確　率　と　乱　数

杉　田　　洋

大阪大学大学院理学研究科

- **確率論の目的**
- 乱数とは
- 硬貨投げはランダムか？
- 極限定理
- モンテカルロ法
- ひとこと…

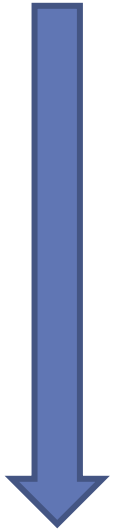
確率論の目的

１．ランダム性の解析（デタラメ、不規則）

２．公平性（ゲーム、経済）

３．予測（天気予報、リスク管理）

４．……



何を調べればよいのか？

「確率」の原意

中国語

Probability = 確率 = 十中八九確実なこと = 「概率」

Probably = たぶん = 十中八九

Probably > likely > maybe ≒ perhaps > possibly

||

Almost surely

Probability theory = 確率論

= 十中八九確実なことに関する理論

確率論の第一の目的

ランダム性の解析

∴

十中八九確実なことを詳しく調べること

？

- 確率論の目的
- **乱数とは**
- 硬貨投げはランダムか？
- 極限定理
- モンテカルロ法
- ひとこと…

「ランダムである」とはどういうことか？

ランダムである  規則的である

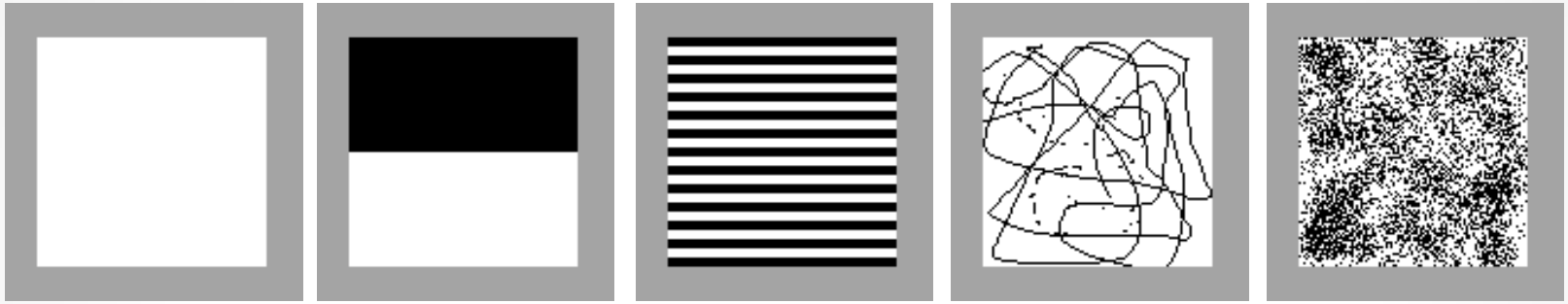
||

最も規則的でない

「規則的である」とはどういうことか？

規則的 ⇒ 圧縮可能

- 画像の圧縮 (zip)



100×100 画素：40054 byte

583 byte　589 byte　666 byte　1853 byte　2967 byte

ランダム ⇒ 圧縮されにくい

規則的 ⇒ 圧縮可能

- 円周率 π の小数点以下5兆桁

3.1415926535 … 9484283852

圧縮

プログラム

定義

どのような方法でもほとんど圧縮することができないとき**ランダム**という．

抽象化、普遍化、…

数、文章、音声、画像、映像、動き、…

デジタル符号化

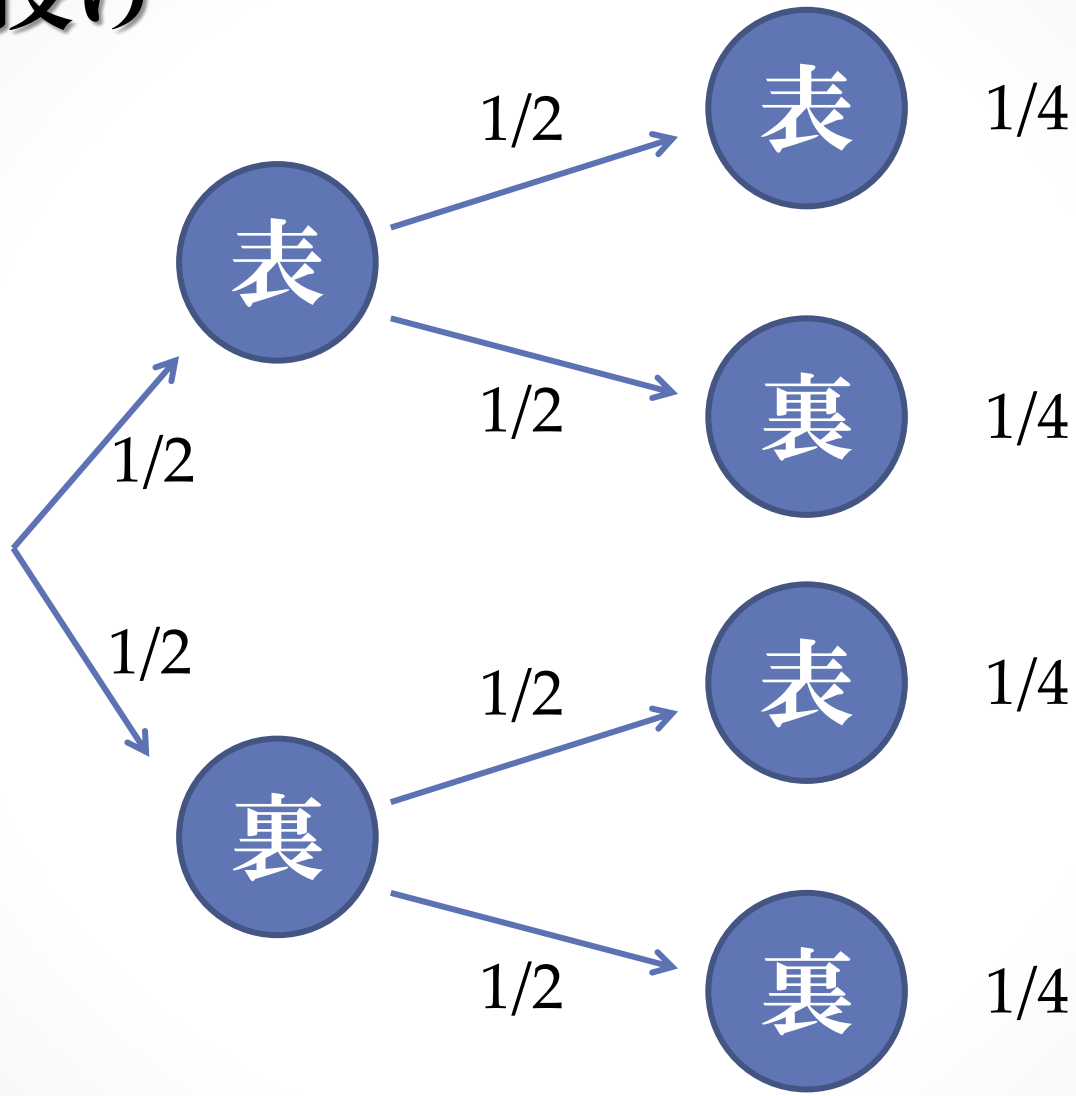
{0,1}-列

定義

どのような方法でもほとんど圧縮することができない長い {0,1}-列を**乱数**という．

- 確率論の目的
- 乱数とは
- **硬貨投げはランダムか？**
- 極限定理
- モンテカルロ法
- ひとこと…

硬貨投げ



硬貨投げ

= 1

= 0

n

2^n

			確率
00…000	$1/2 \times \ldots \times 1/2$	=	$1/2^n$
00…001	$1/2 \times \ldots \times 1/2$	=	$1/2^n$
……			
11…110	$1/2 \times \ldots \times 1/2$	=	$1/2^n$
11…111	$1/2 \times \ldots \times 1/2$	=	$1/2^n$

硬貨投げでできる{0,1}-列は乱数か？

例：　0 0 … 0 0 0 1 1 … 1 1 1

5000　　5000

圧縮できるので乱数ではない．

硬貨投げでできる{0,1}-列は乱数でないこともある！

長さ 10000 の {0,1}-列のうち圧縮すると
長さ 9990 以下の {0,1}-列になるものの
個数は？

高々　$2^1 + 2^2 + \cdots + 2^{9990} = 2^{9991} - 2$

∵
- 長さ k の {0,1}-列の総数 $= 2^k$ 個
- 長さ 10000 の {0,1}-列のうち圧縮すると長さ k の {0,1}-列になるようなものの個数は高々 2^k 個

∵
- $2^1 + 2^2 + \cdots + 2^{9990} = x$ とおくと
- $2x = 2^2 + 2^3 + \cdots + 2^{9991} = x + 2^{9991} - 2^1$
- 両辺から x を引いて $x = 2^{9991} - 2$ を得る．

10000 回の硬貨投げでできる {0,1}-列が圧縮すると長さ 9990 以下の {0,1}-列になる確率は？

- 長さ 10000 の {0,1}-列の総数 = 2^{10000} 個
- 長さ 10000 の {0,1}-列のうち圧縮すると長さ 9990 以下の {0,1}-列になるようなものの個数は高々

$$2^{9991} - 2$$

したがって求める確率は高々

$$\frac{2^{9991}-2}{2^{10000}} < \frac{2^{9991}}{2^{10000}} = \frac{1}{2^9} = \frac{1}{512}$$

10000 回の硬貨投げでできる {0,1}-列が長さ 9990 以下の {0,1}-列に圧縮できない確率は？

$$1 - \frac{1}{512} = \frac{511}{512} = 0.998 \text{ 以上}$$

硬貨投げでできる長い{0,1}-列は、**非常に高い確率**で乱数である．

乱数：まとめ

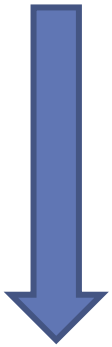
- どのような方法でもほとんど圧縮できない長い {0,1}-列を乱数という．
- 長い {0,1}-列の圧倒的大多数は乱数である．
- 硬貨投げの結果で得られる長い {0,1}-列は非常に高い確率で乱数である．
- 乱数はコンピュータでも生成することができない．
- 乱数は「乱数であること」を証明することができない．

- 確率論の目的
- 乱数とは
- 硬貨投げはランダムか？
- **極限定理**
- モンテカルロ法
- ひとこと…

確率論の目的

１．ランダム性の解析（デタラメ、不規則）

２．公平性（ゲーム、金融、経済）
３．予測（天気予報、リスク管理）
４．……



何を調べればよいのか？

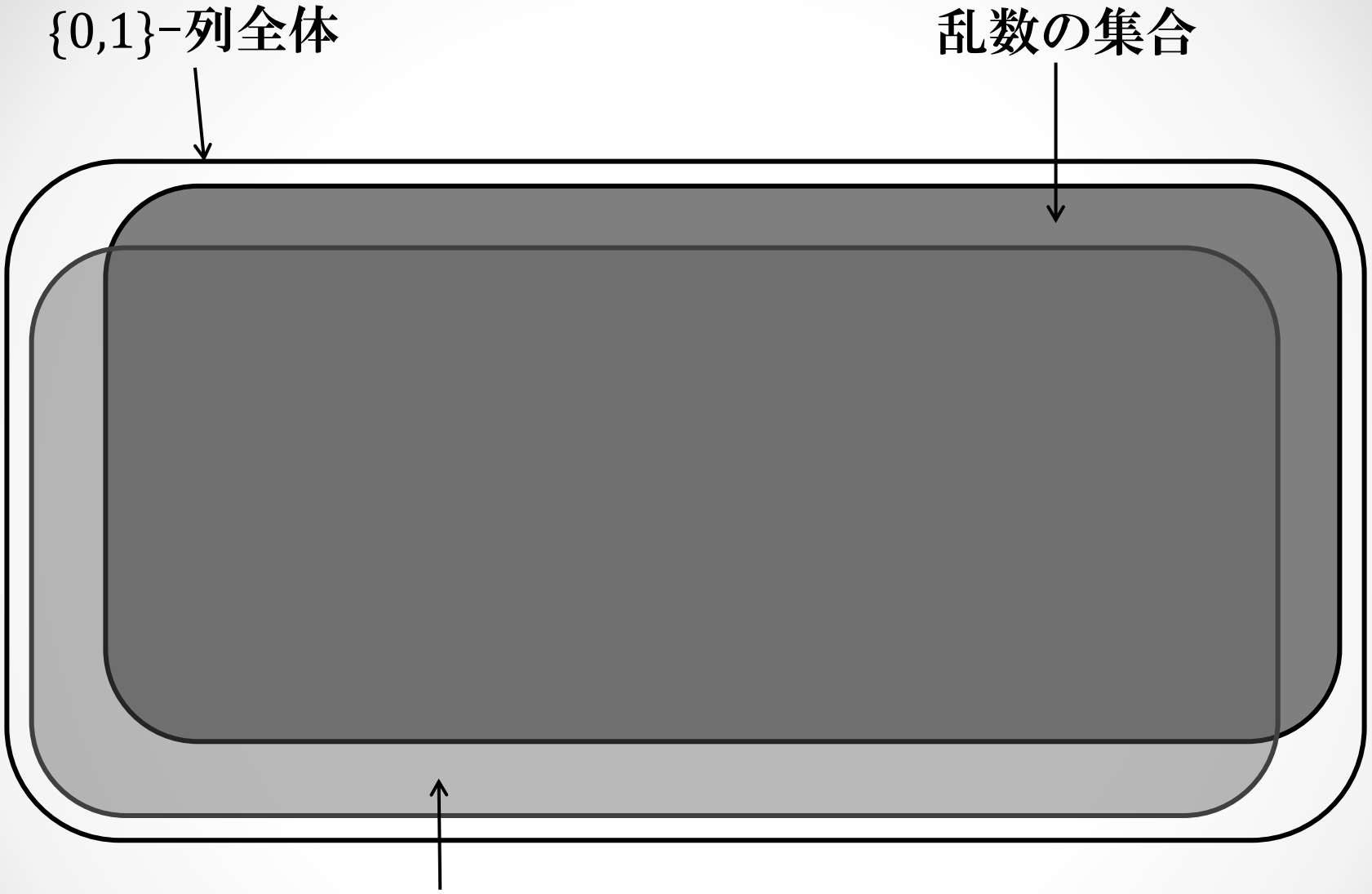
乱数の性質を調べればよい．

乱数の性質はどうやって調べるか？

- 硬貨投げでできる長い{0,1}-列は、非常に高い確率で乱数である．

だから…

- 硬貨投げでできる長い{0,1}-列に関する事象で非常に高い確率で起こるものについて調べれば、ほぼ乱数の性質が分かる．


{0,1}-列全体
乱数の集合
非常に高い確率とされる{0,1}-列の集合 = 十中八九確実なこと

確率論の第一の目的

ランダム性の解析

=

乱数の性質の解析

∴

十中八九確実なことを詳しく調べること

=

極限定理

ベルヌーイの定理 — 大数の法則

実験：100回硬貨を投げて表と裏の出た回数を数える．

1110110101 1011101101 0100000011 0110101001 0101000100
0101111101 1010000000 1010100011 0100011001 1101111101

表(=1)の出た回数 = 51

硬貨を投げる回数を増やすと、表の出る回数の比率が 1/2 に近づく．

- 100回硬貨投げるとき表が k 回出る場合の数は

$$ {}_{100}\mathrm{C}_k = \frac{100!}{(100-k)!\,k!} \qquad k = 0,1,\ldots,100 $$

- 表の出る回数が 50 ± 15 以内である場合の数は

$$ \sum_{k=35}^{65} {}_{100}C_k = 126538159389309434317301668 2006 $$

- 表裏の出方の総数は

$$ 2^{100} = 1267650600228229401496703205376 $$

- 表の出る回数が 50 ± 15 以内である確率は

$$ \frac{126538159389309434317301668 2006}{1267650600228229401496703205376} = 0.9982100696 $$

- 1000回硬貨投げるとき表の出る回数が 500 ± 150以内である確率は

$$\frac{\sum_{k=350}^{650} {}_{1000}\mathrm{C}_k}{2^{1000}} = 0.99999999999999999999135703693$$

- 1000回硬貨投げるとき表の出る回数が 500 ± 50以内である確率は

$$\frac{\sum_{k=450}^{550} {}_{1000}\mathrm{C}_k}{2^{1000}} = 0.99860825840557791377674858896$$

- 10000回硬貨投げるとき表の出る回数が 5000 ± 150以内である確率は

$$\frac{\sum_{k=4850}^{5150} {}_{10000}\mathrm{C}_k}{2^{1000}} = 0.99869463104666066521012366458$$

まとめると…

投げる回数	表の出る回数の区間	区間の幅 / 投げる回数	確率
100	50±15	0.30	99.8%
1000	500±50	0.10	99.9%
10000	5000±150	0.03	99.9%

長さ 10000 の乱数は「表の出る回数が 5000 ± 150以内である」という性質を持つ.

一般の硬貨投げ：表の確率 $= p$，裏の確率 $= 1 - p$

定理（ベルヌーイ）

どんな小さな $\varepsilon > 0$ に対しても、硬貨を投げる回数 n を大きくすると

$$\frac{\text{表の出る回数}}{n}$$

が区間 $p \pm \varepsilon$ に入る確率が 1 に収束する．

- 確率論の目的
- 乱数とは
- 硬貨投げはランダムか？
- 極限定理
- **モンテカルロ法**
- ひとこと…

例題

硬貨を100回投げるとき、表が続けて 6 回以上出る確率 p を求めよ.

解法

- 「硬貨を100回投げる」を 1000000 回繰り返す.
- そのうちで「表が続けて 6 回以上出る」が起こった回数を S とする.
- $S/1000000$ を p の推定値とする.

S は「表の出る確率が p の硬貨を 1000000 回投げたとき表が出る回数」と思うことができるから、

ベルヌーイの定理（チェビシェフ不等式）

$$P\left(\left|\frac{S}{1000000}-p\right|\leq\frac{1}{200}\right)\geq\frac{99}{100}$$

- S/1000000 と p の差が 1/200 以下になる確率は 99%以上である.

実際の手続き

- 硬貨を投げる回数は $100 \times 1000000 = 10^8$.
- 本物の硬貨を投げるのは実行不可能.
- コンピュータで生成する「疑似乱数」を用いて計算する.

デモンストレーションをご覧下さい。

疑似乱数 (RWS法)

例題を解くための疑似乱数生成器

g : 長さ 238 の {0,1}−列 → 長さ 10^8の {0,1}−列

ω : 長さ 238 の {0,1}-列 (**種**)

$g(\omega)$:長さ 10^8の {0,1}−列 (**疑似乱数**)

より $S(g(\omega))/1000000$ を計算して p を推定する.

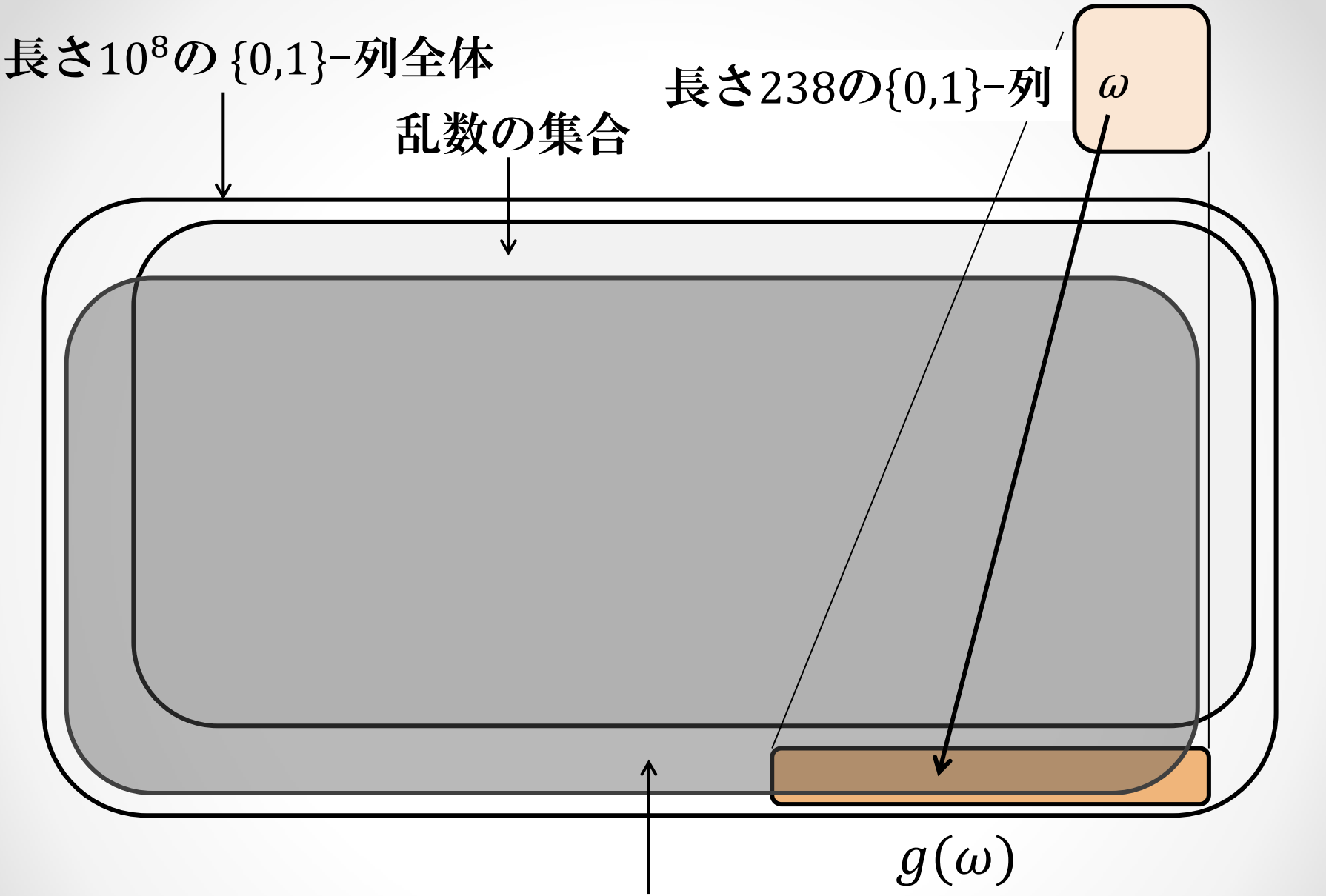
ンをご覧下さい。

定理（RWS 法）

$$P\left(\left|\frac{S(g(\omega))}{1000000} - \boldsymbol{p}\right| \leq \frac{1}{200}\right) \geq \frac{99}{100}$$

- $S(g(\omega))/1000000$ と $\boldsymbol{p}$ の差が 1/200 以下になる確率は 99%以上である．

RWS 法を用いれば、10^8回の硬貨投げの代わりに 238 回の硬貨投げで済む．


長さ10^8の {0,1}−列全体
乱数の集合
長さ238の{0,1}−列
ω
$g(\omega)$
チェビシェフ不等式によって99% 以上とされる{0,1}−列の集合

RWS法 $g:\{0,1\}^{238}\to\{0,1\}^{10^8}$ の定義

- 種を半分に分ける：$\omega=(x,\alpha)\in\{0,1\}^{119}\times\{0,1\}^{119}$
- $x=(x_1,\cdots,x_{119})$ および $\alpha=(\alpha_1,\cdots,\alpha_{119})$ から 次を定義する：

$$\tilde{x}:=\sum_{i=1}^{119}x_i2^{-i}\qquad\tilde{\alpha}:=\sum_{i=1}^{119}\alpha_i2^{-i}$$

- $\tilde{z}_k:=\tilde{x}+k\tilde{\alpha}\ (\mathrm{mod}\ 1),\quad k=1,\cdots,10^6$
- $\tilde{z}_k$ の2進数表記の小数第 i 桁目を $d_i(\tilde{z}_k)\in\{0,1\}$ と表すとき

$$g_k:=(d_1(\tilde{z}_k),\cdots,d_{100}(\tilde{z}_k))\in\{0,1\}^{100},\quad k=1,\cdots,10^6$$

- $g_k=g_k(\omega)$ を並べて

$$g(\omega):=(g_1(\omega),\cdots,g_{10^6}(\omega)),\qquad\omega\in\{0,1\}^{238}$$

- 確率論の目的
- 乱数とは
- 硬貨投げはランダムか？
- 極限定理
- モンテカルロ法
- ひとこと…

年譜

確率論

- 1560年　カルダノ「サイコロあそびについて」
- 1713年　ベルヌーイ「推測法」
- 1820年　ラプラス「確率の解析的理論」
- 1933年　コルモゴロフ「確率論の基礎概念」

乱数

- 1919年　　ミーゼス
- 1960年代　コルモゴロフ、チャイティン、ソロモノフ

疑似乱数

- 1940年代　ノイマン　「モンテカルロ法」
- 1980年代　ブラム、ヤオ　「暗号理論」
- 2002年　　「RWS法」

- 応用数学とは
- 不可能から可能へ
- 奇想天外 vs 荒唐無稽

宿題：「数学=Mathematics」の原意は？

ありがとうございました！

をご覧下さい。