第 13 回　大数の法則と中心極限定理（8）

村澤 康友

2023 年 11 月 6 日

今日のポイント

1. 確率変数列 $\{X_i\}$ の標本平均 $\bar{X}_n := (X_1 + \cdots + X_n)/n$ の分布を近似する．
2. $\{X_i\}$ が平均 μ，分散 σ^2 の独立かつ同一な分布をもつなら $\{\bar{X}_n\}$ は μ に確率収束（大数の法則）．
3. $\{Z_i\}$ が平均 0，分散 1 の独立かつ同一な分布をもつなら $\{\sqrt{n}\bar{Z}_n\}$ は $\mathrm{N}(0,1)$ に分布収束（中心極限定理）．したがって $\bar{Z}_n \overset{a}{\sim} \mathrm{N}(0,1/n)$．

目次

1　標本平均（pp. 149, 183）

$\{X_i\}$ を確率変数列とする．

定義 1. $(X_1,\dots,X_n)$ の**標本平均**は

$$\bar{X}_n := \frac{X_1 + \cdots + X_n}{n}$$

注 1. 確率変数の平均（期待値）とは異なる．

定理 1. $X_1,\dots,X_n$ が平均 μ の同一な分布をもつなら

$$\mathrm{E}\left(\bar{X}_n\right) = \mu$$

証明. 期待値の線形性より

$$\begin{aligned}\mathrm{E}\left(\bar{X}_n\right) &= \mathrm{E}\left(\frac{X_1 + \cdots + X_n}{n}\right)\\ &= \frac{\mathrm{E}(X_1) + \cdots + \mathrm{E}(X_n)}{n}\\ &= \frac{\mu + \cdots + \mu}{n}\\ &= \mu\end{aligned}$$

□

定理 2. $X_1,\dots,X_n$ が分散 σ^2 の独立かつ同一な分布をもつなら

$$\operatorname{var}\left(\bar{X}_n\right) = \frac{\sigma^2}{n}$$

証明. $X_1,\dots,X_n$ は独立なので

$$\begin{aligned}\operatorname{var}\left(\bar{X}_n\right) &= \operatorname{var}\left(\frac{X_1 + \cdots + X_n}{n}\right)\\ &= \frac{\operatorname{var}(X_1 + \cdots + X_n)}{n^2}\\ &= \frac{\operatorname{var}(X_1) + \cdots + \operatorname{var}(X_n)}{n^2}\\ &= \frac{\sigma^2 + \cdots + \sigma^2}{n^2}\\ &= \frac{\sigma^2}{n}\end{aligned}$$

□

2　大数の法則

2.1　確率収束（p. 162）

$\{x_n\}$ を実数列，$\{X_n\}$ を確率変数列とする．

定義 2. 任意の $\epsilon > 0$ について，ある自然数 $N(\epsilon)$ が存在し，

$$n \geq N(\epsilon) \Longrightarrow |x_n - c| < \epsilon$$

なら $\{x_n\}$ は c に**収束**するという．

注 2. $\lim_{n\to\infty} x_n = c$ または $x_n \to c$ と書く．

定義 3. 任意の $\epsilon > 0$ について

$$\lim_{n\to\infty} \Pr[|X_n - c| < \epsilon] = 1$$

なら $\{X_n\}$ は c に**確率収束**するという．

注 3. $\operatorname{plim}_{n\to\infty} X_n = c$ または $X_n \xrightarrow{p} c$ と書く．

注 4. 確率変数列の収束の概念は他にもある．

2.2　大数の法則（p. 160）

定理 3 (チェビシェフの大数の弱法則)**.** $\{X_i\}$ が平均 μ，分散 σ^2 の独立かつ同一な分布をもつなら

$$\operatorname*{plim}_{n\to\infty} \bar{X}_n = \mu$$

証明. チェビシェフの不等式より，任意の $\epsilon > 0$ について

$$\Pr\left[\left|\bar{X}_n - \mathrm{E}\left(\bar{X}_n\right)\right| \geq \epsilon\right] \leq \frac{\operatorname{var}\left(\bar{X}_n\right)}{\epsilon^2}$$

すなわち

$$\Pr\left[\left|\bar{X}_n - \mu\right| \geq \epsilon\right] \leq \frac{\sigma^2/n}{\epsilon^2}$$

余事象の確率は

$$\Pr\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] > 1 - \frac{\sigma^2/n}{\epsilon^2}$$

$n \to \infty$ の極限をとると

$$\lim_{n\to\infty} \Pr\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] \geq 1$$

確率は 1 以下なので等号が成立．　□

例 1. コインを 10 回，100 回，1000 回と投げ続けると表の出る割合は 1/2 に近づく（図 1）．

3　中心極限定理

3.1　分布収束

$\{X_n\}$ に対応する cdf の列を $\{F_n(.)\}$ とする．

定義 4. $F(.)$ の任意の連続点 x で

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

なら $\{X_n\}$ は $F(.)$ に**分布（法則）収束**するという．

注 5. $X_n \xrightarrow{d} F(.)$ と書く．

3.2　総乗記号

定義 5.

$$\prod_{i=1}^n x_i := x_1 \cdots x_n$$

練習 1. 以下の公式を示しなさい．

1. $\prod_{i=1}^n ax_i = a^n \prod_{i=1}^n x_i$
2. $\prod_{i=1}^n a^{x_i} = a^{\sum_{i=1}^n x_i}$
3. $\prod_{i=1}^n x_i y_i = \prod_{i=1}^n x_i \prod_{i=1}^n y_i$

3.3　中心極限定理（p. 162）

定理 4 (リンドバーグ＝レヴィの中心極限定理)**.** $\{Z_i\}$ が平均 0，分散 1 の独立かつ同一な分布をもつなら

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \xrightarrow{d} \mathrm{N}(0,1)$$

証明. $(1/\sqrt{n}) \sum_{i=1}^n Z_i$ の mgf が $\mathrm{N}(0,1)$ の mgf に収束することを示せばよい．すなわち

$$\lim_{n\to\infty} M_{\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i}(t) = \mathrm{e}^{t^2/2}$$

を示したい．$Z_1, \ldots, Z_n$ は独立かつ同一な分布をもつので

$$\begin{aligned} M_{\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i}(t) &:= \mathrm{E}\left(\mathrm{e}^{\frac{t}{\sqrt{n}} \sum_{i=1}^n Z_i}\right) \\ &= \mathrm{E}\left(\prod_{i=1}^n \mathrm{e}^{\frac{t}{\sqrt{n}} Z_i}\right) \\ &= \prod_{i=1}^n \mathrm{E}\left(\mathrm{e}^{\frac{t}{\sqrt{n}} Z_i}\right) \\ &= M_Z\left(\frac{t}{\sqrt{n}}\right)^n \end{aligned}$$

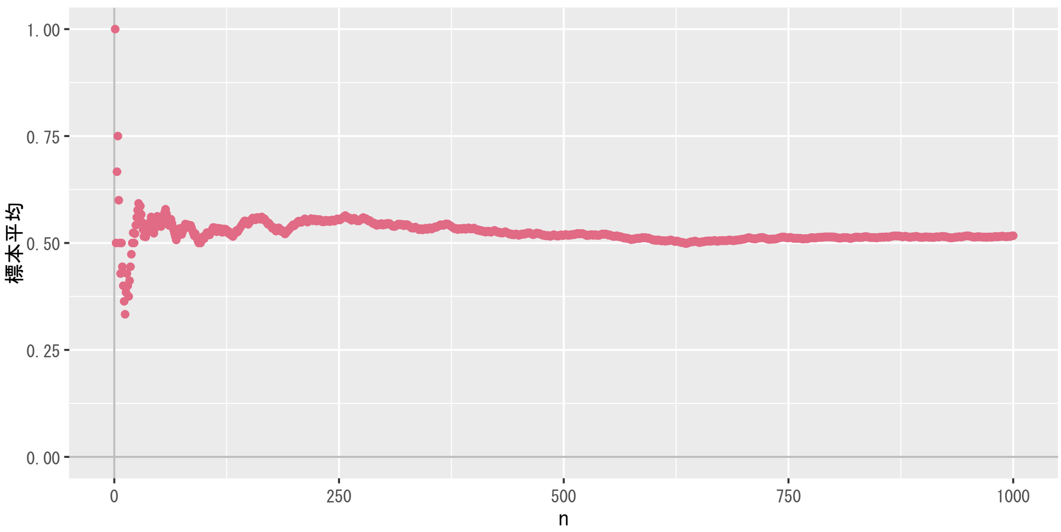

図 1　n 回のコイントスにおける表の割合

$M'_Z(0) = 0$, $M''_Z(0) = 1$ なので，マクローリン展開より

$$
\begin{aligned}
&M_Z\left(\frac{t}{\sqrt{n}}\right) \\
&= M_Z(0) + M'_Z(0)\frac{t}{\sqrt{n}} + \frac{M''_Z(0)}{2}\left(\frac{t}{\sqrt{n}}\right)^2 + \cdots \\
&= 1 + \frac{t^2}{2n} + \cdots
\end{aligned}
$$

したがって

$$
\begin{aligned}
\lim_{n\to\infty} M_Z\left(\frac{t}{\sqrt{n}}\right)^n &= \lim_{n\to\infty}\left(1 + \frac{t^2}{2n} + \cdots\right)^n \\
&= \mathrm{e}^{t^2/2}
\end{aligned}
$$

□

注 6. 式変形すると

$$\sqrt{n}\bar{Z}_n \xrightarrow{d} \mathrm{N}(0,1)$$

系 1. $\{X_i\}$ が平均 μ, 分散 σ^2 の独立かつ同一な分布をもつなら

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma} \xrightarrow{d} \mathrm{N}(0,1)$$

証明. $Z_i := (X_i - \mu)/\sigma$ として前定理を適用. □

注 7. 式変形すると

$$\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \xrightarrow{d} \mathrm{N}(0,1)$$

定義 6. n が大きいときの X_n の近似分布を**漸近分布**という.

注 8. 中心極限定理より

$$
\begin{aligned}
\sqrt{n}\bar{Z}_n &\overset{a}{\sim} \mathrm{N}(0,1) \\
\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} &\overset{a}{\sim} \mathrm{N}(0,1)
\end{aligned}
$$

すなわち

$$
\begin{aligned}
\bar{Z}_n &\overset{a}{\sim} \mathrm{N}\left(0, \frac{1}{n}\right) \\
\bar{X}_n &\overset{a}{\sim} \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)
\end{aligned}
$$

ただし $\overset{a}{\sim}$ は漸近分布を表す.

例 2. 指数乱数の標本平均の分布（図 2）.

3.4　正規乱数の生成（p. 171）

$\{U_i\}$ を $[0,1]$ 上の一様乱数の列とすると

$$
\begin{aligned}
\mathrm{E}(U_i) &= \frac{1}{2} \\
\mathrm{var}(U_i) &= \frac{1}{12}
\end{aligned}
$$

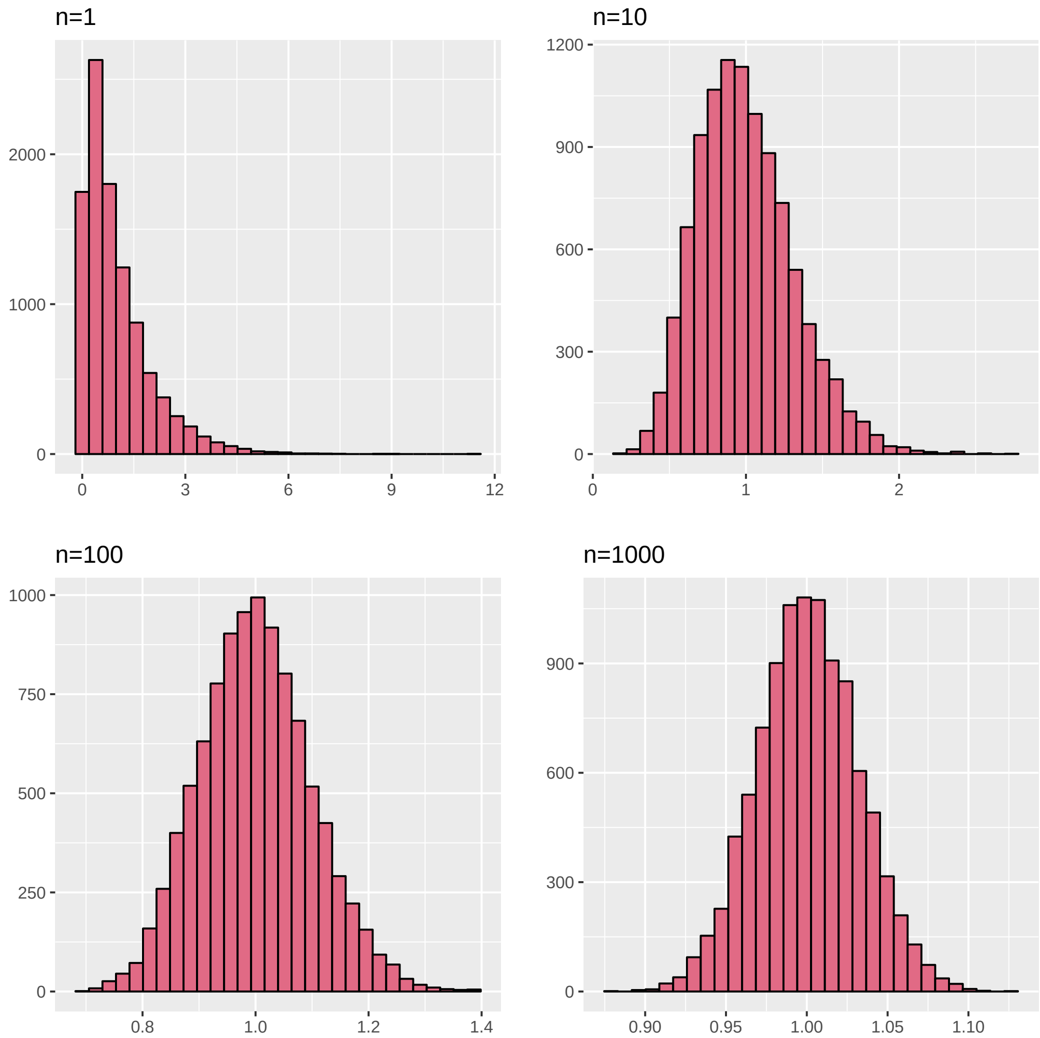

図 2　指数乱数の標本平均の分布

$X := U_1 + \cdots + U_{12} - 6$ とすると

$$
\begin{aligned}
\mathrm{E}(X) &= 0 \\
\mathrm{var}(X) &= 1
\end{aligned}
$$

中心極限定理より

$$X \overset{a}{\sim} \mathrm{N}(0, 1)$$

これを利用して一様乱数から標準正規乱数が生成できる（図 3）.

4　今日のキーワード

標本平均，確率収束，大数の法則，分布収束，中心極限定理，漸近分布

5　次回までの準備

提出　宿題 4，復習テスト 9–13

復習　教科書第 8 章，復習テスト 13

試験　(1) 教科書を読む (2) 用語の定義を覚える (3) 復習テストを自力で解く (4) 過去問に挑戦

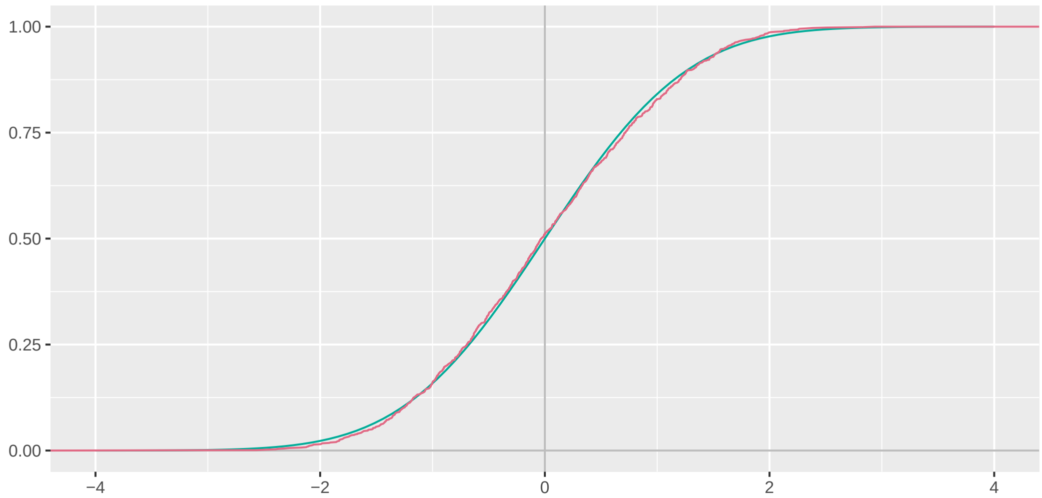

図 3　1000 個の標準正規乱数の累積相対度数グラフと N(0, 1) の cdf