

確率・統計 B

確率変数と分布の収束

若木宏文

wakaki@math.sci.hiroshima-u.ac.jp

2013.10.16

table of contents

確率収束

概収束

分布収束

中心極限定理

漸近公式

確率収束 (4/6)

定義 5.1 (確率収束 (convergence in probability))

$\{X_n\}$: 確率変数列, θ : 定数, X : 確率変数

$$\forall \varepsilon > 0 \lim_{n \rightarrow \infty} P(|X_n - \theta| > \varepsilon) = 0$$

が成り立つとき, $\{X_n\}$ は θ に確率収束するといい,

$$X_n \xrightarrow{p} \theta \quad (n \rightarrow \infty)$$

とかく.

$$\forall \varepsilon > 0 \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0$$

が成り立つとき, $\{X_n\}$ は X に確率収束するといい,

$$X_n \xrightarrow{p} X \quad (n \rightarrow \infty)$$

とかく.

確率収束 (5/6)

定理 5.1 (大数の法則 (弱法則))

 $\{X_n\}$: 互いに独立な確率変数列,

$$E(X_n) = \mu, \text{Var}(X_n) = \sigma^2 \quad (n = 1, 2, \dots)$$

 \Rightarrow

$$\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j \xrightarrow{p} \mu \quad (n \rightarrow \infty)$$

概収束 (1/6)

定義 5.3 (概収束 (almost surely convergence))

 $\{X_n\}$: 確率変数列, X : 確率変数 $P(\lim_{n \rightarrow \infty} X_n = X) = 1$, すなわち

$$\exists E \subset \Omega \text{ s.t. } P(E) = 0, \forall \omega \in \Omega_0 = \Omega - E \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$$

が成り立つとき $\{X_n\}$ は X に概収束, または確率 1 で収束するとい
いい,

$$X_n \xrightarrow{a.s.} X \quad (n \rightarrow \infty)$$

とかく.

分布収束 (1/12)

定義 5.2 (分布収束 (convergence in distribution))

F_n : 確率変数 X_n の分布関数 ($n = 1, 2, \dots$),

F : 確率変数 X の分布関数

F の任意の連続点 x で

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

が成り立つとき, X_n は X に分布収束 (または 法則収束) するという. P_X を X_n の極限分布といい,

$$X_n \xrightarrow{d} X \quad (n \rightarrow \infty) \quad \text{または} \quad X_n \xrightarrow{d} P_X \quad (n \rightarrow \infty)$$

とかく. (P_X には $N(0, 1)$ などの分布を表す記号こともある)

分布収束 (6/12)

定理 5.3

確率変数列 X_1, X_2, \dots が X に確率収束すれば, その確率変数列は X に分布収束している. すなわち

$$X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$$

が成り立つ. とくに, $X \sim U(c)$ のときは, 逆も成り立つ. したがって

$$X_n \xrightarrow{d} c \Leftrightarrow X_n \xrightarrow{p} c.$$

が成り立つ.

分布収束 (7/12)

[証明]

F : X の分布関数, F_n : X_n の分布関数 $\varepsilon > 0$

$X_n > x$ かつ $|X_n - X| \leq \varepsilon \Rightarrow X > x - \varepsilon$

$X \leq x - \varepsilon \Rightarrow X_n \leq x$ または $|X_n - X| > \varepsilon$

$P(X \leq x - \varepsilon) \leq P(X_n \leq x) + P(|X_n - X| > \varepsilon)$

$F(x - \varepsilon) \leq F_n(x) + P(|X_n - X| > \varepsilon)$

$F(x - \varepsilon) \leq \liminf_{n \rightarrow \infty} F_n(x)$

分布収束 (8/12)

$$X > x + \varepsilon \text{ かつ } |X_n - X| \leq \varepsilon \Rightarrow X_n > x$$

対偶をとると

$$X_n \leq x \Rightarrow X \leq x + \varepsilon \text{ または } |X_n - X| > \varepsilon$$

$$P(X_n \leq x) \leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon)$$

$$F_n(x) \leq F(x + \varepsilon) + P(|X_n - X| > \varepsilon)$$

$$\limsup_{n \rightarrow \infty} F_n(x) \leq F(x + \varepsilon)$$

分布収束 (9/12)

$$F(x - \varepsilon) \leq \liminf_{n \rightarrow \infty} F_n(x) \leq \limsup_{n \rightarrow \infty} F_n(x) \leq F(x + \varepsilon)$$

$\varepsilon \rightarrow 0$ とすると, x が F の連続点ならば

$$F(x) \leq \liminf_{n \rightarrow \infty} F_n(x) \leq \limsup_{n \rightarrow \infty} F_n(x) \leq F(x)$$

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

分布収束 (10/12)

$P(X = c) = 1$ とし, 逆を示す. このとき

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

であり, $x \neq c \Rightarrow \lim_{n \rightarrow \infty} F_n(x) = F(x)$.

$$\begin{aligned} P(|X_n - c| \leq \varepsilon) &\geq P(c - \varepsilon < X_n \leq c + \varepsilon) \\ &= F_n(c + \varepsilon) - F_n(c - \varepsilon) \end{aligned}$$

$\varepsilon > 0$ ならば

$$\lim_{n \rightarrow \infty} P(|X_n - c| \leq \varepsilon) \geq F(c + \varepsilon) - F(c - \varepsilon) = 1 - 0 = 1.$$

$$\lim_{n \rightarrow \infty} P(|X_n - c| > \varepsilon) = 0.$$

分布収束 (11/12)

例 5.4 (定理 5.3 の逆の反例)

$$\Omega = \{1, 2, 3, 4\}, \quad P(\{\omega\}) = \frac{1}{4}, \quad \omega = 1, 2, 3, 4$$

とし, $n = 1, 2, \dots$, に対して

$$X_n(\omega) = \begin{cases} 1, & \omega = 1, 2 \\ 0, & \omega = 3, 4 \end{cases} \quad X(\omega) = \begin{cases} 0, & \omega = 1, 2 \\ 1, & \omega = 3, 4 \end{cases}$$

とすると $X_n \xrightarrow{d} X$ ($n \rightarrow \infty$) であるが

$$\forall \omega \in \Omega \quad |X_n(\omega) - X(\omega)| = 1$$

なので, 確率収束しない.

分布収束 (12/12)

定理 5.4 (レヴィの連続定理)

X_n ($n = 1, 2, \dots$) の特性関数を φ_n ($n = 1, 2, \dots$) とする. $\varphi_n(t)$ が各点で $\varphi(t)$ に収束し, $\varphi(t)$ が $t = 0$ で連続ならば $\varphi(t)$ は, ある確率分布 P_X の特性関数であって $X_n \xrightarrow{d} P_X$ ($n \rightarrow \infty$).

[証明は省略 (テキスト参照)]

分布収束と特性関数の収束

$\varphi_n(t)$: 確率変数 X_n の特性関数 ($n = 1, 2, \dots$)

$\varphi(t)$: 確率変数 X の特性関数

$$X_n \xrightarrow{d} X \quad (n \rightarrow \infty) \quad \Leftrightarrow \quad \forall t \quad \varphi_n(t) \rightarrow \varphi(t) \quad (n \rightarrow \infty)$$

中心極限定理 (1/5)

定理 5.5 (中心極限定理)

$\{X_n\}$: 独立に同一分布に従う確率変数列,

$$E(X_j) = \mu, \text{Var}(X_j) = \sigma^2, \quad j = 1, \dots, n$$

$$Z_n = \frac{1}{\sqrt{n}\sigma} \sum_{j=1}^n (X_j - \mu) = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{d} N(0, 1) \quad (n \rightarrow \infty)$$

[証明は板書で]

中心極限定理 (2/5)

注 5.6 (基準化)

確率変数 X に対して

$$Z = \frac{X - E(X)}{\sqrt{\text{Var}(X)}}$$

を X の基準化 という。

- (1) $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$ の基準化も, $S_n = \sum_{j=1}^n X_j$ の基準化も, $\sqrt{n}(\bar{X} - \mu)/\sigma$ となる.
- (2) a, b を定数とするとき, X の基準化と, $aX + b$ の基準化は同じ.

中心極限定理 (3/5)

注 5.7 (和の分布の近似)

$\{X_n\}$: 独立に同一分布に従う確率変数列

中心極限定理を用いると和 $S_n = \sum_{j=1}^n X_j$ の分布を

$$\begin{aligned} P(a < S_n \leq b) &= P\left(\frac{a - n\mu}{\sqrt{n}\sigma} < \sqrt{n}(\bar{X} - \mu)/\sigma \leq \frac{b - n\mu}{\sqrt{n}\sigma}\right) \\ &\approx \Phi\left(\frac{b - n\mu}{\sqrt{n}\sigma}\right) - \Phi\left(\frac{a - n\mu}{\sqrt{n}\sigma}\right) \end{aligned}$$

によって近似できる.

中心極限定理 (4/5)

注 5.8 (連続性の補正)

$\{X_n\}$: 独立に同一分布に従う確率変数列. 整数値のみをとる.

$$S_n = \sum_{j=1}^n X_j$$

S_n も整数値しかとらないので, a, b を整数値とするとき,

$$P(a \leq S_n \leq b) = P(a - 1 < S_n \leq b) \quad (1)$$

$$= P(a - 0.5 < S_n \leq b + 0.5) \quad (2)$$

中心極限定理 (5/5)

注 5.8 (連続性の補正 (つづき))

$E(X_j) = \mu, \text{Var}(X_j) = \sigma^2$ とする.

(1), (2) それぞれの, 中心極限定理による確率の近似式は

$$\Phi\left(\frac{b - n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{a - 1 - n\mu}{\sigma\sqrt{n}}\right) \quad (3)$$

$$\Phi\left(\frac{b + 0.5 - n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{a - 0.5 - n\mu}{\sigma\sqrt{n}}\right) \quad (4)$$

例 5.5 (2項分布の正規近似)

コインを 200 回投げて, 表が 95 回以上 105 回以下である確率

$\mu = 1/2, \sigma^2 = 1/4, n = 200, a = 95, b = 105$

正確な値 : 0.56325...

近似値 : (3) 0.56222(0.00103), (4) 0.56331(0.000006)

漸近公式 (1/9)

$$X_n \xrightarrow{a.s.} X \quad \Rightarrow \quad X_n \xrightarrow{p} X \quad \Rightarrow \quad X_n \xrightarrow{d} X$$

定理 5.11

- (1) $X_n \xrightarrow{p} a$ かつ $g(x)$ は $x = a$ で連続 $\Rightarrow g(X_n) \xrightarrow{p} g(a)$.
- (2) $X_n \xrightarrow{p} a, Y_n \xrightarrow{p} b$ かつ $g(x, y)$ は $(x, y) = (a, b)$ で連続 $\Rightarrow g(X_n, Y_n) \xrightarrow{p} g(a, b)$.
- (3) $X_n \xrightarrow{d} X$ かつ $g(x)$ は X の値域 ($= R(X)$) で連続 $\Rightarrow g(X_n) \xrightarrow{d} g(X)$.
- (4) $X_n \xrightarrow{d} X, Y_n \xrightarrow{p} b$ かつ $g(x, y)$ は $D = \{(x, b) \mid x \in R(X)\}$ で連続 $\Rightarrow g(X_n, Y_n) \xrightarrow{d} g(X, b)$.
- (5) $X_n \xrightarrow{a.s.} X$ かつ $g(x)$ は $R(X)$ で連続 $\Rightarrow g(X_n) \xrightarrow{a.s.} g(X)$.

[証明は (1) のみ, 板書で]

漸近公式 (2/9)

定理 5.12 (スラスキーの定理)

$X_n \xrightarrow{d} X, Y_n \xrightarrow{p} c$ 定数 とする. ただし, c は定数である. このとき

(1) $X_n + Y_n \xrightarrow{d} X + c.$

(2) $X_n Y_n \xrightarrow{d} cX.$

(3) $X_n/Y_n \xrightarrow{d} X/c \quad (c \neq 0, P(Y_n \neq 0) = 1).$

補題 1

分布関数の不連続点は高々可算個

[証明は板書で]

漸近公式 (3/9)

[証明]

- (1) $F_X(z)$: X の分布関数, $F_{X+c}(z)$: $X+c$ の分布関数
 $F_{X_n+Y_n}(z)$: $X_n + Y_n$ の分布関数

$$F_{X+c}(z) = P(X+c \leq z) = F_X(z-c)$$

$z-c$ が F_X の連続点のとき次を示せばよい.

$$\lim_{n \rightarrow \infty} F_{X_n+Y_n}(z) = F_X(z-c)$$

任意の $\varepsilon > 0$ に対して $0 < \varepsilon' < \varepsilon$ かつ $z-c \pm \varepsilon'$ が F_X の連続点となるような ε' をとる.

$$\begin{aligned} P(X_n + Y_n \leq z) &\leq P(X_n + Y_n \leq z, |Y_n - c| \leq \varepsilon') + P(|Y_n - c| > \varepsilon') \\ &\leq P(X_n \leq z - c + \varepsilon') + P(|Y_n - c| > \varepsilon') \end{aligned}$$

$$\limsup_{n \rightarrow \infty} F_{X_n+Y_n}(z) \leq F_X(z-c+\varepsilon')$$

漸近公式 (4/9)

$$\begin{aligned} \mathbf{P}(X_n \leq z - c - \varepsilon') &\leq \mathbf{P}(X_n \leq z - c - \varepsilon', |Y_n - c| \leq \varepsilon') \\ &\quad + \mathbf{P}(|Y_n - c| > \varepsilon') \\ &\leq \mathbf{P}(X_n + Y_n \leq z) + \mathbf{P}(|Y_n - c| > \varepsilon') \end{aligned}$$

$$F_X(z - c - \varepsilon') \leq \liminf_{n \rightarrow \infty} F_{X_n + Y_n}(z)$$

ε' はいくらでも小さくとれるので

$$F_X(z - c) \leq \liminf_{n \rightarrow \infty} F_{X_n + Y_n}(z) \leq \limsup_{n \rightarrow \infty} F_{X_n + Y_n}(z) \leq F_X(z - c)$$

漸近公式 (5/9)

$$(2) \quad X_n Y_n = X_n c + X_n (Y_n - c), \quad X_n c \xrightarrow{d} Xc \quad (n \rightarrow \infty)$$

なので, $X_n(Y_n - c) \rightarrow 0$ を示せばよい. 任意の $\varepsilon > 0$ に対して, $\{1 - F_X(M)\} + F_X(-M) < \varepsilon$ かつ $\pm M$ が F_X の連続点であるような M をとると, 任意の $\delta > 0$ に対して

$$\begin{aligned} & P(|X_n(Y_n - c)| > \delta) \\ & \leq P(X_n \leq -M) + P(X_n > M) + P(|Y_n - c| > \delta/M) \end{aligned}$$

$$\limsup_{n \rightarrow \infty} P(|X_n(Y_n - c)| > \delta) \leq F_X(-M) + \{1 - F_X(M)\} < \varepsilon$$

より $X_n(Y_n - c) \rightarrow 0$

$$(3) \quad \frac{1}{Y_n} \rightarrow \frac{1}{c} \quad (n \rightarrow \infty) \quad \text{なので, (2) より } X_n \cdot \frac{1}{Y_n} \rightarrow X \cdot \frac{1}{c}$$

漸近公式 (6/9)

定義 5.1 (多次元確率ベクトルと分布の収束)

$F(x)$: X を k 次元確率ベクトルの分布関数.

- (1) 任意の正数 ε に対して $\lim_{n \rightarrow \infty} P(\|X_n - X\| > \varepsilon) = 0$ が成り立つとき, $\{X_n\}$ は X に確率収束するといい, $X_n \xrightarrow{P} X$ とかく. ここで, $\|\cdot\|$ はユークリッドノルム,
- (2) $F(x)$ のすべての連続点で $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ が成り立つとき, X_n は X に分布収束 (または, 法則収束) するといい, X の分布を P_X とするとき, $X_n \xrightarrow{d} X$, または, $X_n \xrightarrow{d} P_X$ とかく.
- (3) $P(\lim_{n \rightarrow \infty} X_n = X) = 1$ が成り立つとき, $\{X_n\}$ は X に概収束する, または確率 1 で収束するといい, $X_n \xrightarrow{a.s.} X$ とかく.

漸近公式 (7/9)

命題 1 (確率ベクトルの収束と成分の収束)

$\mathbf{X}_n = (X_{n_1}, \dots, X_{n_k})'$, $\mathbf{X} = (X_1, \dots, X_k)'$ とするとき, 次が成り立つ.

$$(1) X_{n_i} \xrightarrow{p} X_i, i = 1, \dots, k \Leftrightarrow \mathbf{X}_n \xrightarrow{p} \mathbf{X}.$$

$$(2) X_{n_i} \xrightarrow{a.s.} X_i, i = 1, \dots, k \Leftrightarrow \mathbf{X}_n \xrightarrow{a.s.} \mathbf{X}.$$

漸近公式 (8/9)

例 5.6

(X_n, Y_n) : 次の確率密度関数を持つ 2 次元連続型確率ベクトル.

$$f(x, y) = \begin{cases} 1/2, & -1 < x < 1, -1 < y < 1, xy > 0 \\ 0, & \text{その他} \end{cases}$$

(X, Y) : 矩形 $(-1, 1) \times (-1, 1)$ 上の 2 次元一様分布に従う確率ベクトル.

X_n, Y_n, X, Y の周辺分布はすべて区間 $(-1, 1)$ 上の一様分布なので, $X_n \xrightarrow{d} X, Y_n \xrightarrow{d} Y$

(X_n, Y_n) の極限分布は, $(-1, 1) \times (-1, 1)$ 上の 2 次元一様分布ではない.

漸近公式 (9/9)

定理 5.13 (確率ベクトルの分布収束と特性関数の収束)

$\varphi(\mathbf{t})$: k 次元確率ベクトル \mathbf{X} 特性関数

$\varphi_n(\mathbf{t})$: k 次元確率ベクトル \mathbf{X}_n , $n = 1, 2, \dots$ の特性関数

次の (1) ~ (3) は同値.

(1) $\mathbf{X}_n \xrightarrow{d} X (n \rightarrow \infty)$.

(2) 任意の有界連続関数 $g(x)$ に対して
 $\lim_{n \rightarrow \infty} E\{g(\mathbf{X}_n)\} = E\{g(\mathbf{X})\}$.

(3) $\lim_{n \rightarrow \infty} \varphi_n(\mathbf{t}) = \varphi(\mathbf{t})$.

(参考: 「確率論」西尾真喜子, 第 5 章)