演習「統計的処理について」1

蛭川　潤一

新潟大学・自然科学系

2018 年 5 月 21 日

大数の法則

大数の法則（law of large numbers）:
概念的にデータと確率変数を結びつける重要な事実

X をゆがみのないサイコロの目とする
$\Rightarrow P(X=1)=\frac{1}{6}$

このサイコロを n 回振った時の 1 の目の度数を f_1 とする

$\Rightarrow$ 1 の目の相対頻度は $r_1=\frac{f_1}{n}$

Question ?

n が十分大きい時に $r_1 \approx \frac{1}{6}$ と近似できると考えてもよいであろうか？

相対頻度の極限が確率に一致するか?

大数の法則

Theorem (大数の（弱）法則（(weak) law of large numbers））

確率変数列 $(X), X_1, \ldots, X_n, \ldots$ が互いに独立に
いずれも平均 $E(X)=\mu$，分散 $V(X)=\sigma^2$ を持つ同一分布に従うとし
$\overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$ とする
$\Rightarrow$ 任意の正数 ε について

$$\lim_{n\to\infty} P_X\left\{\left|\overline{X}-\mu\right|>\varepsilon\right\}=0 \tag{1}$$

が成り立つ

(1) のような収束を
$\overline{X}$ が μ に確率収束する（convergence in probability）といい
$\overline{X} \xrightarrow{P} \mu \ (n\to\infty)$ と表す

相対頻度と確率

1 の目の定義関数:

$$\chi_{\{x=1\}}=\begin{cases}1 & \text{if} \quad x=1\\ 0 & \text{otherwise}\end{cases}$$

$Y=\chi_{\{X=1\}}$ とする
$\Rightarrow$

$$E(Y)=E\left(\chi_{\{X=1\}}\right)=1\times P(X=1)+0\times P(X\neq 1)=P(X=1)$$

$\Rightarrow$
確率 $P(X=1)$ は定義関数 $Y=\chi_{\{X=1\}}$ の期待値である

相対頻度と確率

$X_1, \dots, X_n$ を n 回サイコロを投げたときのそれぞれの目とする
$\Rightarrow$
1 の目の度数:

$$\begin{aligned} f_1 &= \chi_{\{X_1=1\}} + \cdots + \chi_{\{X_n=1\}} \\ &:= Y_1 + \cdots Y_n \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} r_1 &= \frac{f_1}{n} = \frac{1}{n}\{Y_1 + \cdots Y_n\} \\ &= \overline{Y} = \overline{\chi_{\{X=1\}}} \end{aligned}$$

$\Rightarrow$
相対度数 r_1 は $Y = \chi_{\{X=1\}}$ という確率変数の標本平均として表わされる

相対頻度と確率

Remark

相対頻度および確率は，標本平均および期待値の特殊ケースである
$\overline{Y} = \overline{\chi_{\{X=1\}}} = r_1$
$\mu = E(Y) = E\left(\chi_{\{X=1\}}\right) = P(X=1)$

$\Downarrow$ 大数の法則

任意の正数 ε について

$$\begin{aligned} &\lim_{n\to\infty} P_X\left\{\left|\overline{Y} - \mu\right| > \varepsilon\right\} \\ &= \lim_{n\to\infty} P_X\{|r_1 - P(X=1)| > \varepsilon\} = 0 \end{aligned}$$

$\Rightarrow r_1 \xrightarrow{P} P(X=1)\ (n \to \infty)$

大数の法則

Exercise

1 個の偏りのないサイコロを投げるという試行を独立に繰り返す

Excel によるシミュレーション（模擬実験）：

cf

「ファイル」–>「オプション」–>「アドイン」

管理：*Excel アドイン*「設定（*G*）」
分析ツールにチェック「*OK*」

大数の法則

Step (1)

乱数を発生させる

「データ」–>「データ分析」–>「乱数発生」

「変数の数」1
「乱数の数」1000
「分布」均一
「パラメータ」1 から 7

–>「OK」

大数の法則

Step (2)

（A 列）の乱数を整数にする

B1 <– 「=INT(A1)」

B2 から B1000 <– 「B1 をコピー」

Step (3)

（B 列の）1 の数を数える

C1 <– 「=COUNTIF($B$1:$B1,1)」

C2 から C1000 <– 「C1 をコピー」

大数の法則

Step (4)

（B 列の）1 の数の相対頻度を求める

D1 から D1000 <– 「1 から 1000 を代入」

E1 <– 「=$C1/$D1」

E2 から E1000 <– 「E1 をコピー」

Step (5)

$\frac{1}{6}$ と比べる

F1 から F1000 <– 「$\frac{1}{6}$ を代入」

E 列と F 列を選択して
「挿入」 –>「グラフ」 –>「折れ線」 –>

大数の法則

まとめてやってみる（Step(1),(2) は同じ）

Step

（B 列の）$1-6$ の数の相対頻度を求めて，$\frac{1}{6}$ と比べる

C1 から C1000 <– 「1 から 1000 を代入」

D1 <– 「=COUNTIF($B$1:$B1,1)/$C1」
E1 <– 「=COUNTIF($B$1:$B1,2)/$C1」
...
I1 <– 「=COUNTIF($B$1:$B1,6)/$C1」
J1<– 「=1/6」

D1 から J1 を選択し，フィルハンドルをダブルクリック
D 列から J 列を選択して
「挿入」 –>「グラフ」 –>「折れ線」 –>

確率の性質

次の性質 (P1)〜(P3) を満たす $P(\cdot)$ を確率という

Definition

(P1) $0 \leq P(E) \leq 1$
(P2) $P(\Omega) = 1$
(P3) $E_1, E_2, \ldots$ が互いに素ならば

$$P\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

(P3) について
$E_1, E_2, \ldots$ が互いに素 $\Leftrightarrow$ 任意の i, j $(i \neq j)$ について E_i と E_j が互いに素

E_i と E_j が互いに素 $\Leftrightarrow E_i \cap E_j = \emptyset \Leftrightarrow E_i$ と E_j に共通の要素がない

幾何学的確率

平面 $\mathbb{R}^2$ における有界領域 Ω が与えられてるとする
Ω の部分集合 E に対して

$$P(E) = \frac{|E|}{|\Omega|} \quad (|\cdot| \text{ は集合の面積を表す})$$

とおく
さらに、互いに素な Ω の部分集合 $E_1, E_2, \ldots$ に対し

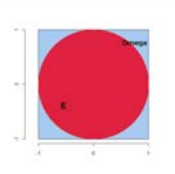
$$P\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

と定める

$\Rightarrow$ $P(\cdot)$ は性質 (P1)〜(P3) を満たす
$\Rightarrow$ $P(\cdot)$ は確率である

幾何学的確率

例：$\Omega = [-1,1] \times [-1,1]$, $E = \{(x,y) \mid x^2 + y^2 \leq 1\}$

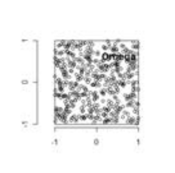


$$p := P(E) = \frac{|E|}{|\Omega|} = \frac{\pi}{4}$$

幾何学的確率

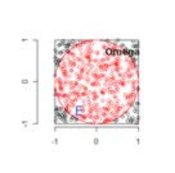
$P(E)$ は $\Omega = [-1,1] \times [-1,1]$ 上の一様分布が
E に入る確率になる

$\Omega = [-1,1] \times [-1,1]$ 上の一様分布からの 500 個の標本



幾何学的確率

$\Omega = [-1,1] \times [-1,1]$ 上の一様分布からの 500 個の標本の内
E に含まれる標本の割合



$$\hat{p} = \frac{E \text{ に含まれる標本の数}}{\text{全標本数}} = \frac{396}{500}$$

幾何学的確率

$$\hat{p} = \frac{E \text{ に含まれる標本の数}}{\text{全標本数}} \quad \text{は}$$

$$p := P(E) = \frac{|E|}{|\Omega|} = \frac{\pi}{4}$$

の推定量になる

$$\Rightarrow \quad \hat{\pi} = 4\hat{p} \quad \text{は } \pi \text{ の推定量}$$

幾何学的確率

標本数 500：

$$\hat{\pi} - \pi = 4\hat{p} - \pi = 4 \times \frac{396}{500} - \pi = 3.168 - \pi = 0.02640735$$

標本数 5000：

$$\hat{\pi} - \pi = 4\hat{p} - \pi = 4 \times \frac{3943}{5000} - \pi = 3.1544 - \pi = 0.01280735$$

標本数 50000：

$$\hat{\pi} - \pi = 4\hat{p} - \pi = 4 \times \frac{39216}{50000} - \pi = 3.13728 - \pi = -0.004312654$$