

Abstract

Przeprowadziliśmy kompleksową analizę statystyczną właściwości liczby · 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 · jest

maksymalnie złożone, statystycznie losol 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	
Kluczowe testy PASS (Potwierdzenie lokalnej losowości)	5
3.1.2	
Krytyczne testy FAIL (Granice losowości)	6
3.2	
Wizualizacje wyników	
.....	8
3.3	
Test Frequency - szczegółowy wyniki	11
3.4	
Test Kompresji - szczegółowy wyniki	11
3.5	
Test Entropii - szczegółowy wyniki	
.....	12
3.6	
Tabela wyników wszystkich testów	12
4	
Szczegółowa 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 dzisiajne wykazuje właściwości statystyczne nieodróżnialne od losowych danych. W niniejszej pracy przedstawiamy empiryczny analizę właściwości na podstawie 10,000,000,000 cyfr.

2

Metodologia

2.1

Próbka danych

Analiza została przeprowadzona na próbce 10,000,000,000 cyfr dzisiajnych 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 szczegółowo opisy kolejnego zastosowanych testów statystycznych, wraz z wyjaśnieniem celu, zastosowania oraz wzorów matematycznych.

2.2.1

Test 1: Frequency Test (NIST)

Cel testu:

Test statystyczny służący do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchylienia od idealnie losowego rozkładu w sekwencji cyfr.

2.2.2

Test 2: Runs Test (NIST)

Cel testu:

Test statystyczny służący

2.2.18

Test 18: BirthdaySpacings Test (SmallCrush)

Cel testu:

Test statystyczny służy do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.19

Test 19: Collision Test (SmallCrush)

Cel testu:

Test statystyczny służy do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.20

Test 20: Gap Test (SmallCrush)

Cel testu:

Test statystyczny służy do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.21

Test 21: SimplePoker Test

Cel testu:

Test statystyczny służy do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.22

Test 22: CouponCollector Test

Cel testu:

Test statystyczny służy do oceny losowości sekwencji cyfr.

Zastosowanie:

Służy do wykrywania odchyłek od idealnie losowego r

S^au»y do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.25

Test 25: MatrixRank Test

Cel testu:

Test statystyczny s^au»jcy do oceny losowości sekwencji cyfr.

Zastosowanie:

S^au»y do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.26

Test 26: HammingIndep Test

Cel testu:

Test statystyczny s^au»jcy do oceny losowości sekwencji cyfr.

Zastosowanie:

S^au»y do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.2.27

Test 27: RandomWalk1 Test

Cel testu:

Test statystyczny s^au»jcy do oceny losowości sekwencji cyfr.

Zastosowanie:

S^au»y do wykrywania odchyłek od idealnie losowego rozkładu w sekwencji cyfr ..

2.3

Parametry analizy

Parametr

Wartość

Próbka

10,000,000,000 cyfr

Liczba testów

27

Poziom istotności

$\alpha = 0.05$

Ciekawy 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 - szczegółów wyniki

0
1
2
3
4
5
6
7
8
9

Cyfra

0.0
0.2
0.4
0.6
0.8
1.0

Cz sto

wyst pie
1e9
-32005
+37790
+17271
-23517
-62312
+7928
-14269
+41330
-8228
+36012

Test 01: Cz sto ci Cyfr 0-9

Oczekiwana: 1,000,000,000

Figure 7: Cz±sto±ci cyfr 0-9 w te±cie Frequency. Czerwona linia oznacza oczekiwani; cz±sto±c.

3.4

Test Kompresji - szczegółów wyniki

0.0
0.1
0.2
0.3
0.4

Wspóczynnik kompresji R(N)

Kompresja zlib

Test Kompresji - Wspóczynnik Kompresji

Oczekiwana dla losowych (~0.47)

Górna granica (~0.49)

Figure 8: Wspóczynnik kompresji dla testu kompresji. Zielona linia oznacza oczekiwani; warto±c dla losowych danych.

3.5

Test Entropii - szczegółów wyniki

103

104

105

106

107

108

109

1010

D ugo

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 dugości 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 odchyłek od losowości

2

Runs Test (NIST)

0.278108

1211.2

Brak odchyłek od losowości

3

Block Frequency Test (NIST)

1.000000

301.7

Brak odchyłek 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 doeryjacji równomiernego rozkładu bitów w ciągu binarnym. Testuje hipotezę zerową, że sekwencja jest losowa poprzez porównanie częstości występowania każdej cyfry z oczekiwanej częstości.

4.1.2

Wzory matematyczne

Test opiera się na następujących wzorach matematycznych:

$$\cdot 2 =$$

$$9$$

$$X$$

$$i=0$$

$$(f_i \cdot n/10)2$$

$$n/10$$

$$(1)$$

$$E[f_i] = n$$

10 = oczekiwana częstość każdej cyfry

$$(2)$$

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

$$(3)$$

gdzie: f_i = częstość cyfry i (0-9), n = całkowita liczba cyfr

$$(4)$$

4.1.3

Metodologia badania

^ Próbka: 10,000,000,000 cyfr dziesiętnych liczb.

^ 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 oczekiwanej częstością.

Mierzy ile losowa 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)

$$H_m =$$

$$m$$

$$X$$

$$k=1$$

$$1$$

k = liczba harmoniczna

(83)

$$m = 10 = \text{liczba różnych wartości (cyfry 0-9)}$$

(84)

$$Z = \text{observed_mean} - E[\text{length}]$$

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

(85)

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

(86)

4.22.3

Metodologia badania

^ Próbka: 10,000,000 cyfr dziesiętnych liczb.

^ 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 wykazał brak statystycznie istotnych odchyleń od hipotezy loso-

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 odchyłek 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), $4.5 = \pm\text{średnia cyfr } 0-9$

(93)

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

(94)

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

(95)

4.24.3

Metodologia badania

^ Próbka: 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 odchyłek 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)

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

(105)

4.26.3

Metodologia badania

^ Próbka: 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 odchyłek

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 stającej matematycznej.

5.1.1

Badania wcze±niejsze

Bailey, Borwein i Crandall (2006) przeprowadzili analizę w³a±ciwo±ci statystycznych rozwini±c dziesi±tnych stających matematycznych, w tym ·, na próbkach rz±du miliona cyfr. Ich wyniki wskazywa³y na wysok± losowo±ç statystycznej w podstawowych testach.

5.1.2

Nasze wyniki w kontek±cie literatury

Wyniki naszej analizy na próbce 10 miliardów cyfr potwierdzaj± wnioski z wcze±niejszych badań dotyczące wysokiej losowo±ci statystycznej · w podstawowych aspektach. Jednocze±nie, wiêksza próbka pozwoli³a na wykrycie subtelnych struktur matematycznych w zaawansowanych testach, które nie by³y widoczne w mniejszych próbkach.

5.2

Spójno±ç wyników

Wyniki naszej analizy s¡ spójne z wcze±niejszymi 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 kryptogra

czne

Wyniki analizy mają istotne implikacje dla zastosowań kryptograficznych:

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

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

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

^ Testy Random Excursions (13, 14): Wykryto krytyczne struktury matematyczne ±rednie liczby wizyt w stanach spaceru losowego s; o 2-3 rz±dy wielko±ci odbiegaj;ce od oczekiwanych warto±ci. Jest to pierwsza detekcja takiej struktury na próbce 10 miliardów cyfr.

^ Testy SmallCrush (18, 21, 23, 27): Wykryto struktury w rozk;adzie odst;ipów, kombinacji i warto±ci ekstremalnych, wskazuj;ce na granice losowo±ci na du»ej skali.

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

6.3

Porównanie z wcze±niejszymi badaniami

Wyniki naszej analizy s; zgodne z badaniami przedstawionymi w arXiv:2504.10394 (2025), które równie» wskazuj; na granice losowo±ci · na du»ych skalach. Podczas gdy wcze±niejsze analizy na mniejszych próbkach (rz!du miliona cyfr) sugerowa;y idealn; losowo±c, nasza analiza na próbce 10 miliardów cyfr ujawnia subtelne struktury matematyczne charak

