13　大数の法則・中心極限定理

導入

本講義の最後に，現代の統計・確率論で重要な概念である大数の法則および中心極限定理について紹介する．

13.1　確率変数列

確率変数列という，確率変数からなる列 $\{X_n\}$ を考えることができる．例えばコインを複数回投げる試行において，X_n を n 回目に表が出たら 1，裏が出たら 0 となる確率変数とすれば，$\{X_n\}$ は確率変数列になる．

このような確率変数列が与えられると，その平均

$$\overline{X} = \frac{X_1 + \cdots + X_n}{n}$$

もまた確率変数になっている．大数の法則は，大まかに言えば各確率変数 X_i は互いに独立で，その平均がすべて μ であるとき，確率変数列 $\{X_n\}$ の平均の極限について

$$\lim_{n\to\infty} \overline{X}_n = \lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mu$$

が成り立つというものである．直感的には正しそうに思えるが，確率変数とは「関数」であった．「関数の極限」は決して自明な概念ではない．関数列の「収束」については標準的なものはなく，どの意味での収束なのかを明確にする必要がある．

13.2　大数の法則

定義 13.1

確率変数列 $\{X_n\}_{n=1,2,\ldots}$ と確率変数 X が与えられているとする．X_n が X に**確率収束**するとは，任意の $\varepsilon > 0$ に対して

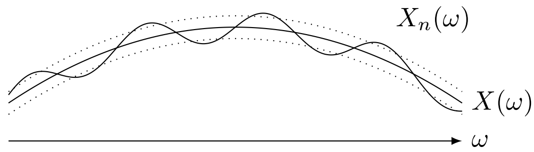
$$\lim_{n\to\infty} P\Big(\big|X_n - X\big| > \varepsilon\Big) = 0$$

のときにいう．

左辺に現れる $P\Big(\big|X_n - X\big| > \varepsilon\Big)$ は，正確に書けば

$$P\Big(\big|X_n - X\big| > \varepsilon\Big) = P\left(\{w \in \Omega;\, |X_n(\omega) - X(\omega)| > \varepsilon\}\right)$$

である．すなわち，$X(\omega)$ との差が ε よりも大きくなるような ω の集合 (下図でいえば点線からはみ出している部分の x 軸部分) が，n を大きくするにつれて限りなく 0 に近づいていくことを表している．



確率変数列 $\{X_n\}$ が**互いに独立**とは，任意の相異なる自然数 i, j に対して X_i と X_j が独立であることをいう．

定理 13.2（大数の法則）

$\{X_n\}$ を確率変数列で，それぞれについて

$$\mathbb{E}[X_n] = \mu, \quad V[X_n] = \sigma^2 < \infty$$

を満たすとする．このとき，この確率変数列の第 n 項までの平均

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

について，$\{X_n\}$ が互いに独立であるならば，$\overline{X}_n$ は μ に確率収束する．

13.3　大数の法則の証明

次のチェビシェフの不等式を利用する．

定理 13.3（チェビシェフの不等式）

任意の実数 $\alpha > 0$ に対して

$$P\Big(\big|X - \mathbb{E}[X]\big| \geqq \alpha\Big) \leqq \frac{V[X]}{\alpha^2}$$

この不等式を利用して，まずは大数の法則を証明しよう．

証明．　この不等式を利用すれば

$$\begin{aligned} P\Big(\big|\overline{X}_n - \mu\big| \geqq \varepsilon\Big) &\leqq \frac{V[\overline{X}_n]}{\varepsilon^2} = \frac{1}{\varepsilon^2}\cdot\frac{1}{n^2} V\Big[\sum_{i=1}^{n} X_i\Big] \\ &= \frac{1}{\varepsilon^2}\cdot\frac{1}{n^2}\sum_{i=1}^{n} V[X_i] = \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \end{aligned}$$

これと $P(|\overline{X}_n - \mu| \geqq \varepsilon) \geqq 0$ であることから，はさみうちの定理より

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| \geqq \varepsilon) = 0$$

となる．　□

チェビシェフの不等式は，より一般的な不等式であるマルコフの不等式の特別な場合である．

定理 13.4（マルコフの不等式）

Z を確率変数とし，u を $\mathbb{R}$ 上で非負の単調非減少関数とする．このとき任意の $u(\alpha) > 0$ なる実数 α に対して

$$P(Z \geqq \alpha) \leqq \frac{E[u(Z)]}{u(\alpha)}$$

証明． (1) まず

$$I_{Z \geqq \alpha} = \begin{cases} 1 & (Z \geqq \alpha), \\ 0 & (Z < \alpha) \end{cases}$$

とすれば，

$$\mathbb{E}[I_{Z \geqq \alpha}] = P(Z \geqq \alpha)$$

となることに注意する．また $u(x)$ の単調非減少性から

$$u(\alpha) I_{Z \geqq \alpha} \leqq u(Z)$$

が成り立つ．この不等式において期待値を計算すれば，非負関数の積分は不等号を保存するので

$$\mathbb{E}\Big[u(\alpha) I_{Z \geqq \alpha}\Big] \leqq \mathbb{E}\big[u(Z)\big]$$

ここで

$$\mathbb{E}\Big[u(\alpha) I_{Z \geqq \alpha}\Big] = u(\alpha)\mathbb{E}\Big[I_{Z \geqq \alpha}\Big] = u(\alpha) P(Z \geqq \alpha)$$

(2) このマルコフの不等式において

$$Z = \big|X - \mathbb{E}[X]\big|, \quad u(x) = \begin{cases} x^2 & (x \geqq 0), \\ 0 & (x < 0) \end{cases}$$

とすればチェビシェフの不等式を得る． □

注意． 大数の強法則というものも存在する．これは，確率変数列により強い仮定を課すことにより，確率収束よりも強い**概収束性**を導けるというものである．なお X_n が X に**概収束**するとは，次が成り立つことをいう．

$$P\left(\lim_{n \to \infty} X_n = X\right) = 1.$$

13.4 中心極限定理の主張・意義

確率変数列 $\{X_n\}$ の平均 $\overline{X}_n$ の極限について，より詳しいことが知られている．大雑把に言えば「平均の極限は正規分布で近似できる」ということである．

定理 13.5（中心極限定理）

$X_1, X_2, \ldots, X_n$ を互いに独立で同じ分布に従う確率変数とする．このとき

$$\mathbb{E}[X_i] = \mu, \quad V[X_i] = \sigma^2 > 0$$

が存在するならば，$\overline{X}_n$ に対して次が成立する．

$$\lim_{n \to \infty} P\left(\frac{\overline{X} - \mathbb{E}[\overline{X}]}{\sqrt{V[\overline{X}]}} \leqq x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

この主張の中には一見すると正規分布が出てきていないように見える．しかし，右辺は標準正規分布 $N(0,1)$ の分布関数である．標準正規分布の確率密度関数は

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

であったので，その分布関数 $F(x)$ は

$$F(x) = P(X \leqq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

である．よってこの定理の主張は「平均 $\overline{X}_n$ を正規化した $\frac{\overline{X}_n - \mathbb{E}[\overline{X}_n]}{\sqrt{V[\overline{X}_n]}}$ の分布関数の極限は，標準正規分布の分布関数と一致する」ということになる．このことを，「n が十分大きいとき近似的に

$$\frac{\overline{X}_n - \mathbb{E}[\overline{X}_n]}{\sqrt{V[\overline{X}_n]}} \ N(0,1)$$

が成立する」ということもある．

またこの定理により，元の X_i の従う分布が何であっても，n が十分大きい場合には $\overline{X}_n$ の分布は一定の分布 (すなわち正規分布) と仮定して議論できることになる．

13.5 証明のアイデア

次の補題を用いる．

補題 13.6

確率変数 X および確率変数列 $\{X_n\}$ を考える．これらに対応する分布関数をそれぞれ F_X, F_{X_n} とする．またこれらに対応するモーメント母関数 M_X, M_{X_n} はすべて存在すると仮定する．このとき次は同値．

- F_X のすべての連続点 x において

$$F_{X_n}(x) \to F_X(x) \quad (n \to \infty)$$

- 0 の近傍において

$$M_{X_n}(t) \to M_X(t) \quad (n \to \infty)$$

この補題より，モーメント母関数の各点収束に帰着させることができる．

証明． まず標準正規分布に従う確率変数 $X \sim N(0,1)$ のモーメント母関数が

$$M_X(t) = e^{\frac{t^2}{2}}$$

となることに注意する．確率変数 Z_i を以下で定義する．

$$Z_i = \frac{X_i - \mu}{\sigma}, \quad \mathbb{E}[Z_i] = 0, \quad V[Z_i] = 1$$

このとき

$$T_n = \frac{\overline{X}_n - \mathbb{E}[\overline{X}_n]}{\sqrt{V[\overline{X}_n]}} = \frac{1}{\sqrt{\sigma}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma} = \frac{1}{\sqrt{\sigma}} \sum_{i=1}^{n} Z_i$$

とすれば，T_n のモーメント母関数は

$$\begin{aligned} M_{T_n}(t) = \mathbb{E}\left[e^{tT_n}\right] &= \mathbb{E}\left[e^{\frac{t}{\sqrt{n}}\sum_i Z_i}\right] \\ = \prod_{i=1}^{n} \mathbb{E}\left[e^{\frac{t}{\sqrt{n}}Z_i}\right] = \prod_{i=1}^{n} M_{Z_i}\left(\frac{t}{\sqrt{n}}\right) &= M_{Z_1}\left(\frac{t}{\sqrt{n}}\right)^n \end{aligned}$$

のようになる．ここで Z_1 のモーメント母関数は

$$M_{Z_1}(t) = 1 + \frac{t^2}{2} + o(t^2)$$

となるので

$$\begin{aligned} & M_{T_n}(t) \\ &= \left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n = \left(1 + \frac{t^2}{2n}\right)^n + n\,o\left(\frac{t^2}{n}\right) \\ &= \left(\left(1 + \frac{t^2}{2n}\right)^{\frac{2n}{t^2}}\right)^{\frac{t^2}{2}} + n\,o\left(\frac{t^2}{n}\right) \to e^{\frac{t^2}{2}} \quad (n \to \infty). \end{aligned}$$

よって T_n のモーメント母関数が標準正規分布のモーメント母関数に各点収束するので，T_n の分布関数が標準正規分布の分布関数に各点収束する． □

注意． 本来は定義できない場合があるモーメント母関数ではなく，いつでも定義可能な特性関数を用いて証明をする．

$$\varphi_X(t) = \mathbb{E}[e^{itX}]$$