

乱数とヒストグラム

モデリングとシミュレーション

1

2020年度

乱数(Random Numbers)

- ▶ でたらめに生成される数の列
- ▶ 例：サイコロの目の列
 - ▶ 出る順番がでたらめ
 - ▶ 次が予測できない
 - ▶ 出る頻度がほぼ均等

乱数とシミュレーション

確率過程

- ▶ ある確率 p で事象 E が起こる
 - ▶ 次の動作が決定的でない
- ▶ 動作が決定的でない場合
 - ▶ 対象の外からの予測不能な影響
 - ▶ 内部に不確かな要素
 - ▶ 構成要素の予測不能な挙動
 - ▶ 本質的に確率的現象
 - ▶ 熱
 - ▶ 量子力学

確率の意味

ある確率 p で事象 E が起こる

- ▶ 試行を十分大きな回数 N 回行う
- ▶ 事象 E が起こった回数 N_E

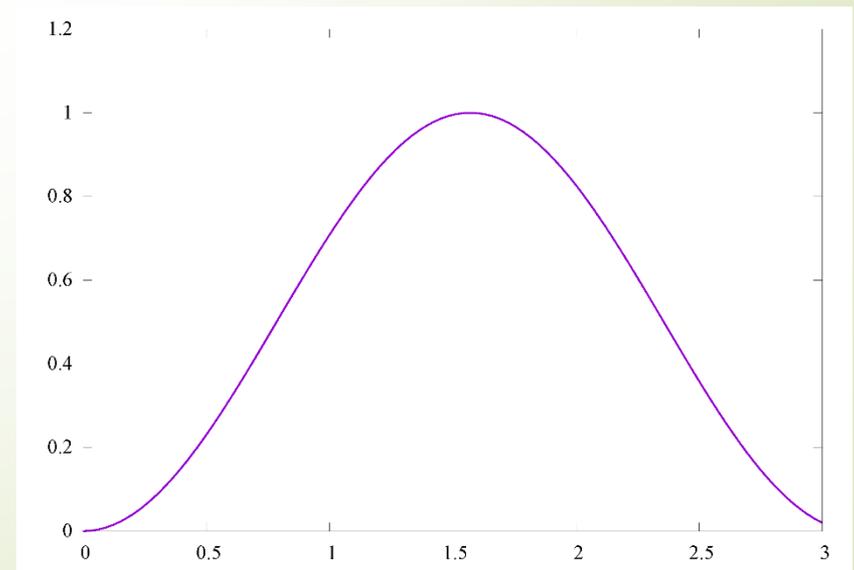
$$\frac{N_E}{N} \sim p \quad \text{大数の法則}$$

- ▶ 例：さいころの1の目が出る確率 $1/6$
 - ▶ 十分大きな回数 N 回、さいころを投げる
 - ▶ 1の目がでる回数は、 $N/6$ に近づく
 - ▶ 「近づく」の意味

乱数とシミュレーション

Monte Carlo 法

- ➡ 乱数を積極的に活用して、近似を行う
- ➡ 積分値の近似
 - ➡ ある曲線で囲まれた領域に二次元の乱数が入る確率



疑似乱数 (Pseudo Random Numbers)

■ 本物の乱数

- さいころ

- 熱雑音

- 放射性原子の崩壊

■ 疑似乱数

- コンピュータで生成

- アルゴリズムがあるので「偽物」

乱数の分布を調べる ヒストグラム (histogram)

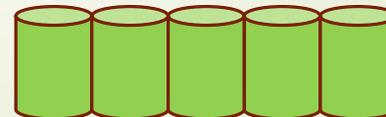
- ▶ ある事象が何回起こったか、その頻度
あるいは相対頻度
- ▶ 0から $N - 1$ までの乱数
 - ▶ $N = 2^{31} \cong 2 \times 10^9$ の場合
 - ▶ それぞれの整数は何回発生する？
 - ▶ そこから何が分かる？

binを使う

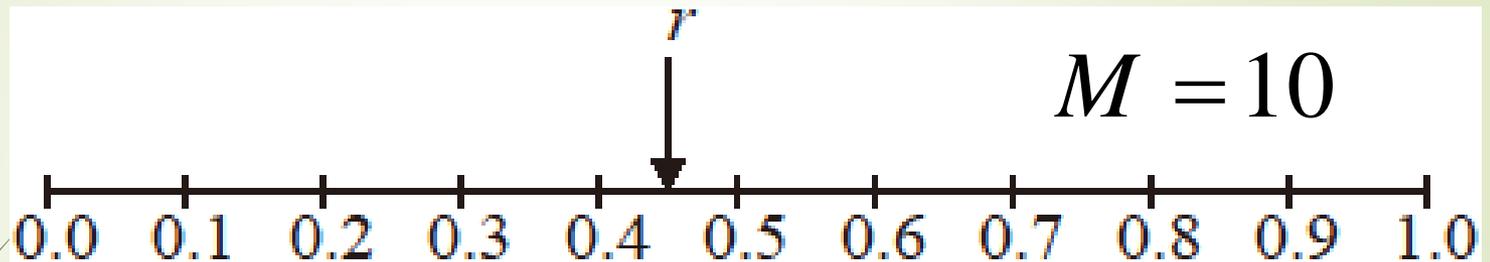
bin:
a container that you put waste in
a large container, usually with a
lid, for storing things in

➡ $[0,1)$ の範囲の乱数

- ➡ 各値の発生回数を数えるのは無意味
- ➡ 区間 $[0,1)$ を M 個に区切る。
 - ➡ 各区間をbinと呼ぶ
- ➡ 各binに何個入ったかの頻度または相対頻度を求める
- ➡ binに入る標本数が大きくなると無意味



bin



- r はどの区間に入っている？
- 区間幅を w とすると

$$k = \lfloor r/w \rfloor$$

$$w = 1/M$$

■ 課題：一般化

- 区間 $[a, b)$ を M 個に分割した場合
 - ある数 r が入る区間を判定する方法は？
- その方法で、 $[-0.1, 0.9)$ を5個に等分した区間に対して、 $\{-0.02, 0.04, 0.22, 0.57\}$ はそれぞれどの区間に入ると判定されるか？

理想的一様乱数

- ➡ $[0,1)$ の区間に一様乱数が発生したとする
- ➡ 乱数の総数は N 、区間数は M とする
- ➡ ある区間に
 - ➡ 一つの乱数が入る確率 $p = 1/M$
 - ➡ N 個の乱数のうち、区間に k 個が入る確率

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k}$$

二項係数がでてくる理由は？

平均と分散を求めるには

$$\begin{aligned}
 \langle k \rangle &= \sum_{k=0}^N k P(k) = \sum_{k=0}^N k \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k} \\
 &= \sum_{k=1}^N \frac{N(N-1)!}{(k-1)!(N-k)!} p p^{k-1} (1-p)^{N-k} \\
 &= Np \sum_{k=0}^{N-1} \frac{(N-1)!}{k!(N-1-k)!} p^{k-1} (1-p)^{N-k} = Np
 \end{aligned}$$

$$\langle k^2 \rangle = \langle k(k-1) \rangle + \langle k \rangle = N(N-1)p^2 + Np = Np((N-1)p + 1)$$

$$\begin{aligned}
 \langle k(k-1) \rangle &= \sum_{k=0}^N k(k-1) \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k} \\
 &= \sum_{k=2}^N \frac{N(N-1)(N-2)!}{(k-2)!(N-k)!} p^2 p^{k-2} (1-p)^{N-k} \\
 &= N(N-1)p^2
 \end{aligned}$$

$$\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2 = Np((N-1)p + 1) - N^2 p^2 = Np(1-p)$$

もっと効率よく計算したい

確率母関数

Probability generating function

$$G(z) = \sum_{k=0}^N P(k) z^k$$

和の範囲に注意

$$G(1) = \sum_{k=0}^N P(k) = 1$$

$$G'(z) = \sum_{k=1}^N kP(k) z^{k-1}$$

母関数の具体的な形が分かると、平均や二乗平均が容易に求められる

$$G'(1) = \sum_{k=1}^N kP(k) = \sum_{k=0}^N kP(k) = \langle k \rangle$$

$$G''(z) = \sum_{k=2}^N k(k-1)P(k) z^{k-2}$$

$$G''(1) = \sum_{k=2}^N k(k-1)P(k) = \sum_{k=0}^N k(k-1)P(k) = \langle k^2 \rangle - \langle k \rangle$$

二項分布に対する確率母関数

$$\begin{aligned} G(z) &= \sum_{k=0}^N P(k) z^k = \sum_{k=0}^N \binom{N}{k} p^k (1-p)^{N-k} z^k \\ &= \sum_{k=0}^N \binom{N}{k} (zp)^k (1-p)^{N-k} = (zp + 1 - p)^N \end{aligned}$$

$G(z)$ の具体的関数形が分かると微分が具体的に計算可能

$$G'(z) = Np(zp + 1 - p)^{N-1}$$

$$G''(z) = N(N-1)p^2(zp + 1 - p)^{N-2}$$

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

二項分布に対する平均と標準偏差

$$G'(1) = Np = \langle k \rangle$$

重要： $\langle k \rangle \propto N$

$$G''(1) = N(N-1)p^2 = \langle k^2 \rangle - \langle k \rangle$$

$$\langle k^2 \rangle = N(N-1)p^2 + Np$$

重要： $\sigma^2 \propto N$

$$\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2 = N(N-1)p^2 + Np - N^2p^2 = Np(1-p)$$

$$\frac{\sigma}{\langle k \rangle} = \left(\frac{1-p}{p} \frac{1}{N} \right)^{1/2}$$

重要： $\sigma/\langle k \rangle \propto N^{-1/2}$

連続的な確率変数と確率分布

- 確率変数が実数の場合
 - 実数は連続であることに注意
 - 「ある値になる確率」はあり得ない
- 確率変数 $X \in [a, b)$ に対する確率分布
 - $a \leq X < x$ という「範囲」に対して確率を定義

$$F(x) = P(a \leq X < x)$$

確率密度

- ➡ 微小区間($X \in [x, x + \Delta x)$)に入る確率

$$f(x)\Delta x = \frac{d}{dx} F(x)\Delta x$$

$$F(x) = \int_a^x f(y)dy$$

- ➡ $f(x)$ を確率密度と呼ぶ
 - ➡ 確率そのものでないことに注意
 - ➡ 区間幅を乗じて確率になる

相対頻度と確率密度

- ➡ 区間 $[a, b)$ に N 個の乱数、 M 個のbinに分割
- ➡ 各区間の頻度 F_i 、相対頻度 $f_i = \frac{F_i}{N}$
- ➡ 相対頻度は確率分布として理解できる？
 - ➡ 離散的確率と考えると、規格化は

$$\sum_{i=0}^{M-1} f_i = \frac{1}{N} \sum_{i=0}^{M-1} F_i = 1$$

- ▶ 連続的確率密度と考えると

$$\int_a^b p(x)dx = 1$$

- ▶ 積分は面積→小さな台形の和
- ▶ 対応する和は、binの幅 $w = (b - a)/M$ として

$$\sum_{i=0}^{M-1} w \frac{f_i}{w} = 1$$

$$\frac{f_i}{w} \rightarrow p(x_i), x_i = a + \left(i + \frac{1}{2}\right)w$$