + \mathsf{Pr}\big(X_n \ge c+\epsilon\big) \\
&= \mathsf{F}_n(c-\epsilon) + 1 - \mathsf{F}_n(c+\epsilon) \\
&\to \mathsf{F}(c-\epsilon) + 1 - \mathsf{F}(c+\epsilon) \quad (n \to \infty) \\
&= 0 + 1 - 1 = 0
\end{aligned}
$$

となる． □

[3] X の分布関数の不連続点は高々加算個なので, このように点を取れることに注意せよ．

[4] $\mathsf{F}(x)$ は有界非減少関数なので, $\mathsf{F}(x)$ の不連続点は高々可算個である．よってこのように ϵ を取ることとができる．

問 4.1. (4.19) と (4.21) を確認せよ．すなわち

$$\mathsf{Pr}\big(X_n \le x,\, X \ge x+\epsilon\big) \le \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big),$$
$$\mathsf{Pr}\big(X \le x-\epsilon,\, X_n \ge x\big) \le \mathsf{Pr}\big(\big|X_n - X\big| \ge \epsilon\big)$$

を示せ．

4.5.4 定理 4.9 の証明

(1) $\Rightarrow$ (2) の証明: g を任意の有界連続関数とする．一般性を失わずに $|g(x)| \le 1/2\,(x \in \mathbb{R})$ と仮定してよい．任意の $\epsilon > 0$ に対して, 十分大きな有界閉区間 K をとると

$$\mathsf{Pr}(X \in K^c) \le \epsilon \tag{4.23}$$

とできる．そこで, この K を $m\,(m \in \mathbb{N})$ 個の互いに素な区間 $K_1 = [a_1,\, b_1],\, K_j = (a_j,\, b_j]\,(j = 2,\, 2,\, \ldots,\, m); a_j,\, b_j \in \mathbb{R})$ に直和分解する．すなわち, $K = \bigcup_{j=1}^m K_j$ となっている．このとき, m を十分大きく取ると

$$\max_{j\in\{1,2,\ldots,m\}} \max_{x,\,y\in K_j} |g(x) - g(y)| \le \epsilon \tag{4.24}$$

となるように $\{K_j\}_{j=1}^m$ をとることができる．さらに, $\{a_j,\, b_j\}_{j=1}^m$ は X の分布関数 F の連続点[5]に取る．このとき, (1) より

$$\mathsf{Pr}(X_n \in K) \overset{n\to\infty}{\longrightarrow} \mathsf{Pr}(X \in K) \tag{4.25}$$

となる．つぎに, 各 K_j から 1 点 x_j を任意に選び, 関数 g_ϵ を

$$g_\epsilon(x) = \sum_{j=1}^m g(x_j)\mathbb{1}_{K_j}(x) \qquad (x \in \mathbb{R})$$

と定める．すると

$$\max_{x\in K}\big|g(x) - g_\epsilon(x)\big| \le \epsilon \tag{4.26}$$

[5] F の不連続点は高々可算個なので, このように取れることがわかる．

となる．したがって，(4.23) と (4.26) に注意すると

$$
\begin{aligned}
&\left|\mathsf{E}[g(X)] - \mathsf{E}[g_\epsilon(X)]\right| \\
&\quad = \left|\mathsf{E}[\mathbb{1}_K(X)g(X)] + \mathsf{E}[\mathbb{1}_{K^c}(X)g(X)] - \mathsf{E}[\mathbb{1}_K(X)g_\epsilon(X)] \right.\\
&\qquad \left. - \mathsf{E}[\mathbb{1}_{K^c}(X)g_\epsilon(X)]\right| \\
&\quad \le \left|\mathsf{E}[\mathbb{1}_K(X)g(X)] - \mathsf{E}[\mathbb{1}_K(X)g_\epsilon(X)]\right| + \left|\mathsf{E}[\mathbb{1}_{K^c}(X)g(X)] \right.\\
&\qquad \left. - \mathsf{E}[\mathbb{1}_{K^c}(X)g_\epsilon(X)]\right| \\
&\quad \le \mathsf{E}\big[\mathbb{1}_K(X)\big|g(X) - g_\epsilon(X)\big|\big] + \mathsf{E}\big[\mathbb{1}_{K^c}(X)\big|g(X) - g_\epsilon(X)\big|\big] \\
&\quad \le \underbrace{\max_{x\in K}|g(x) - g_\epsilon(x)|}_{\le\epsilon \quad \because (4.26)} \mathsf{E}\big[\mathbb{1}_K(X)\big] + \mathsf{E}\big[\mathbb{1}_{K^c}(X)\big|g(X) - g_\epsilon(X)\big|\big] \\
&\quad \le \epsilon \mathsf{Pr}(X\in K) + \mathsf{Pr}(X\in K^c) \\
&\quad \le 2\epsilon
\end{aligned}
\tag{4.27}
$$

となる．(4.25) から，n を十分大きく取ると

$$
\Big|\underbrace{\mathsf{Pr}(X_n\in K^c)}_{=1-\mathsf{Pr}(X_n\in K)} - \underbrace{\mathsf{Pr}(X\in K^c)}_{=1-\mathsf{Pr}(X\in K)}\Big| = \left|\mathsf{Pr}(X\in K) - \mathsf{Pr}(X\in K)\right| \le \epsilon
$$

とできるので

$$
\mathsf{Pr}(X_n\in K^c) \le \mathsf{Pr}(X\in K^c) + \left|\mathsf{Pr}(X_n\in K^c) - \mathsf{Pr}(X\in K^c)\right| \le 2\epsilon
$$

がわかる．同様に

$$
\left|\mathsf{E}[g(X_n)] - \mathsf{E}[g_\epsilon(X_n)]\right| \le \epsilon\mathsf{Pr}(X_n\in K) + \mathsf{Pr}(X_n\in K^c) \le 3\epsilon \tag{4.28}
$$

を得る．g_ϵ の作りかたから n を十分大きくとると

$$
\begin{aligned}
&\left|\mathsf{E}[g_\epsilon(X)] - \mathsf{E}[g_\epsilon(X_n)]\right| \\
&\quad \le \sum_{j=1}^{n}\left|\mathsf{Pr}(X\in K_j) - \mathsf{Pr}(X_n\in K_j)\right| \max_{j=\{1,2,\dots,m\}}|g(x_j)| \\
&\quad \le \frac{\epsilon}{2} < \epsilon
\end{aligned}
\tag{4.29}
$$

となる．よって，(4.27) – (4.29) を合わせると

$$
\begin{aligned}
\left|\mathsf{E}[g(X_n)] - \mathsf{E}[g(X)]\right| &\le \Big|\underbrace{\mathsf{E}[g(X_n)] - \mathsf{E}[g_\epsilon(X_n)]}_{\le 3\epsilon}\Big| + \underbrace{\left|\mathsf{E}[g_\epsilon(X_n)] - \mathsf{E}[g_\epsilon(X)]\right|}_{\le\epsilon} \\
&\quad + \underbrace{\left|\mathsf{E}[g_\epsilon(X)] - \mathsf{E}[g(X)]\right|}_{\le 2\epsilon} \\
&\le 6\epsilon
\end{aligned}
$$

がわかるので, 主張は証明された.
(2) $\Rightarrow$ (3) の証明: 明らか.
(2) $\Rightarrow$ (4) の証明: 非負値連続関数 g と $M>0$ に対して, $g_M(x)=\min\{g(x),M\}\geq 0\,(x\in\mathbb{R})$ と定める g_M は有界連続となる. よって

$$\liminf_{n\to\infty}\mathsf{E}[g(X_n)]\geq\liminf_{n\to\infty}\mathsf{E}[g_M(X_n)]=\mathsf{E}[g_M(X)]\qquad\big(\because (2)\big)$$

となる. 両辺で $M\to\infty$ とすることにより, 有界単調収束定理から (4) を得る.
(4) $\Rightarrow$ (2) の証明: g を有界連続関数とする. するとある $M>0$ が存在して, $|g|\leq M$ とできる. すると $M\pm g\geq 0$ は非負値連続関数となるので, (4) を用いると (2) を得る.
(3) $\Rightarrow$ (5) の証明: 開集合 O と $x\in\mathbb{R}$ に対して $g_M(x)\uparrow \mathbb{1}_O(x)\,(M\to\infty)$ となるような Lipschitz 連続関数列 $g_M\geq 0$ を取ることができるので, (3) から

$$\liminf_{n\to\infty}\mathsf{Pr}(X_n\in O)\geq\liminf_{n\to\infty}\mathsf{E}[g_M(X_n)]=\mathsf{E}[g_M(X)]$$

となる. この両辺で $M\to\infty$ とすると, 単調収束定理により (5) を得る.
(5) $\Leftrightarrow$ (6) の証明: お互いの補集合をとればよい.
(5) + (6) $\Rightarrow$ (7) の証明: B° を B の内部, $\mathrm{cl}(B)$ を B の閉包とすると $\partial B=\mathrm{cl}(B)\setminus B^\circ$ である. このことに注意すると

$$\begin{aligned}\mathsf{Pr}\big(X\in B^\circ\big)&\leq\liminf_{n\to\infty}\mathsf{Pr}\big(X_n\in B^\circ\big)\leq\limsup_{n\to\infty}\mathsf{Pr}\big(X_n\in\mathrm{cl}(B)\big)\\&\leq\mathsf{Pr}\big(X_n\in\mathrm{cl}(B)\big)\end{aligned}$$

となる. $\mathsf{Pr}(X\in\partial B)=0$ より, 上式の右辺と左辺は等しいので, (7) が得られる.
(7) $\Rightarrow$ (1) の証明: $B=(-\infty,x]$ ととればよい. □

4.5.5 定理 4.11 の証明

(6) の証明: 関数 g は $x_0\,(\in\mathbb{R})$ で連続なので, 任意の $\epsilon>0$ に対して, $\exists\delta>0$ があって

$$\forall x\in\mathbb{R}:\ |x-x_0|<\delta\ \Rightarrow\ |g(x)-g(x_0)|<\epsilon$$

である. よって

$$\{\omega\in\Omega:\ |X(\omega)-X_n(\omega)|<\delta\}\subset\{\omega\in\Omega:\ |g\big(X(\omega)\big)-g\big(X_n(\omega)\big)|<\epsilon\}$$

となる. 上式の補事象を取り, 補題 1.7(4) の後半の主張を適用すると

$$\mathsf{Pr}\big(|X - X_n| \geq \delta\big) \geq \mathsf{Pr}\big(|g(X) - g(X_n)| \geq \epsilon\big)$$

がわかる. よって, $g(X_n) \xrightarrow{P} g(X)$ が示せた.
(7) の証明: $g \circ f$ が有界連続関数となるように有界連続関数 f を取る. $X_n \rightsquigarrow X$ と定理 4.9 から

$$\lim_{n\to\infty} \mathsf{E}[f\big(g(X_n)\big)] = \lim_{n\to\infty} \mathsf{E}[(f \circ g)(X_n)] = \mathsf{E}[(f \circ g)(X)] = \mathsf{E}[f\big(g(X)\big)]$$

から (7) は示される.
(1) の証明: $X_n \xrightarrow{P} X$ かつ $Y_n \xrightarrow{P} Y$ から $(X_n, Y_n) \xrightarrow{P} (X, Y)$ がわかるので, (6) の関数を $g(x, y) = x + y$ ととればよい.
(2) の証明: 実数 $x, y, a, b \in \mathbb{R}$ に対して, 不等式 $\{(x+y) - (a+b)\}^2 \leq 2\{(x-a)^2 + (y-b)\}$ を使えばよい.
(3) の証明: $Y_n \xrightarrow{P} c$ なので, 注意 4.2(2) から $Y_n \rightsquigarrow c$ となる. このことから $(X_n, Y_n) \rightsquigarrow (X, Y)$ となることが証明[6]できる. あとは (7) において, $g(x, y) = x + y$ とすればよい.
(4) の証明: (6) を用いればよい.
(5) の証明: (7) を用いればよい. □

4.5.6 定理 4.23 の証明

$g(x)$ は $x = \mu$ の近傍で連続微分可能である. このことからある $\delta_1 > 0$ が存在して $|x - \mu| < \delta_1$ なる任意の x に対して

$$g(x) - g(\mu) = (x - \mu) \int_0^1 \overset{\bullet}{g}\ \big(\mu + t(x - \mu)\big)\, \mathrm{d}t \tag{4.30}$$

が成立する. また $\overset{\bullet}{g}\ (x)$ は $x = \mu$ で連続なので, 任意の $\epsilon > 0$ に対してある $\delta_2 > 0$ があって

$$|x - \mu| < \delta_2 \ \Rightarrow\ \big|\overset{\bullet}{g}\ (x) - \overset{\bullet}{g}\ (\mu)\big| < \epsilon$$

となる. よって $|x - \mu| < \min(\delta_1, \delta_2) =: \delta$ なる任意の x に対して

$$\begin{aligned}\Big|\int_0^1 \overset{\bullet}{g}\ \big(\mu + t(x-\mu)\big)\, dt - \overset{\bullet}{g}\ (\mu)\Big| &\leq \int_0^1 \big|\overset{\bullet}{g}\ \big(\mu + t(x-\mu)\big) - \overset{\bullet}{g}\ (\mu)\big|\, dt \\ &< \epsilon \int_0^1 dt = \epsilon \end{aligned} \tag{4.31}$$

[6] これは特性関数の収束と分布収束が同値である事実からわかる. あとは特性関数に対して三角不等式を用いればよい. 詳しくは [33, p.135] を参照のこと.

となる. (4.31) より

$$\mathsf{Pr}\left(\left|\int_0^1 \overset{\bullet}{g}\left(\mu+t(Y_n-\mu)\right)dt-\overset{\bullet}{g}(\mu)\right|\geq\epsilon\right)\leq\mathsf{Pr}\left(\left|Y_n-\mu\right|\geq\delta\right) \quad (4.32)$$

が成立する. $\sqrt{n}\big(Y_n-\mu\big)/\sigma \rightsquigarrow \mathsf{N}(0,\,1)$ なので $Y_n\xrightarrow{\mathrm{P}}\mu$ である. (4.32) から

$$\int_0^1 \overset{\bullet}{g}\left(\mu+t(Y_n-\mu)\right)\mathrm{d}t\xrightarrow{\mathrm{P}}\overset{\bullet}{g}(\mu) \quad (4.33)$$

が成立する. $|Y_n-\mu|<\delta$ が起こったとき, (4.30) に $x=Y_n$ を代入すれば

$$\frac{\sqrt{n}\big(g(Y_n)-g(\mu)\big)}{\sigma}=\frac{\sqrt{n}\big(Y_n-\mu\big)}{\sigma}\int_0^1 \overset{\bullet}{g}\left(\mu+t(Y_n-\mu)\right)\mathrm{d}t \quad (4.34)$$

を得る. さらに, $\sqrt{n}(Y_n-\mu)/\sigma \rightsquigarrow \mathsf{N}(0,\,1)$ と (4.33) に注意して定理 4.11(5) を (4.34) に適用すれば

$$\begin{aligned}
&\frac{\sqrt{n}\big(g(Y_n)-g(\mu)\big)}{\sigma}\\
&=\frac{\sqrt{n}\big(g(Y_n)-g(\mu)\big)}{\sigma}\mathbb{1}\{|Y_n-\mu|<\delta\}+\frac{\sqrt{n}\big(g(Y_n)-g(\mu)\big)}{\sigma}\mathbb{1}\{|Y_n-\mu|\geq\delta\}\\
&=\frac{\sqrt{n}(Y_n-\mu)}{\sigma}\underbrace{\int_0^1 \overset{\bullet}{g}\left(\mu+t(Y_n-\mu)\right)\mathrm{d}t\mathbb{1}\{|Y_n-\mu|<\delta\}}_{\overset{\bullet}{g}(\mu)+o_P(1)}\\
&\quad+\underbrace{\frac{\sqrt{n}\big(g(Y_n)-g(\mu)\big)}{\sigma}\mathbb{1}\{|Y_n-\mu|\geq\delta\}}_{=o_P(1)}\\
&\rightsquigarrow\overset{\bullet}{g}(\mu)\mathsf{N}(0,\,1)
\end{aligned}$$

を得る. よって定理は示された. □

4.6 章末注釈と参考文献

第 4.1 節 は [15, 33] を参考にした. 注意 4.4 は [15] からの借用である. 第 4.2 節 は [29] を借用した. 第 4.4 節 は [15, 33] を借用した. 定理 4.13 の証明は [21] からの借用である. 定理 4.23 の証明は [9] からの借用である.

4.7 演習問題

演習問題 4.1. $n \in \mathbb{N}$ とする. 離散型確率変数 X_n は $\left\{\frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}$ 上の一様分布に従うとする. すなわち

$$\mathsf{Pr}\left(X_n = \frac{k}{n}\right) = \begin{cases} \dfrac{1}{n} & (k = 1, 2, \ldots, n) \\ 0 & (\text{その他の場合}) \end{cases}$$

である. 連続型確率変数 X は $(0, 1)$ 上の一様分布に従うとする. すなわち, X は p.d.f.

$$\mathsf{p}(x) = \begin{cases} 1 & (0 < x < 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

をもつ. このとき, 以下の問いに答えよ.
(1) $\dfrac{k}{n} \leq x < \dfrac{(k+1)}{n}$ $(k = 1, \ldots, n)$ に対して

$$\mathsf{Pr}(X_n \leq x)$$

を求めることで, $x \in \mathbb{R}$ に対して $\mathsf{Pr}(X_n \leq x)$ を求めよ.
(2) $x \in \mathbb{R}$ に対して

$$\mathsf{Pr}(X \leq x)$$

を求めよ.
(3)

$$\lim_{n \to \infty} \left| \mathsf{Pr}(X_n \leq x) - \mathsf{Pr}(X \leq x) \right|$$

を評価することで $X_n \rightsquigarrow X$ を示せ.

演習問題 4.2. $\{X_n\}_{n=1}^{\infty}$ を確率変数列とし, $\mathsf{E}[X_n^2] < \infty \, (n \in \mathbb{N})$ とする.
(1) $\mu \in \mathbb{R}$ を定数とする.

$$\mathsf{E}[(X_n - \mu)^2] = \mathsf{Var}[X_n] + \left\{\mathsf{E}[X_n] - \mu\right\}^2$$

が成り立つことを示せ.
(2) 次の条件は同値であることを示せ.

(a) $\lim_{n \to \infty} \mathsf{E}[(X_n - \mu)^2] = 0.$

(b) $\lim_{n \to \infty} \mathsf{E}[X_n] = \mu$ かつ $\lim_{n \to \infty} \mathsf{Var}[X_n] = 0$

演習問題 4.3. $\{X_n\}_{n=1}^{\infty}$, $\{Y_n\}_{n=1}^{\infty}$, X, Y を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ で定義された確率変数列とする．確率ベクトル (X_n, Y_n) が確率ベクトル (X, Y) に確率収束するとは, 任意の $\epsilon > 0$ に対して

$$\lim_{n\to\infty} \mathsf{Pr}\Big(\sqrt{(X_n - X)^2 + (Y_n - Y)^2} \geq \epsilon\Big) = 0$$

が成り立つこととする．これを $(X_n, Y_n) \xrightarrow{\mathrm{P}} (X, Y)$ と記すことにする. このとき, 以下の問いを答えよ.

(1) $X_n \xrightarrow{\mathrm{P}} X$ かつ $Y_n \xrightarrow{\mathrm{P}} Y$ ならば, $(X_n, Y_n) \xrightarrow{\mathrm{P}} (X, Y)$ が成り立つことを示せ.

(2) $(X_n, Y_n) \xrightarrow{\mathrm{P}} (X, Y)$ ならば, $X_n \xrightarrow{\mathrm{P}} X$ かつ $Y_n \xrightarrow{\mathrm{P}} Y$ が成り立つことを示せ.