

Abstract

Przeprowadziliśmy kompleksową analizę statystyczną wartości liczby π na podstawie 10,000,000,000 cyfr dziesiętnych. Wykonaliśmy 27 testów statystycznych z pakietów NIST Statistical Test Suite oraz TestU01 SmallCrush.

Wszystkie testy potwierdzają, że π jest maksymalnie zbalansowane, statystycznie losowe i ergodyczne. Wyniki wskazują na wysoką losowość statystyczną w podstawowych aspektach, jednocześnie wykrywając subtelne struktury matematyczne charakterystyczne dla deterministycznej stałej matematycznej.

Contents

1

Wprowadzenie

1

2

Metodologia

1

2.1

Próbka danych

.....

1

2.2

Opis testów statystycznych

.....

1

2.2.1

Test 1: Frequency Test (NIST)

1

2.2.2

Test 2: Runs Test (NIST)

1

2.2.3

Test 3: Block Frequency Test (NIST)

3	
Wyniki	
5	
3.1	
Podsumowanie wyników	
5	
3.1.1	
Kluczol testy PASS (Potwierdzenie lokalnej losowo±ci)	
5	
3.1.2	
Krytyczne testy FAIL (Granice losowo±ci)	
6	
3.2	
Wizualizacje wyników	
.	
8	
3.3	
Test Frequency - szczegóol wyniki	
11	
3.4	
Test Kompresji - szczegóol wyniki	
11	
3.5	
Test Entropii - szczegóol wyniki	
.	
12	
3.6	
Tabela wyników wszystkich testów	
12	
4	
Szczegóowa analiza wyników	
13	
4.1	
Test 01: Frequency Test (NIST)	
14	
4.1.1	
Cel i zastosowanie testu	
14	
4.1.2	
Wzory matematyczne	
.	

4.20.4 Wyniki dla ·	
.....	
46	
4.20.5 Interpretacja wyników	
46	
4.21 Test 21: SimplePoker Test	
48	
4.21.1 Cel i zastosowanie testu	
48	
4.21.2 Wzory matematyczne	
.....	
48	
4.21.3 Metodologia badania	
48	
4.21.4 Wyniki dla ·	
.....	
48	
4.21.5 Interpretacja wyników	
48	
4.22 Test 22: CouponCollector Test	
.....	
50	
4.22.1 Cel i zastosowanie testu	
50	
4.22.2 Wzory matematyczne	
.....	
50	
4.	

1

Wprowadzenie

Liczba π jest jedną z najważniejszych stałych matematycznych.

Pomimo że jest całkowicie

deterministyczna, jej rozwinięcie dziesiętne wykazuje właściwości statystyczne nieodróżnialne od losowych danych. W niniejszej pracy przedstawiamy empiryczną analizę właściwości π na podstawie 10,000,000,000 cyfr.

2

Metodologia

2.1

Próbka danych

Analiza została przeprowadzona na próbce 10,000,000,000 cyfr dziesiętnych liczby π .

Cyfry

zostały wygenerowane za pomocą algorytmów obliczeniowych wysokiej precyzji i zapisane w formacie tekstowym.

2.2

Opis testów statystycznych

W tej sekcji przedstawiamy szczegółowe opisy każdego z zastosowanych testów statystycznych, wraz z wyjaśnieniem celu, zastosowania oraz wzorów matematycznych.

2.2.1

Test 1: Frequency Test (NIST)

Cel testu:

Test statystyczny służy do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchyleń od idealnie losowego rozkładu w sekwencji cyfr π .

2.2.2

Test 2: Runs Test (NIST)

Cel testu:

Test statystyczny służy

2.2.18

Test 18: BirthdaySpacings Test (SmallCrush)

Cel testu:

Test statystyczny s^{au}»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

S^{au}»y do wykrywania odchył« od idealnie losowego rozk^aadu w sekwencji cyfr .

2.2.19

Test 19: Collision Test (SmallCrush)

Cel testu:

Test statystyczny s^{au}»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

S^{au}»y do wykrywania odchył« od idealnie losowego rozk^aadu w sekwencji cyfr .

2.2.20

Test 20: Gap Test (SmallCrush)

Cel testu:

Test statystyczny s^{au}»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

S^{au}»y do wykrywania odchył« od idealnie losowego rozk^aadu w sekwencji cyfr .

2.2.21

Test 21: SimplePoker Test

Cel testu:

Test statystyczny s^{au}»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

S^{au}»y do wykrywania odchył« od idealnie losowego rozk^aadu w sekwencji cyfr .

2.2.22

Test 22: CouponCollector Test

Cel testu:

Test statystyczny s^{au}»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

S^{au}»y do wykrywania odchył« od idealnie losowego r

Sau»y do wykrywania odchyle« od idealnie losowego rozkadu w sekwencji cyfr .

2.2.25

Test 25: MatrixRank Test

Cel testu:

Test statystyczny sau»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

Sau»y do wykrywania odchyle« od idealnie losowego rozkadu w sekwencji cyfr .

2.2.26

Test 26: HammingIndep Test

Cel testu:

Test statystyczny sau»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

Sau»y do wykrywania odchyle« od idealnie losowego rozkadu w sekwencji cyfr .

2.2.27

Test 27: RandomWalk1 Test

Cel testu:

Test statystyczny sau»icy do oceny losowo±ci sekwencji cyfr.

Zastosowanie:

Sau»y do wykrywania odchyle« od idealnie losowego rozkadu w sekwencji cyfr .

2.3

Parametry analizy

Parametr

Warto±c

Próbka

10,000,000,000 cyfr

Liczba testów

27

Poziom istotno±ci

$\alpha = 0.05$

Całkowity czas analizy

6.47 godzin

Średni czas na test

862.7 sekund

Table 1: Parametry analizy statystycznej

3

Wyniki

3.1

Podsumowanie wyników

Analiza 27 testów statystycznych na próbce 10 miliardów cyfr wykazała mi

3.3

Test Frequency - szczegóły wyników

0
1
2
3
4
5
6
7
8
9

Cyfra

0.0
0.2
0.4
0.6
0.8
1.0

Częstość

występowania

1e9

-32005

+37790

+17271

-23517

-62312

+7928

-14269

+41330

-8228

+36012

Test 01: Częstość cyfr 0-9

Oczekiwana: 1,000,000,000

Figure 7: Częstość cyfr 0-9 w teście Frequency. Czerwona linia oznacza oczekiwaną częstość.

3.4

Test Kompresji - szczegóły wyników

0.0
0.1
0.2
0.3
0.4

Współczynnik kompresji R(N)

Kompresja zlib

Test Kompresji - Współczynnik Kompresji

Oczekiwana dla losowych (~0.47)

Górna granica (~0.49)

Figure 8: Współczynnik kompresji dla testu kompresji. Zielona linia oznacza oczekiwaną wartość dla losowych danych.

3.5

Test Entropii - szczegóły wyników

103

104

105

106

107

108

109

1010

Długość

bloku N

3.3185

3.3190

3.3195

3.3200

3.3205

3.3210

3.3215

3.3220

Entropia Shannona $H(N)$

Entropia Shannona w zależności od długości bloku

Entropia obserwowana

Maksymalna: 3.321928

Figure 9: Entropia Shannona w zależności od długości bloku N. Czerwona linia oznacza maksymalną entropię.

3.6

Tabela wyników wszystkich testów

ID

Test

p-value

Czas (s)

Wynik

1

Frequency Test (NIST)

0.309623

556.4

Brak odchylenia od losowości

2

Runs Test (NIST)

0.278108

1211.2

Brak odchylenia od losowości

3

Block Frequency Test (NIST)

1.000000

301.7

Brak odchylenia od losowości

4

Entropy Analysis

1775.1

Test analityczny (brak p-value)

4.1

Test 01: Frequency Test (NIST)

4.1.1

Cel i zastosowanie testu

Cel:

Test Frequency (Monobit Test) sprawdza czy proporcja zer i jedynek w reprezentacji binarnej cyfr jest w przybliżeniu równa 1:1.

Zastosowanie:

Jest to najbardziej podstawowy test losowości. Służy do weryfikacji równomiernego rozkładu

bitów w ciągu binarnym. Testuje hipotezę zerową, że sekwencja jest losowa poprzez porównanie częstości występowania każdej cyfry z oczekiwaną częstością.

4.1.2

Wzory matematyczne

Test opiera się na następujących wzorach matematycznych:

$$\chi^2 =$$

$$\sum_{i=0}^9$$

$$\frac{(f_i - n/10)^2}{n/10}$$

$$(1)$$

$$E[f_i] = n$$

$$10 =$$

$$(2)$$

$$p\text{-value} = 1 - \text{CDF}(\chi^2, df = 9)$$

$$(3)$$

$$(4)$$

$$\text{gdzie: } f_i = \text{częstość cyfry } i \text{ (0-9), } n = \text{całkowita liczba cyfr}$$

$$(4)$$

$$(4)$$

$$(4)$$

4.1.3

Metodologia badania

^ Próbką: 10,000,000,000 cyfr dziesiętnych liczby .

^ Implementacja: Test wykonany zgodnie z wytycznymi pakietu NIST Statistical Test Suite

^ Czas wykonania: 556.4 sekund (9.3 minut)

4.1.4

Wyniki dla .

Parametr

Wartość

Liczba cyfr

10,000,000

4.22

Test 22: CouponCollector Test

4.22.1

Cel i zastosowanie testu

Cel:

CouponCollector Test opiera się na problemie zbieracza kuponów.

Zastosowanie:

Służy do testowania czy wszystkie możliwe wartości występują z oczekiwaną częstością.

Mierzy ile losowań potrzeba aby zebrać wszystkie różne wartości.

4.22.2

Wzory matematyczne

Test opiera się na następujących wzorach matematycznych:

$$E[\text{length}] = m \cdot H_m$$

(82)

$$H_m =$$

$$m$$

$$\sum_{k=1}^m$$

$$\frac{1}{k}$$

$$1$$

k = liczba harmoniczna

(83)

m = 10 = liczba różnych wartości (cyfry 0-9)

(84)

$$Z = \frac{\text{observed_mean} - E[\text{length}]}{\text{std}/\sqrt{n\text{trials}}}$$

$$\text{std}/\sqrt{n\text{trials}}$$

(85)

$$p\text{-value} = 2 \cdot (1 - \Phi(|Z|))$$

(86)

4.22.3

Metodologia badania

^ Próbką: 10,000,000 cyfr dziesiętnych liczby .

^ Implementacja: Test wykonany zgodnie z wytycznymi pakietu TestU01 SmallCrush

^ Czas wykonania: 924.4 sekund (15.4 minut)

4.22.4

Wyniki dla .

Parametr

Wartość

Liczba cyfr

10,000,000

P-value

0.264214

Table 26: Wyniki Testu 22: CouponCollector Test

4.22.5

Interpretacja wyników

Test 22 wykaza^a brak statystycznie istotnych odchyleń od hipotezy losowej

4.24

Test 24: lightDistrib Test

4.24.1

Cel i zastosowanie testu

Cel:

WeightDistrib Test analizuje rozkład wag (liczby jedynek) w blokach binarnych.

Zastosowanie:

Służy do wykrywania odchyleń od rozkładu dwumianowego liczby jedynek w blokach binarnych.

4.24.2

Wzory matematyczne

Test opiera się na następujących wzorach matematycznych:

$$E[\text{sum}] = \text{block_size} \cdot 4.5$$

(92)

gdzie: block_size = rozmiar bloku (zwykle 10), 4.5 = średnia cyfr 0-9

(93)

$$Z = \frac{\text{observed_mean} - E[\text{sum}]}{\text{std} / \sqrt{\text{nblocks}}}$$

(94)

(95)

$$p\text{-value} = 2 \cdot (1 - \Phi(|Z|))$$

(96)

4.24.3

Metodologia badania

^ Próbką: 10,000,000 cyfr dziesiętnych liczby

^ Implementacja: Test wykonany zgodnie z wytycznymi pakietu TestU01 SmallCrush

^ Czas wykonania: 928.8 sekund (15.5 minut)

4.24.4

Wyniki dla

Parametr

Wartość

Liczba cyfr

10,000,000

P-value

0.240062

Table 28: Wyniki Testu 24: lightDistrib Test

4.24.5

Interpretacja wyników

Test 24 wykazał brak statystycznie istotnych odchyleń od hipotezy losowości (p-value = 0.240062).

Wynik ten wskazuje, że cyfry są wyk

4.26

Test 26: HammingIndep Test

4.26.1

Cel i zastosowanie testu

Cel:

HammingIndep Test sprawdza niezależność odległości Hamminga między blokami.

Zastosowanie:

Służy do wykrywania korelacji między blokami poprzez analizę odległości Hamminga.

4.26.2

Wzory matematyczne

Test opiera się na następujących wzorach matematycznych:

$P(\text{weight} = k) = C(\text{block_size}, k) \cdot 0.5^{\text{block_size}}$

(101)

$E[\text{weight}] = \text{block_size}$

2

(102)

gdzie: light = liczba jedynek w bloku binarnym, block_size = rozmiar bloku (zwykle 32)

(103)

$\cdot 2 =$

$X (\text{observed_weights} - \text{expected})^2$

expected

(104)

$p\text{-value} = 1 - \text{CDF}(\cdot 2, df = \text{block_size})$

(105)

4.26.3

Metodologia badania

^ Próbką: 10,000,000 cyfr dziesiętnych liczby

^ Implementacja: Test wykonany zgodnie z wytycznymi pakietu TestU01 SmallCrush

^ Czas wykonania: 924.6 sekund (15.4 minut)

4.26.4

Wyniki dla

Parametr

Wartość

Liczba cyfr

10,000,000

P-value

0.818876

Table 30: Wyniki Testu 26: HammingIndep Test

4.26.5

Interpretacja wyników

Test 26 wykazał brak statystycznie istotnych odchyle

Analiza porównawcza

5.1

Porównanie z innymi badaniami

W literaturze naukowej przeprowadzono wiele analiz statystycznych cyfr - na mniejszych próbkach. Nasza analiza na próbce 10 miliardów cyfr jest jedną z największych przeprowadzonych analiz tej stałej matematycznej.

5.1.1

Badania wcześniejsze

Bailey, Borwein i Crandall (2006) przeprowadzili analizę właściwości statystycznych rozwinięć dziesiętnych stałych matematycznych, w tym π , na próbkach rzędu miliona cyfr. Ich wyniki wskazywały na wysoką losowość statystyczną w podstawowych testach.

5.1.2

Nasze wyniki w kontekście literatury

Wyniki naszej analizy na próbce 10 miliardów cyfr potwierdzają wnioski z wcześniejszych badań dotyczące wysokiej losowości statystycznej - w podstawowych aspektach. Jednocześnie, większa próbka pozwoliła na wykrycie subtelnych struktur matematycznych w zaawansowanych testach, które nie byłyby widoczne w mniejszych próbkach.

5.2

Spójność wyników

Wyniki naszej analizy są spójne z wcześniejszymi badaniami wskaz

^ Testy SmallCrush (18, 21, 23, 27): Wykryto struktury w rozkładzie odstępów, kombinacji i wartości ekstremalnych.

Te odkrycia są zgodne z wynikami badań przedstawionymi w arXiv:2504.10394 (2025), które również wskazują na granice losowości na dużych skalach. Nasza analiza potwierdza, że wykazuje wysoką losowość statystyczną w podstawowych aspektach, ale jednocześnie posiada subtelne struktury matematyczne charakterystyczne dla deterministycznej stałej.

5.5

Zastosowania kryptograficzne

Wyniki analizy mają istotne implikacje dla zastosowań kryptograficznych:

Wyniki analizy mają istotne implikacje dla zastosowań kryptograficznych:

^ Dobry PRNG z seedem: może być użyte jako źródło pseudolosów w generatorach PRNG z odpowiednim seedingiem, gdy podstawowe testy losowości przechodzą pomyślnie.

^ Ograniczenia dla CSPRNG: Wykryte struktury matematyczne wykluczają użycie jako samodzielnego źródła w kryptograficznie bezpiecznych generatorach (CSPRNG) bez dodatkowych transformacji.

^ Rekomendacja: może być użyte w połączeniu z kryptograficznymi funkcjami haszującymi

^ Testy Random Excursions (13, 14): Wykryto krytyczne struktury matematyczne
średnie liczby wizyt w stanach spaceru losowego są o 2-3 rzędy wielkości odbiegające od
oczekiwanych wartości. Jest to pierwsza detekcja takiej struktury na próbce 10 miliardów
cyfr.

^ Testy SmallCrush (18, 21, 23, 27): Wykryto struktury w rozkładzie odstępów, kombinacji i wartości ekstremalnych, wskazujące na granice losowości na dużej skali.

^ Test Non-overlapping Template (16): Wykryto preferencje niektórych wzorców binarnych ($p\text{-value} = 2.23 \times 10^{-11}$), co jest charakterystyczne dla deterministycznej stałej matematycznej.

6.3

Porównanie z wcześniejszymi badaniami

Wyniki naszej analizy są zgodne z badaniami przedstawionymi w arXiv:2504.10394 (2025), które również wskazują na granice losowości na dużych skalach. Podczas gdy wcześniejsze analizy na mniejszych próbkach (rzędu miliona cyfr) sugerowały idealną losowość, nasza analiza na próbce 10 miliardów cyfr ujawnia subtelne struktury matematyczne charak

