第4章　確率変数列と分布の収束

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

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

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．

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

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 $X_1, X_2, \ldots$ を確率変数列とし, X を同じ確率空間上で定義された別の確率変数とする．各 $X_j\,(j = 1, 2, \ldots)$ 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

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

が成り立つときをいう．このことを $X_n \xrightarrow{P} X$ 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 $X_n \xrightarrow{\text{a.s.}} X$ 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 $X_n \xrightarrow{P} X$ とき $X_n \xrightarrow{P} c$ 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 $\mathsf{E}\big[X_n^2\big] < \infty$, $\mathsf{E}[X^2] < \infty\,(n = 1, 2, \ldots)$ 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

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

が成り立つときをいう．このことを $X_n \xrightarrow{qm} c$ と記す

注意 4.4. $n = 1, 2, \ldots$ に対して $X_n \sim \mathsf{N}(0, 1/n)$ とする．$\{X_n\}$ 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．$F(x)$ 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 $\sqrt{n}X_n \sim \mathsf{N}(0, 1)$ であることに注意する．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

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

最後は $\sqrt{n}x \longrightarrow -\infty\,(n \to \infty)$ よりわかる．

つぎに $x > 0$ のとき

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

最後は $\sqrt{n}x \longrightarrow \infty\,(n \to \infty)$ よりわかる．したがって $x \neq 0$ のとき

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

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

次に確率収束を示そう．Markov の不等式 (定理 3.1) より，$\forall\epsilon > 0$ に対して

$$\mathsf{Pr}\bigl(|X_n| > \epsilon\bigr) = \mathsf{Pr}\bigl(X_n^2 > \epsilon^2\bigr) \le \frac{\mathsf{E}\bigl[X_n^2\bigr]}{\epsilon^2} = \frac{1}{n\epsilon^2} \longrightarrow 0\,(n \to \infty).$$

よって $X_n \xrightarrow{P} 0$ である． □

次に収束のタイプ間の関係について述べる．

定理 4.5. 次に関係が成立する．

(1) $X_n \xrightarrow{qm} X \Longrightarrow X_n \xrightarrow{P} X$.

(2) $X_n \xrightarrow{P} X \Longrightarrow X_n \rightsquigarrow X$.

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

Proof. 節 4.4.5.2 で示す． □

注意 4.6. 定理 4.5(1)(2) の逆は一般に成立しない．
定理 4.5(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

$$\begin{aligned}\mathsf{Pr}\bigl(|X_n| > \epsilon\bigr) &= \mathsf{Pr}\bigl(\sqrt{n}\mathbb{1}_{(0,\,1/n)}(U) > \epsilon\bigr) = \mathsf{Pr}\left(0 < U < \frac{1}{n}\right)\\ &= \mathsf{Pr}\left(0 \le U \le \frac{1}{n}\right) = \frac{1}{n} \longrightarrow 0 \quad (n \to \infty).\end{aligned}$$

よって $X_n \xrightarrow{P} X$．しかし

$$\mathsf{E}\bigl[X_n^2\bigr] = n\int_0^{1/n} du = 1$$

なので $X_n \xrightarrow{qm} X$ は成立しない．
定理 4.5(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 $X_n \sim \mathsf{N}(0,\,1)$．明らかに

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

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

$$\mathsf{Pr}\bigl(|X_n - X| > \epsilon\bigr) = \mathsf{Pr}(|2X| > \epsilon) = \mathsf{Pr}\left(|X| > \frac{\epsilon}{2}\right) \ne 0.$$

よって $X_n \xrightarrow{P} X$ は成立しない． □

定理 4.7. X_n, X, Y_n, Y $(n = 1, 2, \ldots)$ は確率変数列とする．$g : \mathbb{R} \longrightarrow \mathbb{R}$ を連続関数とし，c を定数とする．

(1) $X_n \xrightarrow{P} X$ かつ $Y_n \xrightarrow{P} Y \implies X_n + Y_n \xrightarrow{P} X + Y$.

(2) $X_n \xrightarrow{qm} X$ かつ $Y_n \xrightarrow{qm} Y \implies X_n + Y_n \xrightarrow{qm} X + Y$.

(3) $X_n \rightsquigarrow X$ かつ $Y_n \xrightarrow{P} c \implies X_n + Y_n \rightsquigarrow X + c$.

(4) $X_n \xrightarrow{P} X$ かつ $Y_n \xrightarrow{P} Y \implies X_n Y_n \xrightarrow{P} XY$.

(5) $X_n \rightsquigarrow X$ かつ $Y_n \xrightarrow{P} c \implies X_n Y_n \rightsquigarrow cX$.

(6) $X_n \xrightarrow{P} X \implies g(Y_n) \xrightarrow{P} g(Y)$.

(7) $X_n \rightsquigarrow X \implies g(X_n) \rightsquigarrow g(X)$.

Proof. □

4.2 大数の法則

定理 4.8. (大数の弱法則) $X_1, X_2, \ldots$ は確率空間 $\big(\Omega, \mathcal{A}, \mathsf{Pr}\big)$ 上の i.i.d. 確率変数列とする．$\mathsf{E}\big[|X_1|\big] < \infty$ のとき

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

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．$\mathsf{Var}[X_1] = \sigma^2$ とおく．Markov の不等式 (定理 3.1) より

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

よりわかる． □

定理 4.9. (大数の強法則) $X_1, X_2, \ldots$ は確率空間 $\big(\Omega, \mathcal{A}, \mathsf{Pr}\big)$ 上の i.i.d. 確率変数列とし，$\mathsf{E}\big[|X_1|^4\big] < \infty$ とする．このとき

$$\overline{X}_n := \frac{1}{n}\big(X_1 + X_2 + \cdots + X_n\big) \xrightarrow{\text{a.s.}} \mu, \qquad \mu = \mathsf{E}[X_1] \tag{4.1}$$

が成立する．

Proof.

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

とおく.

まず $\mu = 0$ として (4.1) を示す. 以下の事象の包含関係に注意する.

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

このことより

$$\mathsf{Pr}\left(\sum_{n=1}^{\infty} \overline{X}_n^4 < \infty\right) = 1 \tag{4.2}$$

がわかると

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

となり, $\mu = 0$ のときに (4.1) がわかる.

以下では $\mu = 0$ として (4.2) を示す. そのために多項定理を用いて T_n を展開する.

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

ただし $\ell_1, \ell_2, \ldots, \ell_n$ は 0 以上の整数である. $X_1^{\ell_1}, X_2^{\ell_2}, \ldots, X_n^{\ell_n}$ は独立なので

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

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

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

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

$$\mathsf{E}\big[X_k^2\big] \le \sqrt{\mathsf{E}\big[X_k^4\big]} = \sqrt{K}$$

となるので

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

を得る．よって

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

となる．したがって (4.2) が示せた．

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

$$\mathsf{E}\big[Y_k^4\big] = \mathsf{E}\Big[\big\{|X_k| + |\mu|\big\}^4\Big] \le 8\mathsf{E}\big[|X_k|^4 + |\mu|^4\big] \le 8\big(K + \mu^4\big) < \infty$$

が成り立つ．したがってこの定理の証明の前半部で得られた結果から

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

となる．最後に

$$\frac{1}{n}\big(Y_1 + Y_2 + \cdots + Y_n\big) = 0 \Longleftrightarrow \frac{1}{n}\big(X_1 + X_2 + \cdots + X_n\big) = \mu$$

に注意すればよい．　□

注意 4.10. 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．　□

4.3 中心極限定理

定理 4.11. $X_1,\,X_2,\,\ldots$ 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

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

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

Proof. 中心極限定理の証明は第 G 章でする．　□

例 4.12. $X_1, X_2, \ldots$ を i.i.d. 確率変数列とし，$\mathsf{E}\bigl[X_1\bigr] = \mu$, $\mathsf{Var}\bigl[X_1\bigr] = \sigma^2\,(0 < \sigma < \infty)$ とする．このとき中心極限定理 (定理 4.11) より

$$Z_n = \frac{\sqrt{n}\bigl(\overline{X}_n - \mu\bigr)}{\sigma} \rightsquigarrow \mathsf{N}(0,\, 1). \tag{4.3}$$

ここで

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

とおく．

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

となる．$\mathsf{E}[(X_1 - \mu)^2] = \sigma^2 < \infty$ なので，大数の法則 (定理 4.8) より

$$\frac{1}{n}\sum_{j=1}^{n}\bigl(X_j - \mu\bigr)^2 \xrightarrow{P} \mathsf{E}\bigl[\bigl(X_1 - \mu\bigr)^2\bigr] = \sigma^2, \tag{4.4}$$

$$\overline{X}_n \xrightarrow{P} \mu.$$

定理 4.7(6) より

$$\bigl(\overline{X}_n - \mu\bigr)^2 \xrightarrow{P} 0. \tag{4.5}$$

(4.4), (4.5) と定理 4.7(3)(5) より

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

さらに定理 4.7(6) を用いると

$$\frac{S_n}{\sigma} \xrightarrow{P} 1. \tag{4.6}$$

最後に (4.3), (4.6) と定理 4.7(5) より

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

を得る．

□

定理 4.13. (Berry-Essèen の不等式) $X_1, X_2, \ldots$ を i.i.d. 確率変数列とする．$\mathsf{E}\bigl[|X_1|^3\bigr] < \infty$ のとき

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

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

Proof. 信じることにする. □

定理 4.14. $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ 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

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

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

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. 証明は後の節で行う. □

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 $g : \mathbb{R} \longrightarrow \mathbb{R}$ に対して $g(Y_n)$ の極限分布を求めよう.

定理 4.15. (デルタ法) $Y_1, Y_2, \ldots$ を確率変数列とし

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

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 $\dot{g}(t) = \frac{dg}{dt}(t)$ である. このとき

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

すなわち

$$Y_n \approx \mathsf{N}\left(\mu,\, \frac{\sigma^2}{n}\right) \Longrightarrow g(Y_n) \approx \mathsf{N}\left(g(\mu),\, \bigl(\dot{g}(\mu)\bigr)^2 \frac{\sigma^2}{n}.\right)$$

ただし「≈」は「分布が近似できる」の意味である．

Proof. 節 4.4.5.2 で証明をする． □

例 4.16. $X_1, X_2, \ldots$ を i.i.d. 確率変数列とし，$\mathsf{E}\bigl[X_1\bigr] = \mu$, $\mathsf{Var}\bigl[X_1\bigr] = \sigma^2\,(0 < \sigma < \infty)$ とする．このとき中心極限定理 (定理 4.11) より

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

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 $\dot{g}(x) = e^x$ となる．よってデルタ法 (定理 4.15) より

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

よって

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

□

定理 4.17. (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

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

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 $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ 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

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

となる．

例 4.18.

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

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 4.14) より

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

いま

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

であるので

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

したがって

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

□

4.5　定理 4.5, 4.7, 4.15 の証明

4.5.1　定理 4.5 の証明

(1) $X_n \xrightarrow{qm} X$ 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

$$\mathsf{Pr}\big(|X_n - X| > \epsilon\big) = \mathsf{Pr}\big(|X_n - X|^2 > \epsilon^2\big) \le \frac{\mathsf{E}\big[|X_n - X|^2\big]}{\epsilon^2} \longrightarrow 0\,(n \to \infty).$$

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

$$
\begin{aligned}
F_n(x) &= \mathsf{Pr}\big(X_n \le x\big) \\
&= \mathsf{Pr}\big(X_n \le x,\, X \le x+\epsilon\big) + \mathsf{Pr}\big(X_n \le x,\, X > x+\epsilon\big) \\
&\le \mathsf{Pr}\big(X \le x+\epsilon\big) + \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big) &\quad (4.7)\\
&= F(x+\epsilon) + \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big). &\quad (4.8)
\end{aligned}
$$

同様に

$$
\begin{aligned}
\mathsf{F}(x-\epsilon) &= \mathsf{Pr}\big(X \le x-\epsilon\big) \\
&= \mathsf{Pr}\big(X_n \le x-\epsilon,\, X_n \le x\big) + \mathsf{Pr}\big(X \le x-\epsilon,\, X_n > x\big) \\
&\le \mathsf{Pr}\big(X_n \le x\big) + \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big) &\quad (4.9)\\
&= \mathsf{F}_n(x) + \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big). &\quad (4.10)
\end{aligned}
$$

(4.8) と (4.10) を合わせると

$$
\mathsf{F}(x-\epsilon) - \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big) \le F_n(x) \le \mathsf{F}(x+\epsilon) + \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big).
$$

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

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

(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[1]．このとき

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

□

問 4.1. (4.7) と (4.9) を確認せよ．すなわち

$$
\begin{aligned}
&\mathsf{Pr}\big(X_n \le x,\, X > x+\epsilon\big) \le \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big), \\
&\mathsf{Pr}\big(X \le x-\epsilon,\, X_n > x\big) \le \mathsf{Pr}\big(\big|X_n - X\big| > \epsilon\big)
\end{aligned}
$$

を示せ．

[1] $\mathsf{F}(x)$ 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．

4.5.2　定理 4.7 の証明

□

4.5.3　定理 4.15 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

$$g(x)-g(\mu)=(x-\mu)\int_0^1 \dot g\bigl(\mu+t(x-\mu)\bigr)\,dt \tag{4.11}$$

が成立する．また $\dot g(x)$ 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

$$|x-\mu|<\delta_2 \Longrightarrow \bigl|\dot g(x)-\dot g(\mu)\bigr|<\epsilon$$

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

$$\begin{aligned}\left|\int_0^1 \dot g\bigl(\mu+t(x-\mu)\bigr)\,dt-\dot g(\mu)\right| &\le \int_0^1 \bigl|\dot g\bigl(\mu+t(x-\mu)\bigr)-\dot g(\mu)\bigr|\,dt\\ &\le \epsilon\int_0^1 dt=\epsilon \end{aligned}\tag{4.12}$$

となる．(4.12) より

$$\mathsf{Pr}\left(\left|\int_0^1 \dot g\bigl(\mu+t(Y_n-\mu)\bigr)\,dt-\dot g(\mu)\right|<\epsilon\right)>\mathsf{Pr}\bigl(|Y_n-\mu|<\delta\bigr) \tag{4.13}$$

が成立する． $\sqrt{n}\bigl(Y_n-\mu\bigr)/\sigma \rightsquigarrow \mathsf{N}(0,\,1)$ なので $Y_n\xrightarrow{P}\mu$ である．(4.13) から

$$\int_0^1 \dot g\bigl(\mu+t(Y_n-\mu)\bigr)\,dt\xrightarrow{P}\dot g(\mu) \tag{4.14}$$

が成立する．(4.11) に $x=Y_n$ を代入すれば

$$\frac{\sqrt{n}\bigl(g(Y_n)-g(\mu)\bigr)}{\sigma}=\frac{\sqrt{n}\bigl(Y_n-\mu\bigr)}{\sigma}\int_0^1 \dot g\bigl(\mu+t(Y_n-\mu)\bigr)\,dt. \tag{4.15}$$

$\sqrt{n}(Y_n-\mu)/\sigma\rightsquigarrow\mathsf{N}(0,\,1)$ と (4.14) に注意して定理 4.7(5) を (4.15) に適用すれば

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

を得る．よって定理は示された． □

4.6　章末注釈と参考文献

4.7　演習問題