Chapter 5　大数の法則と中心極限定理

5.1　例題

確率論を担当しているＩ教授が，「今日の講義はこれで終わりです．ただし，まだこの部屋を出てはいけません．これから皆さんに『正しく作られたコイン』を 10 回ずつ投げていただきます．そして，投げたコインが『正しく作られたコイン』であることを実証できた人，すなわち，表の出現比率が 0.5 ± 0.1 の範囲におさまった人に帰宅を許すことにします．では，コインを投げて下さい」と告げた．今年度の確率論の受講者は 100 名であるとする．そのうちの何名が，10 回投げ終わった時点で帰宅を許されるのだろうか.

コインをｎ回投げるとき，ｋ回目の結果を表す確率変数をつぎのように定めよう．

$$X_k = \begin{cases} 1, & \text{k 回目に投げたとき「表」が出る} \\ 0, & \text{k 回目に投げたとき「裏」が出る} \end{cases}$$

このとき，$X_1, \ldots, X_n$ は，$P\{X_k = 1\} = \mu$，$P\{X_k = 0\} = 1 - \mu$ を満たす独立なベルヌーイ型確率変数であると考えてよい.「正しく作られたコイン」であるならば，$\mu = 0.5$ である．

正しく作られたコインを投げるとしても，ｎ回投げたときの表の出現比率 $\overline{X} = \dfrac{1}{n} \sum_{k=1}^{n} X_k$ が 0.5 になるとは限らない．0.47 や 0.52 などの値になることもあるはずである．そこで,「正しいコインであること」を主張するにも,「表の出現比率が 0.5 ± 0.1」などといった幅をもたせた表現が必要になる．

さて,

$$P\{0.4 \le \frac{1}{n} \sum_{k} X_k \le 0.6\}$$

を計算してみよう．例えば，10 回投げたときを考えてみると,

$$P\{0.4 \le \frac{1}{10} \sum_{k=1}^{10} X_k \le 0.6\} = 0.66$$

この結果は，確率論の講義終了後に 10 回コインを投げた 100 名の受講者のうち,「投げたコインは，正しく作られたコインであった」と認定されて帰宅を許されるのは 66 名程度であり，34 名程度の受講者は引き続き教室に残らねばならないことになるであろうという予想を示している．

では，講義終了時にコインを投げる回数が，10 回ではなく 20 回であったなら，帰宅を許される受講者は何名になるのだろうか．これについては，つぎの計算が示すように，100 名の受講者のうち 74 名程度が帰宅できるであろうということができる．

$$P\{0.4 \le \frac{1}{20}\sum_{k=1}^{20} X_k \le 0.6\} = 0.74$$

同様に，講義終了時にコインを投げる回数が 30 回，40 回，50 回，あるいは 100 回と定められた場合に帰宅を許される受講者の数は，つぎの計算結果から予想することができる．

$$P\{0.4 \le \frac{1}{30}\sum_{k=1}^{30} X_k \le 0.6\} = 0.80$$

$$P\{0.4 \le \frac{1}{40}\sum_{k=1}^{40} X_k \le 0.6\} = 0.85$$

$$P\{0.4 \le \frac{1}{50}\sum_{k=1}^{50} X_k \le 0.6\} = 0.88$$

$$P\{0.4 \le \frac{1}{100}\sum_{k=1}^{100} X_k \le 0.6\} = 0.97$$

これらの計算は，投げたコインが正しく作られたコインであったことを表の出現比率で立証しようとするとき，多くの回数を投げるようにしたほうが，「本来あるべき姿」あるいは「正しく作られたコインらしさ」を再現することができる可能性が高いことを暗示している．実際，それが正しいことを主張するのが，大数の弱法則（weak law of large numbers）である．

一方，大数の法則には，強法則（strong law of large numbers）とよばれるものがある．こちらは，関心ある事象（A としよう）の生起・非生起を調べる独立試行を際限なく続けていけば，やがて事象 A が生起する相対度数は A の発生確率 $P\{A\}$ に近づいていくことを主張するものである．

確率論の講義終了後のコイン投げが，「受講生は正しく作られたコインを投げ続け，それまでの表の出現比率が 0.5 ± 0.1 に入った時点で帰宅してよい」というルールのもとで行われたとすると，10 回投げ終わった時点でまだ 34 名ほどが教室に残っているが，引き続いて 10 回コインを投げ終わる頃には教室に残っている人は 26 名ほどに減っている．そして，コインを投げ始めて 100 回を過ぎる頃には，教室に残っている受講生は 3 名ほどにまで減っているが，さらにコインを投げ続けるうちに，やがて誰もいなくなる．

5.2　大数の法則

事象 A について $P\{A\}$ を推定したい．n 回の独立反復試行に対して A の起こる相対度数を $\dfrac{r}{n}$ とするとき，$n \to \infty$ としていくと $\dfrac{r}{n} \to P\{A\}$ とならないだろうか. ベルヌーイは, 次式の成立を示した.

$$P\left\{\left|\frac{r}{n} - P\{A\}\right| \le \varepsilon\right\} \to 1 \quad (n \to \infty) \quad \forall \varepsilon > 0$$

このことは統計的確率（1.1 参照）のひとつの根拠となるものであり，大数の弱法則（weak law of large numbers）として，つぎのように表現される.

Thm. 1（大数の弱法則）　$X_1, \ldots, X_n$ は，たがいに独立，かつ，同一の分布に従う（independent and identically distributed : iid と略記）確率変数とする．なお，$E(X_i) = \mu < \infty \ (\forall i)$ を仮定する．このとき，任意の $\varepsilon > 0$ に対し，次式が成立する.

$$P\left\{\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| \le \varepsilon\right\} \to 1 \quad (n \to \infty)$$

（証明）X_i の ch.f を $c(t)$ とすると，$X_1 + \ldots + X_n$ の ch.f は $c(t)^n$ と書くことができる．さらに $\left(X_1 + \ldots + X_n\right) / n$ の ch.f は $c\left(\dfrac{t}{n}\right)^n$ で与えられる（5.2（P3）参照）．$c(t) = 1 + it\mu + o(t)$ と書けることから（5.2（P4）参照），

$$c\left(\frac{t}{n}\right)^n = \left(1 + \frac{it\mu}{n} + o\left(\frac{t}{n}\right)\right)^n$$

$$= \left[\left(1 + \frac{it\mu}{n} + o\left(\frac{t}{n}\right)\right)^{\frac{n}{it\mu}}\right]^{it\mu} \to \quad e^{it\mu} \quad (n \to \infty)$$

ここで，$e^{it\mu}$ は定数 μ の ch.f であることに注意し，Cdf と ch.f の一対一対応に注意すれば，$\left(X_1 + \ldots + X_n\right) / n$ の Cdf は, 定数値をとる確率変数 $X \equiv \mu$ の Cdf に収束することが示されたことになる．定数値確率変数 $X(\omega) \equiv \mu$ の Cdf を F とすると，

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

また，確率変数$\left(X_1+...+X_n\right)/n$の Cdf を$F_n$とすると，

$$P\left\{\left|\frac{X_1+...+X_n}{n}-\mu\right|\leq\varepsilon\right\}=P\left\{\frac{\sum X_i}{n}\leq\mu+\varepsilon\right\}-P\left\{\frac{\sum X_i}{n}<\mu-\varepsilon\right\}$$

$$\geq P\left\{\frac{\sum X_i}{n}\leq\mu+\varepsilon\right\}-P\left\{\frac{\sum X_i}{n}\leq\mu-\varepsilon\right\}$$

$$=F_n\left(\mu+\varepsilon\right)-F_n\left(\mu-\varepsilon\right)$$

$$\geq F_n\left(\mu+\varepsilon\right)-F_n\left(\mu-\frac{\varepsilon}{2}\right)\to F\left(\mu+\varepsilon\right)-F\left(\mu-\frac{\varepsilon}{2}\right)\qquad(n\to\infty)$$

$$=1-0=1$$

（注）大数の弱法則 Thm. 1 はつぎのように書いてもよい.

$$P\left\{\left|\frac{X_1+...+X_n}{n}-\mu\right|>\varepsilon\right\}\ \to\ 0\qquad(n\to\infty)$$

弱法則におけるX_iをベルヌーイ型確率変数（事象Aの生起を$X_i=1$，非生起を$X_i=0$とする）と解釈すれば，n回の独立反復試行における事象Aの生起・非生起を表す標本空間Ωは，つぎのような2^n個のωから成ると考えられる.

$$\omega_1=(0...0)\ ,\ \omega_2=(10...0)\ ,\ \omega_3=(010...0)\ ,\ ...,\ \omega_{2^n}=(1...1).$$

$E_n\triangleq\left\{\omega:\left|\frac{\sum X_i(\omega)}{n}-\mu\right|\leq\varepsilon\right\}$とするとき，弱法則の主張は，「任意の$\varepsilon>0$に対し，$P\{E_n\}\to1\ \ (n\to\infty)$」，すなわち，「あらかじめ定められた$n$に対して求まる相対度数$\sum x_i/n$は，$n$が十分大きくとってあれば，$P\{A\}$に十分近い値を与えている可能性が大きい」ということである．この可能性の大きさは$P\{E_n\}$で評価される.

しかし，このことは「延々と試行を続けていけば，やがて相対度数は$P\{A\}$に近づいていく」ということを意味しない．このことを述べるものは，大数の強法則（strong law of large numbers）である.

Thm. 2 （大数の強法則） $X_1, \ldots, X_n$ は iid であり， $E(X_i) = \mu < \infty \ (\forall i)$ とするとき

$$\frac{X_1 + \ldots + X_n}{n} \ \to \ \mu \qquad a.s.\ (n \to \infty)$$

（注） $a.s.$ は「$almost\ surely$」の略記である．

（注）上のことを， $P\left\{\frac{\sum X_i}{n} \to \mu\right\} = 1$ あるいは $\frac{\sum X_i(\omega)}{n} \ \to \ \mu \qquad w.p.\ 1$ とも書く．

強法則における標本空間 Ω は 0 か 1 の値をとる要素の無限長系列の集合である．例として，(00･･･)，(010･･･)，(0010･･･)，(11･･･) 等が含まれている． Ω は非可算集合であることに注意せよ．

強法則の主張をことばで述べるとつぎの通り．

「事象 A の生起・非生起を確認する独立反復試行を延々と続けることは，Ω に含まれるいずれかの無限系列に対応する． $n \to \infty$ のとき，無限系列における最初の n 個の要素の平均 $\left(\sum x_i \ / \ n\right)$ が $\mu \left(= P\{A\}\right)$ に収束しないような無限系列は，全ての無限系列の集合 Ω の中では極めて稀なものとして無視することができる」．

すなわち，

$$X_i = \begin{cases} 1 & i \text{ 回目の試行で事象} A \text{が観測される} \\ 0 & \text{それ以外} \end{cases}$$

（$X_1 + \ldots + X_n$ ：最初の n 回の試行中，事象 A の観測回数）

としたとき， n をどんどん大きくして $\sum x_i \ / \ n$ の更新を続けていけば，その値はほとんど確実に（almost surely） $P\{A\}$ に近づく．

5.3　中心極限定理

独立な確率変数の和は，その項数が大きいとき，個々の確率変数の分布に関わりなく，一定の分布に従うことがある．その一例が中心極限定理（central limit theorem）である．

Thm. （中心極限定理） $X_1, \ldots, X_n$ は iid な確率変数であり，

$E(X_i) = \mu < \infty, Var(X_i) = \sigma^2 < \infty \quad (\forall i)$ を満たすものとする．このとき，

$\dfrac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}$ の Cdf は，$\mathcal{N}(0, 1)$ の Cdf に収束する．

（証明） $X_i - E(X_i)$ の ch.f を $c(t)$ とすると，$t = 0$ の近傍で $c(t) = 1 - \dfrac{\sigma^2 t^2}{2} + o\left(t^2\right)$ と書ける．

$$\left(\sum X_i - n\mu\right) / \sigma\sqrt{n} = \sum \left(X_i - E(X_i)\right) / \sigma\sqrt{n}$$

の ch.f は，$c\left(\dfrac{t}{\sigma\sqrt{n}}\right)^n$ で与えられる．

$$c\left(\frac{t}{\sigma\sqrt{n}}\right)^n = \left[1 - \frac{\sigma^2}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + o\left(\left(\frac{t}{\sigma\sqrt{n}}\right)^2\right)\right]^n$$

$$= \left(1 - \frac{t^2}{2} + o\left(\frac{t^2}{n}\right)\right)^n \to e^{-\frac{t^2}{2}} \quad (n \to \infty)$$

ここで，$e^{-\frac{t^2}{2}}$ は $\mathcal{N}(0, 1)$ に従う確率変数の ch.f であることから，題意が証明された．

（注 1）この Thm から，$\dfrac{X_1 + \ldots + X_n}{n}$ の Cdf が $\mathcal{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ の Cdf に収束することがわかる．

これは，大数の弱法則の精密化であるということができる．すなわち，十分大きな n に対し

$$P\left\{\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| < \frac{\sigma}{\sqrt{n}}\,\varepsilon\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-\frac{x^2}{2}}\, dx$$

（注 2）中心極限定理は，統計学における推定ならびに検定において重要な役割を果たす．例えば，あるテレビ番組の視聴率を知る，あるいは現在の内閣の支持率を知ることは，統計学でいう「推定」に相当し，つぎのように表すことができる．

（例）TV ドラマ A の視聴率 μ を推定したい（A 内閣の支持率 μ を推定したい）．そこで，無作為に n 人を選択し，電話インタビューを行うことにした．その結果から，μ はどのように推定できるか？

（解）無作為に選ばれた n 人は，母集団から抽出された大きさ n の標本（$X_1,\ldots,X_n$）と考えることができる．ここで，各 X_k をつぎのように定義する．

$$X_k = \begin{cases} 1, & \text{k 番目の人は TV ドラマ A を見ている（A 内閣を支持している）} \\ 0, & \text{k 番目の人は TV ドラマ A を見ていない（A 内閣を支持していない）} \end{cases}$$

$X_1,\ldots,X_n$ は，独立なベルヌーイ型確率変数（$P\{X_k=1\}=\mu,\ P\{X_k=0\}=1-\mu$）と考えてよい．

さて，電話インタビューの結果に基づいて計算される $\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$ は，視聴率（あるいは支持率）μ の推定値となる．

例えば，A 内閣の支持率を推定するため，全国から 1,000 人を無作為に抽出して電話インタビューを行ったところ，340 人が「支持している」と答えたとする．このとき，A 内閣の支持率は 34%であると推定することができる．

ここで関心があるのは，おそらくつぎの（1），（2）であろう．

（1）支持率の推定値「34%」には，どれくらいの誤差が含まれていると考えるべきか．

（2）その誤差をある一定の値より小さくしたい（もし，「34%」に含まれる誤差が 3%であったとき，この誤差を 1%以内にしたい）なら，標本の大きさをどれくらいにする必要があるか．

これらの議論のために活躍するのが，「中心極限定理」である．詳しくは，3 年生向けに開講する「統計学」で述べる．