統数研共同研究集会「経済物理学とその周辺」
H24年度第一回研究会

RMTテストの性能検証
～NIST乱数検定との比較～

2012年8月27日～28日

鳥取大学大学院工学研究科情報エレクトロニクス専攻
三賀森悠大　楊欣　糸井良太　田中美栄子

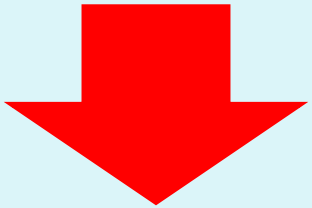
目次

はじめに　～乱数度とは～

乱数列：無作為に並べられた数字の羅列

73	50	1	8	81	10	77	33	82	76
95	97	84	38	54	32	55	95	89	80
55	39	94	10	29	24	・・・			

数の並びが如何に**ランダム**かという基準



乱数度

はじめに　～乱数度の定義～

並び方が無作為であり、特徴をつかみにくい

→ 乱数度が**高い**

並び方に規則性があり、特徴をつかみやすい

→ 乱数度が**低い**

はじめに　～乱数検定～

乱数検定ツール

NISTで推奨されたツール

プログラムの実行により
複数の検定を
まとめて行える

はじめに　～乱数検定～

乱数検定ツール

検定するデータ形式の制限

データ形式指定の必要あり

推奨データ長

データ長に応じた信頼性の問題

複数の基準の併用が必要

検定に時間がかかる

はじめに　～乱数検定～

乱数検定ツール

検定するデータ形式の制限

データ形式を選ばない

推奨データ長

データ長を自由に決められる

複数の基準の併用が必要

一つのプログラムで検定
数値のみで評価可能

前述の問題点を克服する為の乱数度測定器として・・・

RMTテスト

はじめに　～先行研究～

RMTテストとは？

※楊、田中による提案

RMT公式を用いた
乱数度評価法

RMT：Random Matrix Theory
(ランダム行列理論)

はじめに　～先行研究～

RMT公式

乱数列から作成した行列を
自己相関行列に変換

↓

RMTの理論式と
自己相関行列の固有値分布を比較

↓

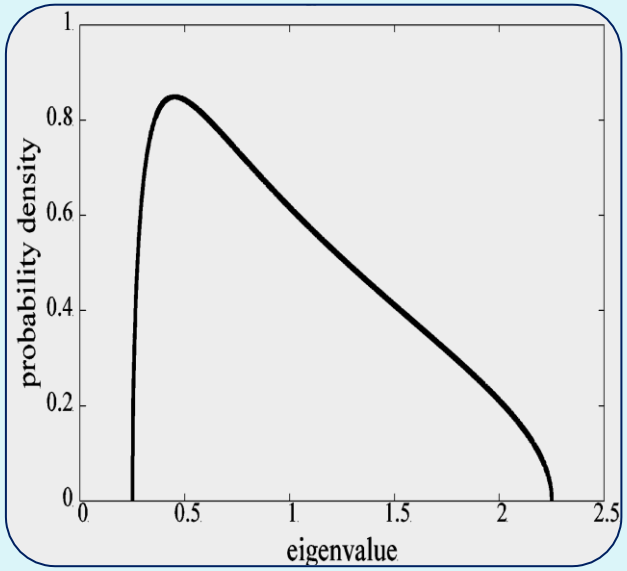
ランダム部分と有意な部分に分別

はじめに　～先行研究～

$$Q = \frac{L}{N} \qquad (\mathrm{L} \to \infty,\ \ N \to \infty)$$

$$\lambda_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}$$

$$P_{RMT}(\lambda) = \frac{Q}{2\pi\lambda}\sqrt{(\lambda_{+} - \lambda)(\lambda - \lambda_{-})}$$



L：データ長　　λ_{+}：最大固有値

N：サンプル数　　λ_{-}：最小固有値

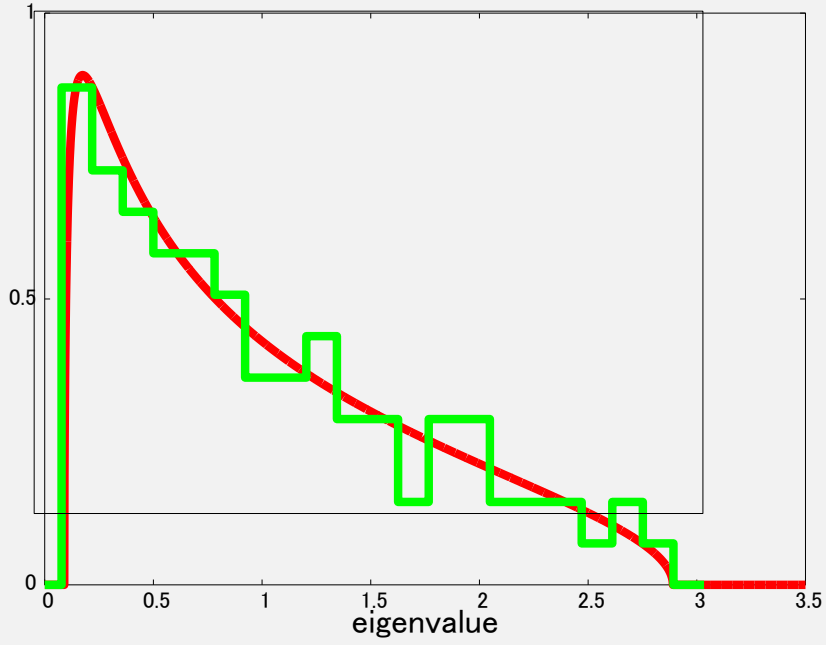
相関行列の固有値分布はQのみに依存

はじめに　～先行研究～

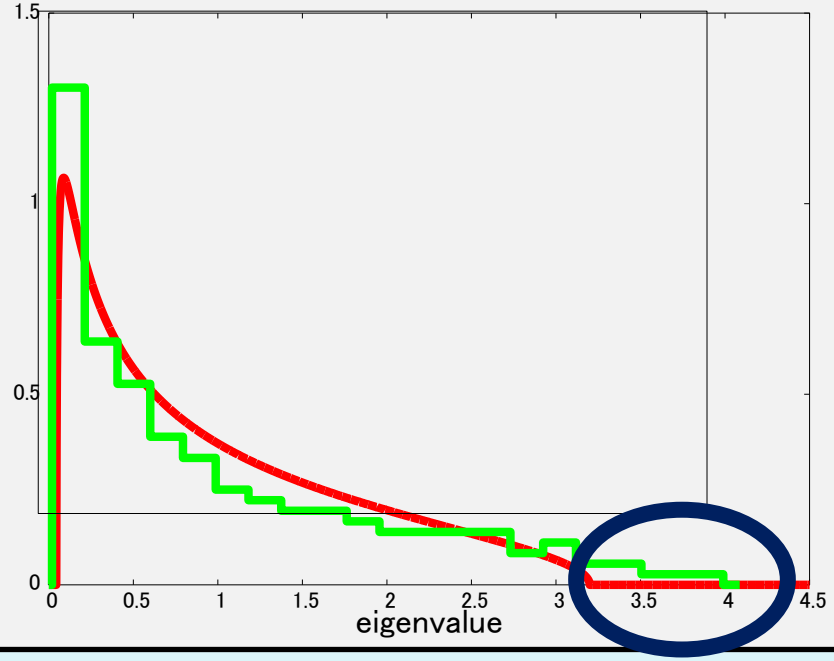
これを応用して・・・

固有値分布を
理論式と比較

乱数度が高い例



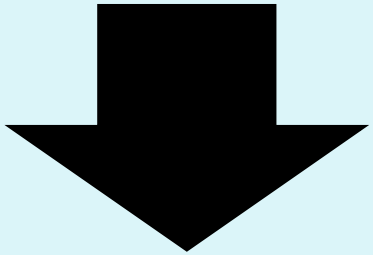
乱数度が低い例



はじめに

数値化するために・・・

理論式と自己相関行列の固有値分布のずれ具合



モーメント法により

誤差(%)で**定量化**

研究目的

暗号でよく使われる
NISTの検定ツールを用いて
その結果をRMTテストの結果と
対比させると

どの程度の誤差の値で
良い乱数と見なされるのか

研究目的

他手法との比較から

RMTテストの

基準を詳しく調査

研究目的

先行研究

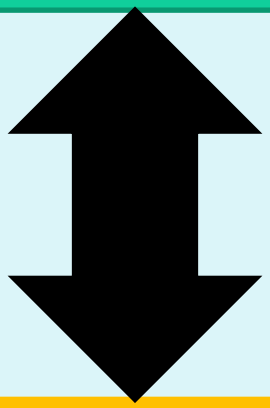
理論式と固有値分布の
照合による判断

本研究

別の検定法による
結果との照合で判断

研究目的　～手法～

RMTテスト



比較

NIST乱数検定

研究目的　～NIST乱数検定～

以下の15種類を採用

- 1次元度数検定
- ブロック単位の頻度検定
- 累積和検定
- 連の検定
- ブロック単位の最長連検定
- 2値行列ランク検定
- 離散フーリエ変換検定
- 線形複雑度検定
- 系列検定
- 重なりの無いテンプレート適合検定
- 重なりのあるテンプレート適合検定
- Maurerのユニバーサル統計検定
- 近似エントロピー検定
- ランダム偏差検定
- 種々のランダム偏差検定

RMTテストに向けた数値処理

乱数データから行列A作成

正規化(行列Gの作成)

自己相関行列Cの作成

乱数列処理

Cの対角要素の平均算出

理論値との誤差算出

数値処理

RMTテストに向けた数値処理

乱数データから行列A作成

乱数データ

$$A_1 \quad A_2 \quad A_3 \quad \cdots \quad A_{NL}$$

長さLの乱数列をN個用意

長さLずつ区切る

RMTテストに向けた数値処理

乱数データから行列A作成

乱数データ

乱数列１　　乱数列２　　乱数列N

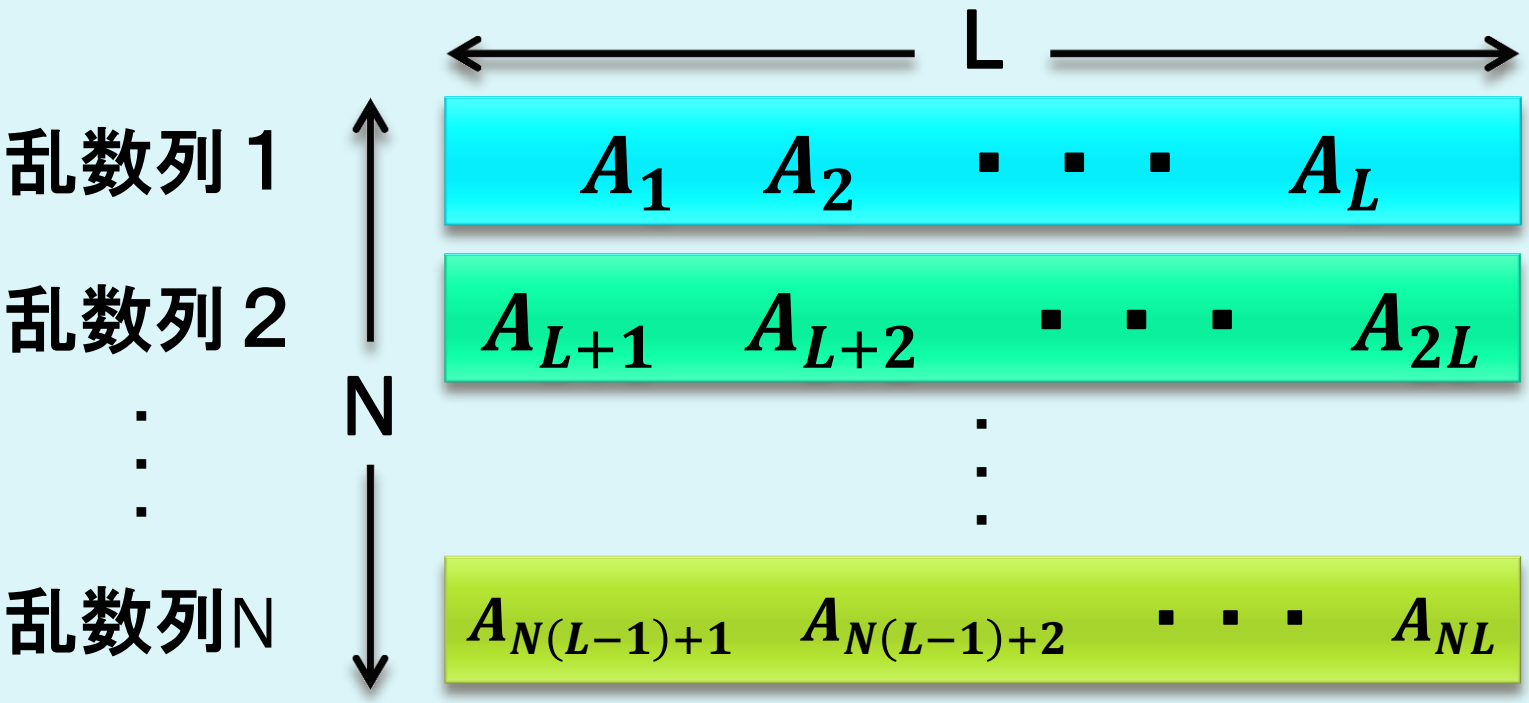
$A_1 \cdots A_L$　$A_{L+1} \cdots A_{2L}$ ・・・ $A_{(N-1)L+1} \cdots A_{NL}$

長さLの乱数列をN個用意

RMTテストに向けた数値処理

乱数データから行列A作成

行列A



RMTテストに向けた数値処理

乱数データから行列A作成

正規化(行列Gの作成)

自己相関行列Cの作成

乱数列処理

Cの対角要素の平均算出

理論値との誤差算出

数値処理

RMTテストに向けた数値処理

正規化（例）

$$\langle G \rangle = \frac{1+2+3+4}{4}$$

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 1 & 6 \\ 2 & 1 & 7 & 4 \end{pmatrix}$$

$$\langle G^2 \rangle = \frac{1^2+2^2+3^2+4^2}{4}$$

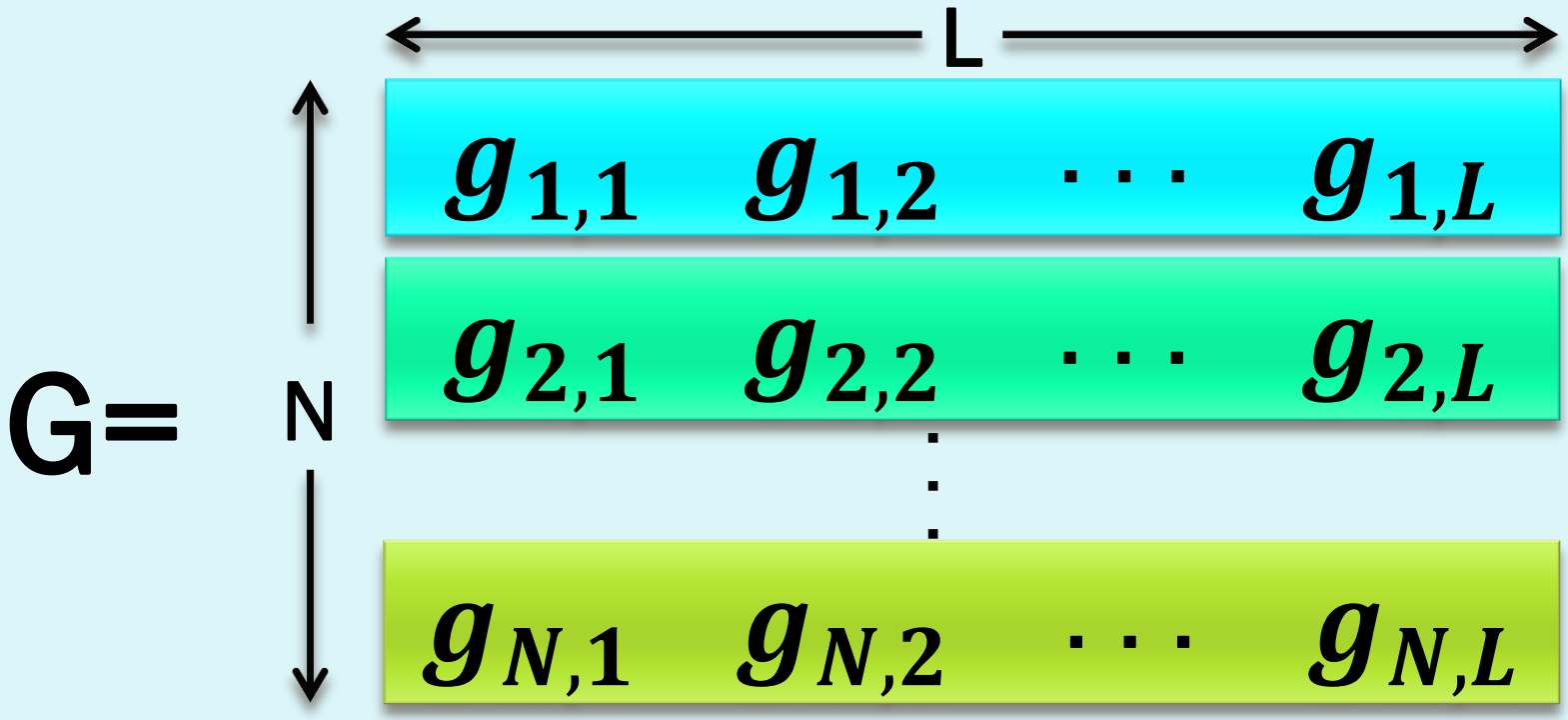
$$g(t) = \frac{G(t) - \langle G \rangle}{\sqrt{\langle G^2 \rangle - \langle G \rangle^2}}$$

$$G = \begin{pmatrix} g_{1,1} & \cdots & g_{1,L} \\ \vdots & \ddots & \vdots \\ g_{N,1} & \cdots & g_{N,L} \end{pmatrix}$$

RMTテストに向けた数値処理

正規化

平均０
分散１



RMTテストに向けた数値処理

乱数データから行列A作成

正規化(行列Gの作成)

自己相関行列Cの作成

乱数列処理

Cの対角要素の平均算出

理論値との誤差算出

数値処理

RMTテストに向けた数値処理

自己相関行列Cの作成

$$\mathrm{C} = \frac{1}{L} G G^{T}$$

$$\frac{1}{L} \times \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,L} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,L} \\ \vdots & & & \\ g_{N,1} & g_{N,2} & \cdots & g_{N,L} \end{bmatrix} \times \begin{bmatrix} g_{1,1} & g_{2,1} & \cdots & g_{N,1} \\ g_{1,2} & g_{2,2} & & g_{N,2} \\ \vdots & \vdots & & \vdots \\ g_{1,L} & g_{2,L} & & g_{N,L} \end{bmatrix}$$

RMTテストに向けた数値処理

乱数データから行列A作成

正規化(行列Gの作成)

自己相関行列Cの作成

乱数列処理

Cの対角要素の平均算出

理論値との誤差算出

数値処理

RMTテストに向けた数値処理

Cの対角要素の平均算出

$$C^k = \begin{bmatrix} \left(C^k\right)_{1,1} & \cdots & \left(C^k\right)_{1,N} \\ \vdots & \ddots & \vdots \\ \left(C^k\right)_{N,1} & \cdots & \left(C^k\right)_{N,N} \end{bmatrix}$$

RMTテストに向けた数値処理

Cの対角要素の平均算出

評価の基準

$$C^k = \begin{bmatrix} \left(C^k\right)_{1,1} & \cdots & \left(C^k\right)_{1,N} \\ \vdots & \ddots & \vdots \\ \cdots & & \left(C^k\right)_{N,N} \end{bmatrix}$$

実測値

RMTテストに向けた数値処理

乱数データから行列A作成

正規化(行列Gの作成)

自己相関行列Cの作成

乱数列処理

Cの対角要素の平均算出

理論値との誤差算出

数値処理

RMTテストに向けた数値処理

理論値との誤差算出

$$\text{誤差}(\%) = \left| \left(\frac{m_k}{\mu_k} - 1 \right) \times 100 \right|$$

m_k：実測値　　μ_k：理論値

対角要素の平均　　既に求められている

０％に近いほど乱数度が高い

※比較実験にはk=6の場合を使用

RMTテストに向けた数値処理

理論値

$$\mu_k = \int_{\lambda_-}^{\lambda_+} \lambda^k P_{RMT}(\lambda) d\lambda$$

より導出

$$\mu_1 = 1$$

$$\mu_2 = 1 + \frac{1}{Q}$$

$$\mu_3 = 1 + \frac{3}{Q} + \frac{1}{Q^2}$$

$$\mu_4 = 1 + \frac{6}{Q} + \frac{6}{Q^2} + \frac{1}{Q^3}$$

$$\mu_5 = 1 + \frac{10}{Q} + \frac{20}{Q^2} + \frac{10}{Q^3} + \frac{1}{Q^4}$$

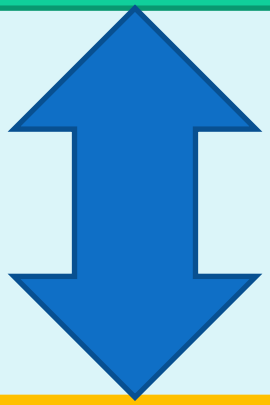
$$\mu_6 = 1 + \frac{15}{Q} + \frac{50}{Q^2} + \frac{50}{Q^3} + \frac{15}{Q^4} + \frac{1}{Q^5}$$

目次

実験説明

RMT
テスト

乱数度調査
実測値と理論値の誤差により判断



比較

NIST
乱数検定

15種類の検定の合否調査
合格検定数により判断

NISTとは

NIST SP 800-22

米国国立標準技術研究所(NIST)
で開発された

乱数検定用ツール

15種の検定

合格数が多いほど乱数度が高い

NISTとは

NIST乱数検定の項目　例えば・・・

連の検定

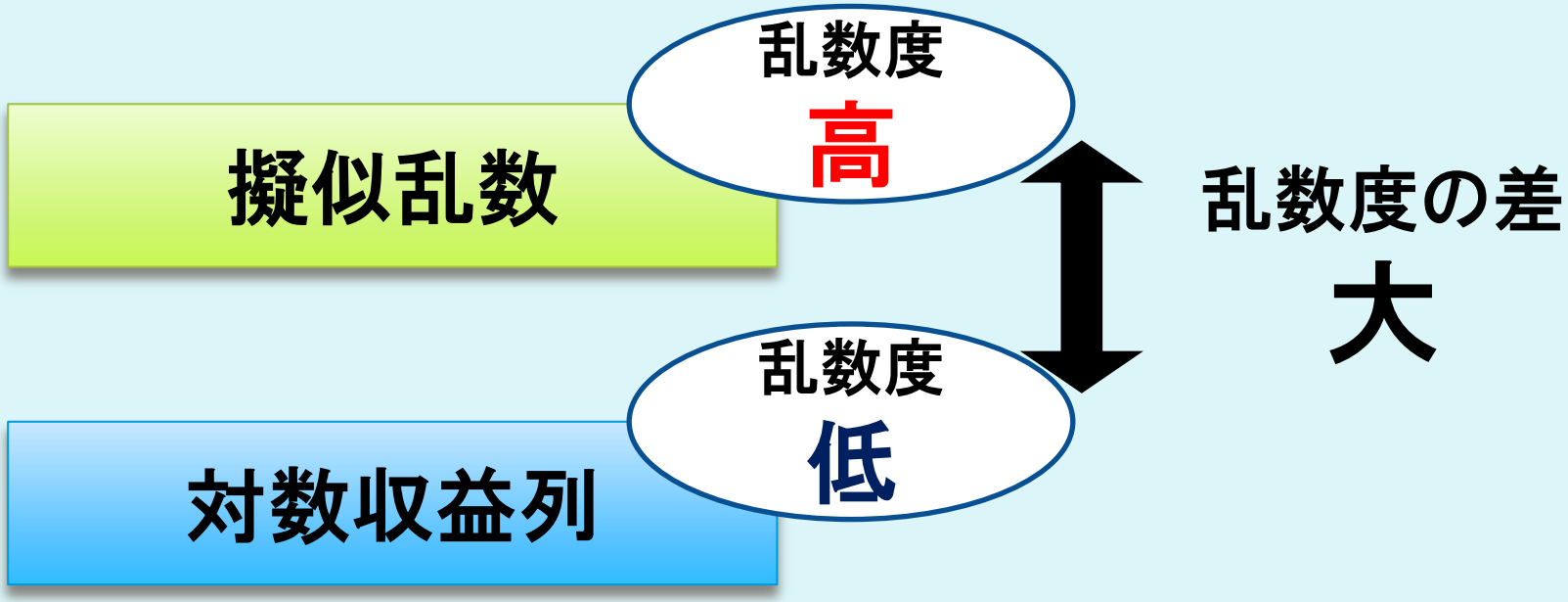
- 0または1の、連続している個数の偏りを調べる

テンプレート適合検定

- 乱数列を幾つかのブロックに分解
- 同じテンプレートの度数の偏りを調べる

使用データ　～異なる乱数度データの作成～

先行研究で用いた乱数データ



基準推定には
まだ不十分

使用データ　～異なる乱数度データの作成～

乱数度の違い(データ長100万、55サンプル平均)

乱数列の種類	元の数列(%)	対数収益列(%)
LCG	0.2831	99.30
日立	0.1597	98.87
東芝	0.0026	99.25
東京エレクトロンデバイス	0.1194	98.76

この中間の乱数度を持つ乱数データが必要

使用データ　～中間の乱数度データの作成～

規則的な数列をシャッフル

0,1の規則正しい数列

0,0,0,...0,1,1,1,...,1,・・・・・・・・・1,1,1,...1

シャッフル

００１１１０１１０１００・・・

使用データ

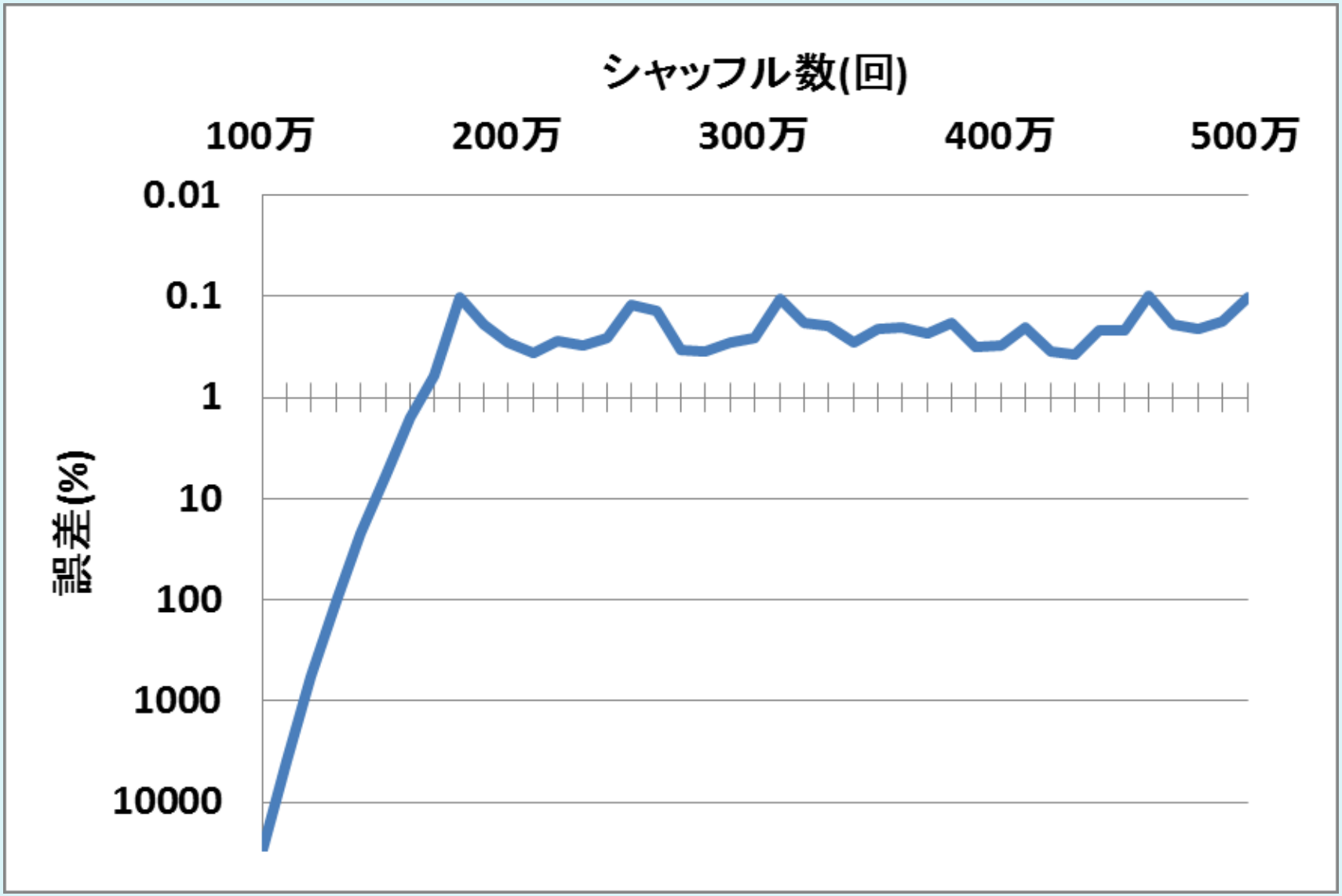
検証データ

範囲：シャッフル１００万～５００万回
１０万回刻みでファイルに出力して作成

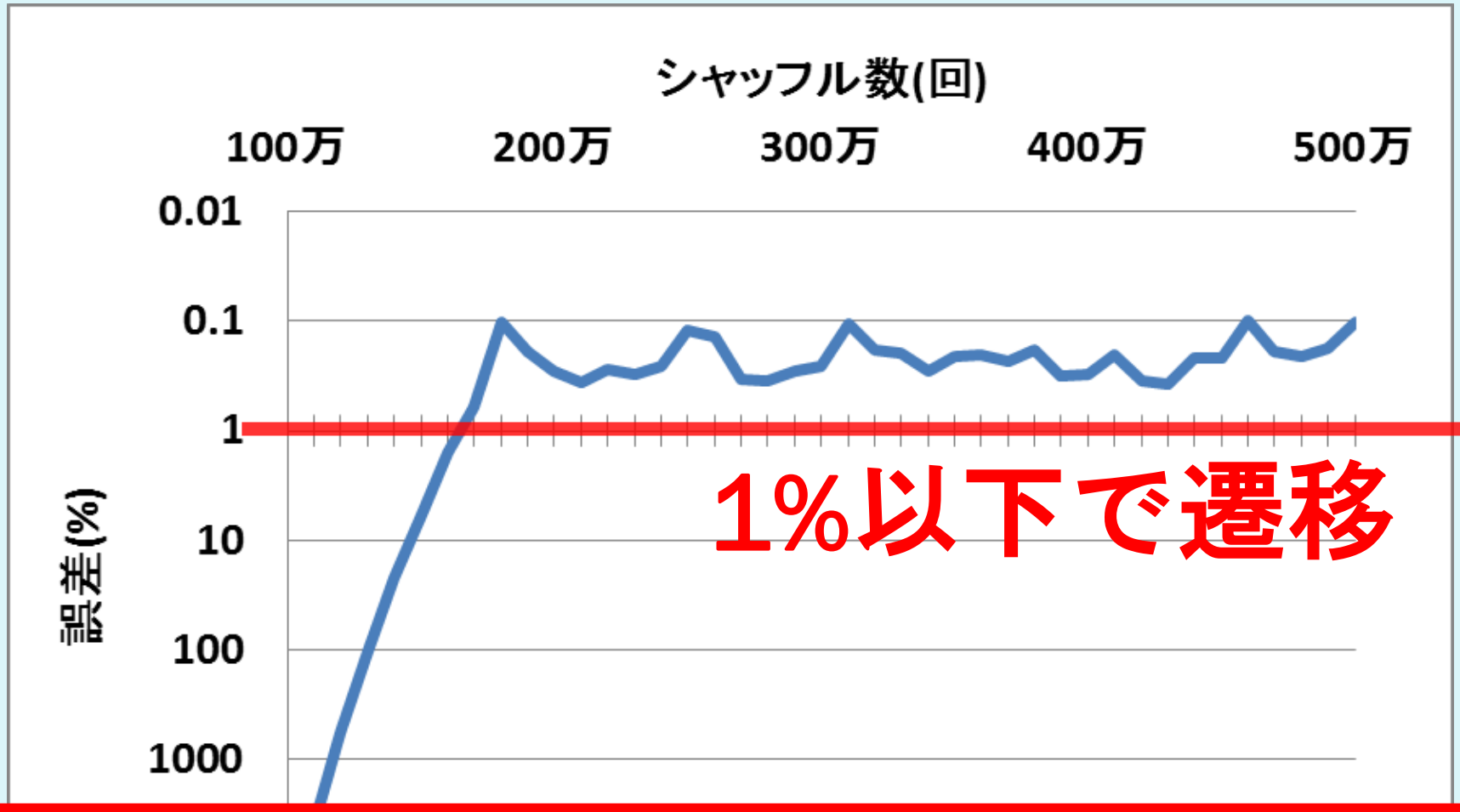
NIST乱数検定の条件に合わせる為に・・・

- 乱数列1サンプルのデータ長は**100万**
- 統計的に有意な結果を得るために
  55サンプル用意
- ０と１のみから構成される数列

結果　～RMTテスト～

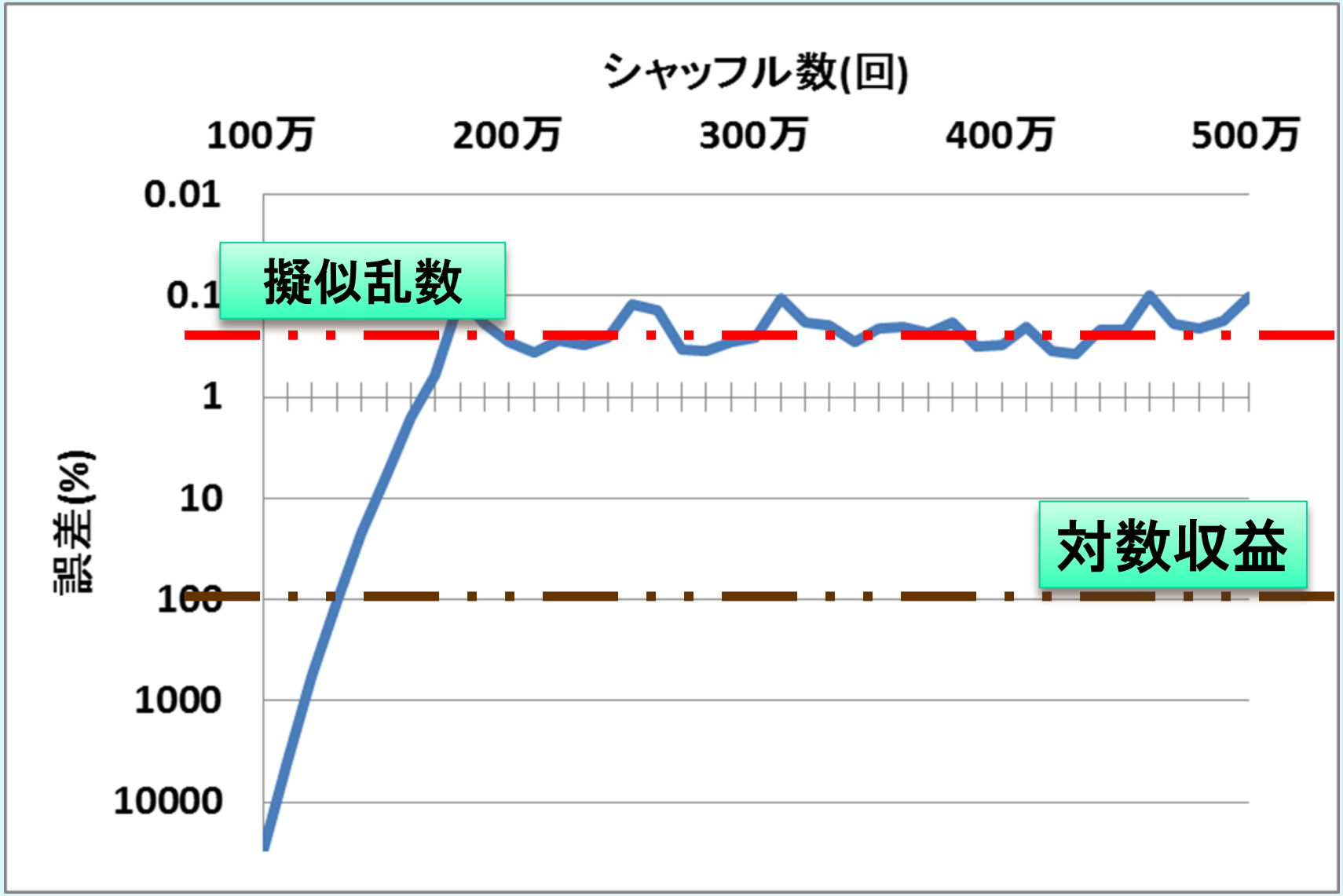


結果　～RMTテスト～

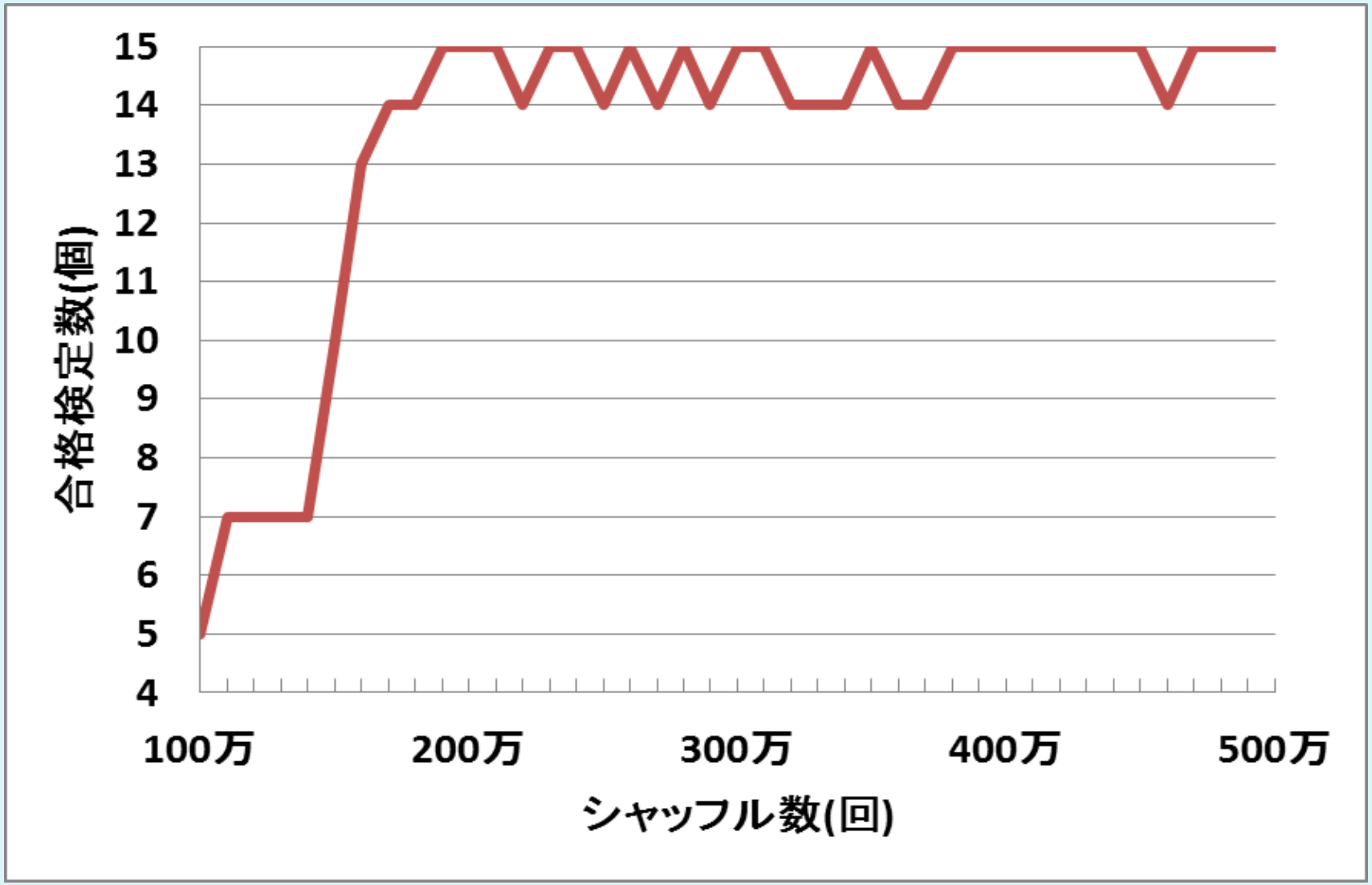


シャッフルによる乱数度向上を確認

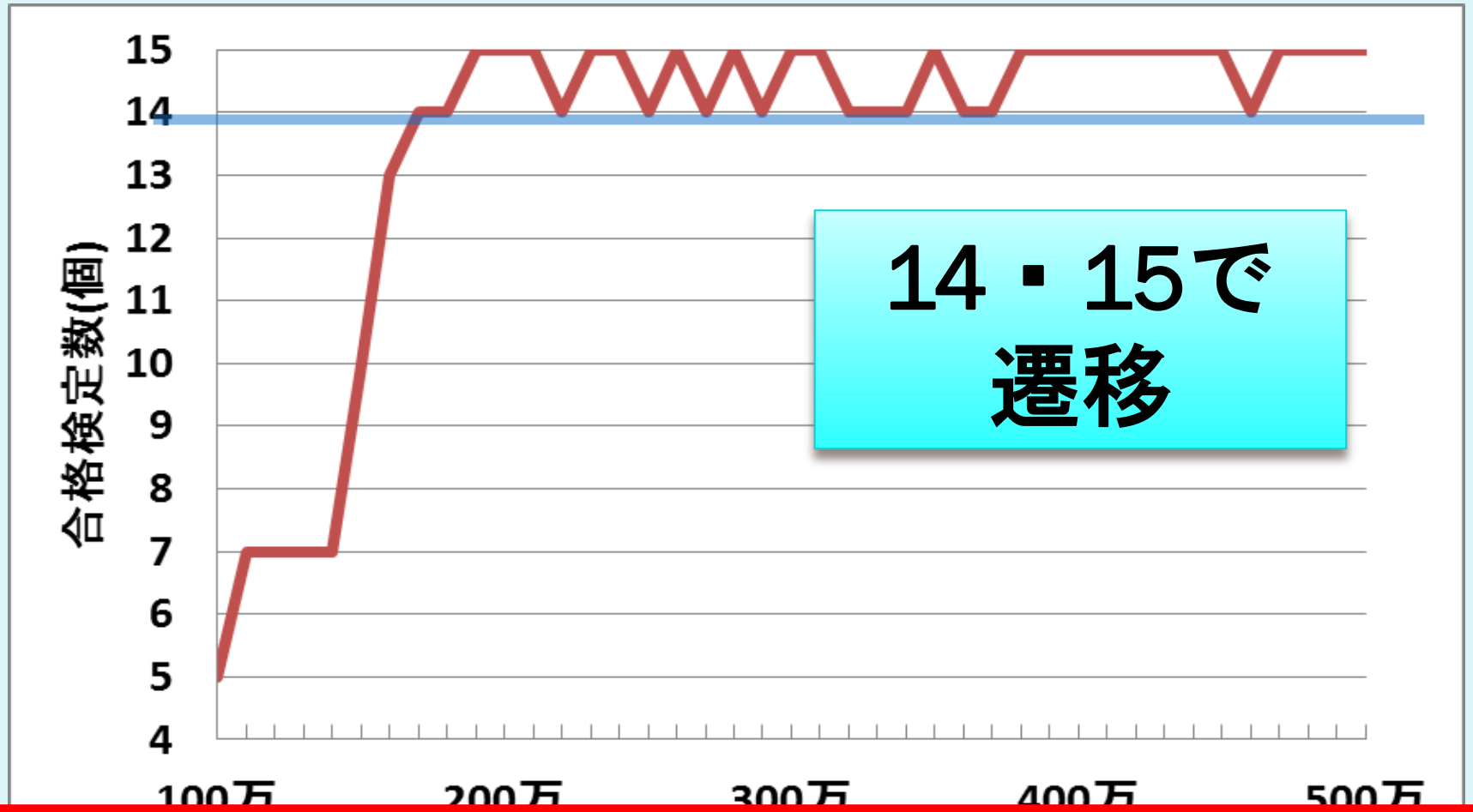
結果　～RMTテスト～


シャッフル数(回)
100万
200万
300万
400万
500万
0.01
0.1
1
10
100
1000
10000
誤差(%)
擬似乱数
対数収益

結果　～NIST乱数検定～

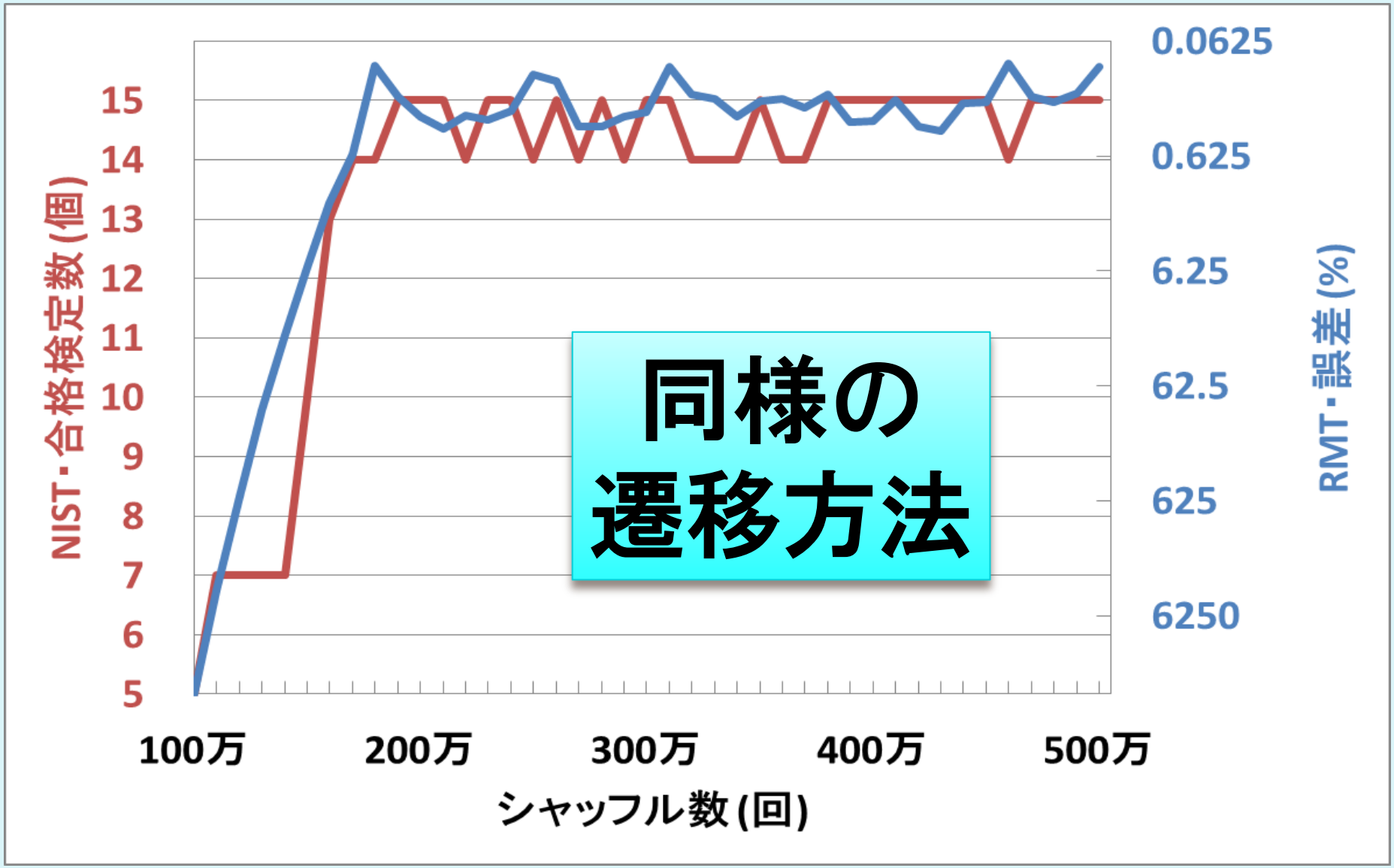

15
14
13
12
11
10
9
8
7
6
5
4
合格検定数(個)
100万
200万
300万
400万
500万
シャッフル数(回)

結果　～NIST乱数検定～

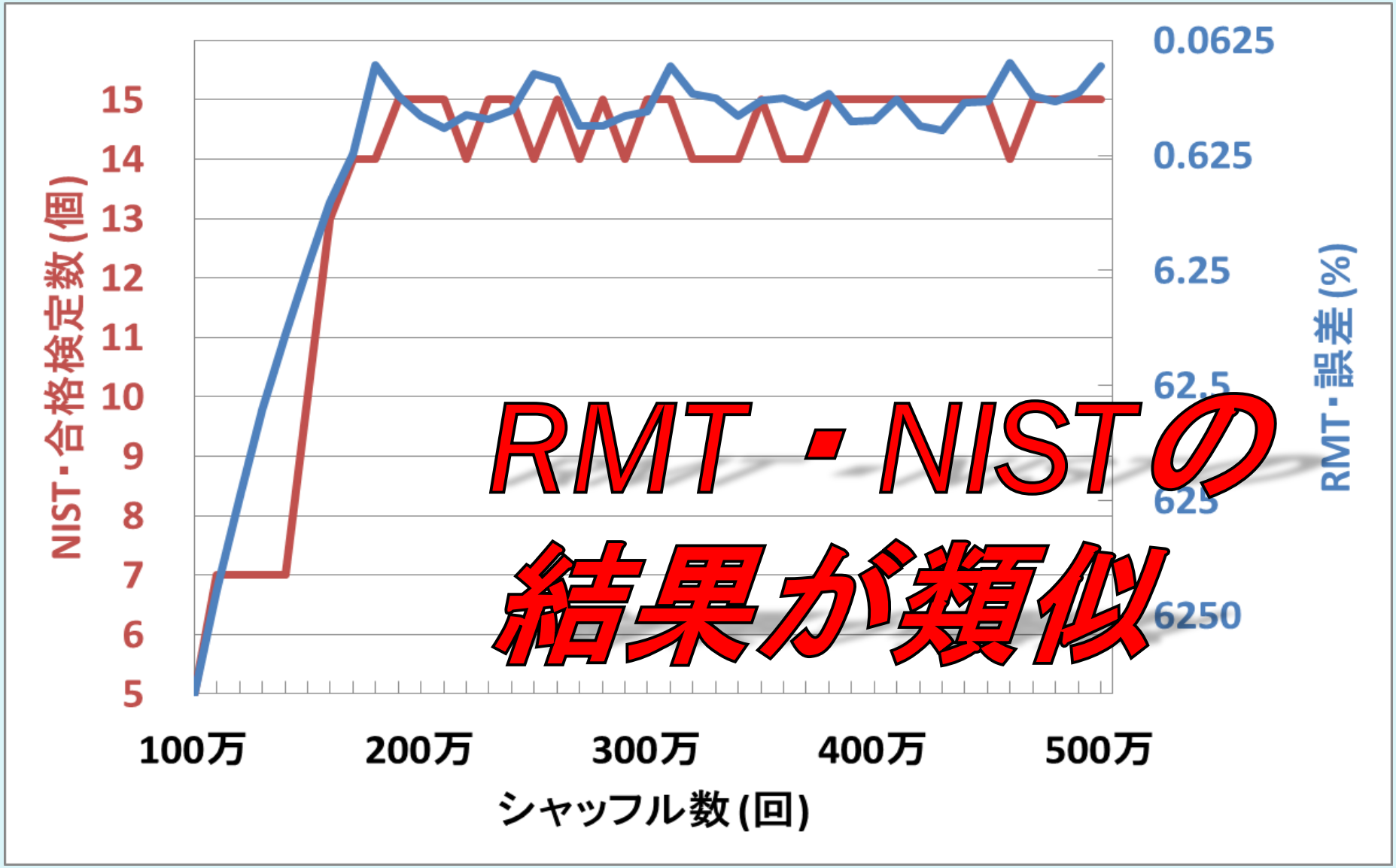


シャッフルによる乱数度向上を確認

結果　～比較～


NIST・合格検定数(個)
15
14
13
12
11
10
9
8
7
6
5
0.0625
0.625
6.25
62.5
625
6250
RMT・誤差(%)
100万
200万
300万
400万
500万
シャッフル数(回)
同様の
遷移方法

結果　～比較～


NIST・合格検定数(個)
15
14
13
12
11
10
9
8
7
6
5
0.0625
0.625
6.25
62.5
625
6250
RMT・誤差(%)
100万
200万
300万
400万
500万
シャッフル数(回)
RMT・NISTの
結果が類似

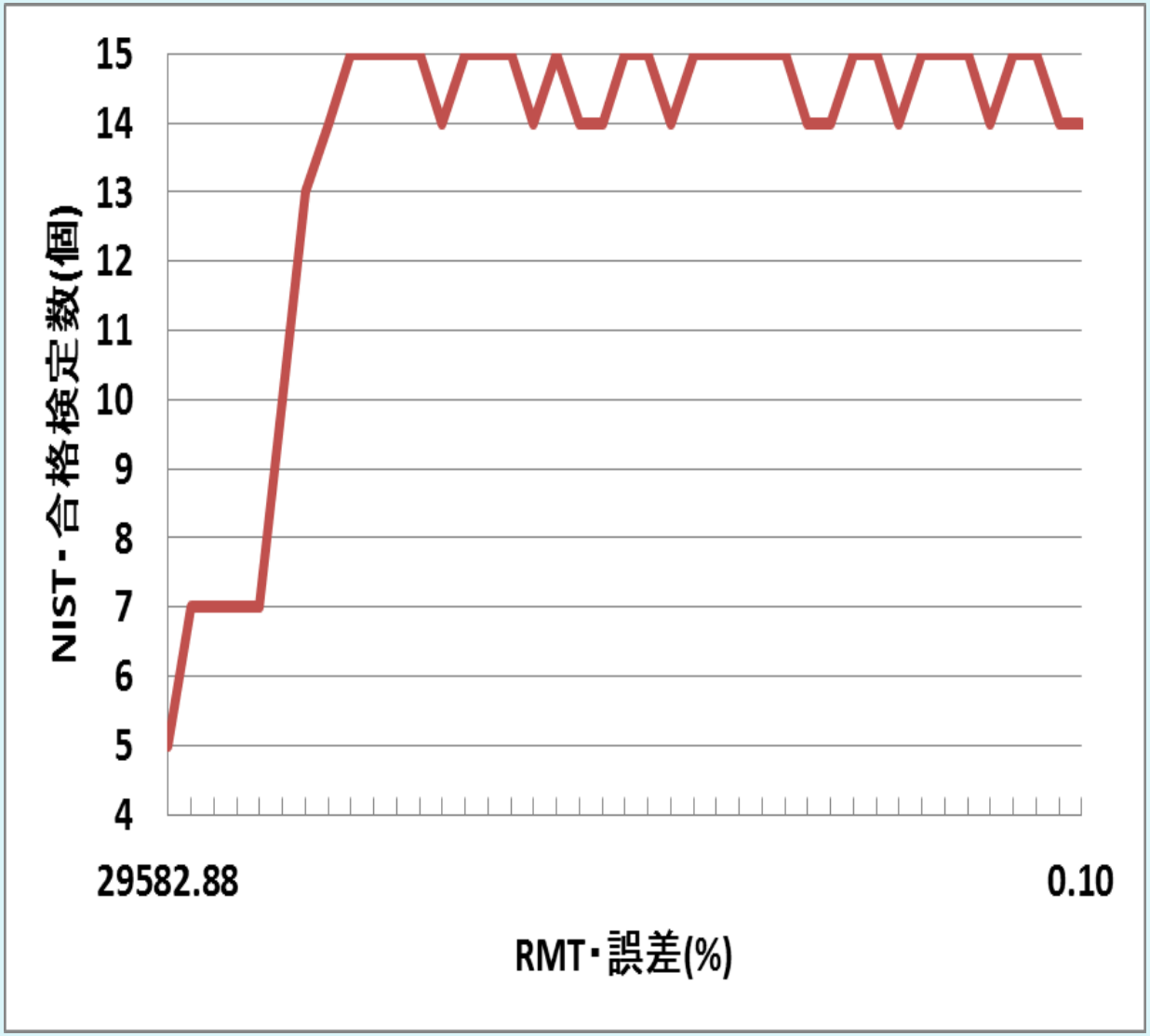
結果　～比較～

NIST乱数検定の結果は
RMTテストの結果との
比較対象として

有用

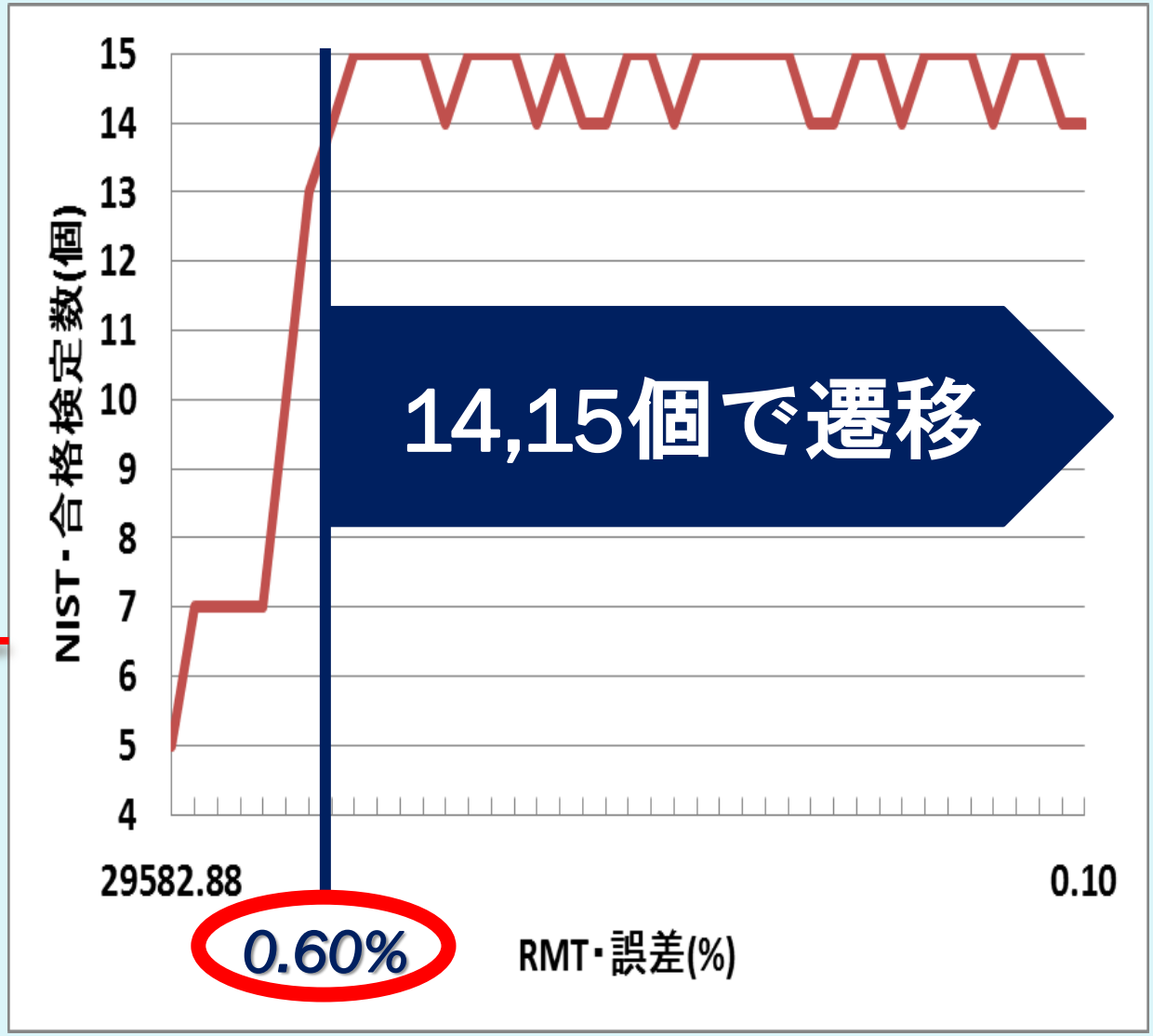
結果　～比較～

RMT誤差(%)	NIST合格率
29582.88	5/15
3803.87	7/15
572.09	7/15
101.80	7/15
22.27	7/15
5.79	10/15
1.60	13/15
0.60	14/15
0.38	15/15
0.36	15/15
0.35	15/15
0.34	15/15



結果　～比較～

RMT誤差(%)	NIST合格率
29582.88	5/15
3803.87	7/15
572.09	7/15
101.80	7/15
22.27	7/15
5.79	10/15
1.60	13/15
0.60	14/15
0.38	15/15
0.36	15/15
0.35	15/15
0.34	15/15



考察

良い乱数と見なせる誤差の基準

NISTと比較した場合

RMT誤差＝0.60%

考察　～基準のずれの原因～

NISTと比較した場合

- **データ長などのパラメータが違う**

  ↑NIST乱数検定のパラメータの
  条件を合わせるためにデータ長100万とした

- **標準偏差の違い**

考察

NIST:合格検定数14の時

毎回、同じ検定で不合格

RMTテストでは
検出できない特徴がある

考察

不合格の検定

重なりの無い
テンプレート適合検定

終わりに

NIST乱数検定を
比較対象とした
RMTテストの
基準を定めることができた

終わりに

今後の課題

- 誤差0.60%付近をより精密に検証
- NIST以外の比較の検討
- 不合格検定についての調査