MC-Proj1

Alicja Wiączkowska

2024-11-22

Contents

Testy

W celu zbadania czy dany generator zwraca liczby pseudolosowe, które można interpretować jako realizacje rozkładu jednostajnego na pewnym zbiorze - uzyskane liczby poddaje się testom zrwacającym wynik: p-wartość. Zazwyczaj przy ustalonym poziomie istotności, który w tym projekcie został ustalony jako $\alpha = 0.05$ hipotezę zerową o pochodzeniu liczb pseudolosowych z zadanego rozkładu odrzuca się przy uzyskaniu p-wartości poniżej α. Jest to tzw. pierwszopoziomowe (*first level testing*).

Warto jednak zauważyć, że przy prawdziwych realizacjach liczb losowych z rozkładu jednostajnego, p-wartości są liczbmi losowymi o rozkładzie $\mathcal{U}(0,1)$. Można zatem sprawdzić również hipotezę o losowości i jednostajnym rozkładzie p-wartości uzyskanych przy wielokrotnym powrórzeniu *first level testing*. Ostatecznym wynikiem będzie p-wartość wynikająca z przeprowadzenia tzw. testowania drugopoziomowego (*second level testing*).

Przy testowaniu pierwszopoziomowym test przeprowadzano na próbie wielkości $n = 2^{20}$.

Frequency Monobit Test

Test ten opiera się na badaniu częstości występowania zer i jedynek w zadanym ciągu bitów. W pierwszym kroku każdy z bitów b_i jest przekształcany do elementu zbioru $\{-1, 1\}$ funkcją $x_i = 2b_i + 1$. Statystyka testowa ma postać

$$T_n^{obs} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i$$

Z Centralnego Twierdzenia Granicznego przy rosnącym n statystyka ta zbiega według rozkładu do $\mathcal{N}(0,1)$, z czego można wywnioskować, że p-wartość w przybliżeniu wynosi

$$p_{val} := \mathbb{P}(|T_n^{teoret}| > |T_n^{obs}|) \approx \mathbb{P}(|N| > |T_n^{obs}|) = 2 \cdot \big(1 - \phi(|T_n^{obs}|)\big)$$

Second-level testing

Zauważmy, że dla prawdziwie losowej próby p-wartość otrzymana w teście jest zmienną loswą o rozkładzie $\mathcal{U}(0,1)$. Każdą liczbę z przedziału $(0,1)$ możemy przedstawić w systemie binarnym jako

$$p_{val} = 0.b_1b_2b_3b_4... := \sum_{i=1}^{\infty} b_i \cdot 2^{-i},$$

gdzie dla każdego $i \in \{1, 2, 3...\}$ zachodzi $b_i \in \{0, 1\}$. Weźmy obcięcie tego szeregu do długości n.

$$p_{val}^n = 0.b_1b_2b_3b_4...b_n := \sum_{i=1}^{n} b_i \cdot 2^{-i}.$$

Można pokazać indukcyjnie, że gdy $p_{val} \sim \mathcal{U}(0,1)$ to $\{b_i\}_{i=1}^n$ jest ciągiem losowych bitów dlugości n.

$$1^o \quad p_{val} \sim \mathcal{U}(0,1) \implies \mathbb{P}(p_{val} < 2^{-1}) = \mathbb{P}(b_1 = 0) = \frac{1}{2} = \mathbb{P}(p_{val} \geq 2^{-1}) = P(b_1 = 0)$$

$$2^o \quad założenie : \mathbb{P}(b_{i-1} = 0) = \frac{1}{2}, \qquad teza : \mathbb{P}(b_i = 0) = \frac{1}{2}$$

$$b_i = 1 \iff p_{val} - \sum_{k=1}^{i} b_k \cdot 2^{-k} \geq 2^{-i}$$

wartość różnicy $p_{val} - \sum_{k=1}^{i} b_k \cdot 2^{-k}$ należy do przedziału $\left[0, 2^{-(i-1)}\right]$, a ponieważ p_{val} była rozłożona jednostajnie, to $P(b_i = 1) = \frac{1}{2}$. $\mathbb{C}.\mathbb{K}.\mathbb{D}.$

Test χ^2

Test ten opiera się na podzieleniu zaobserwowanych wartości na k przedziałów (kategorii, koszyków, kubełków) oraz porównywaniu faktycznej liczby wartości wpadających do każdego z koszyków. Przy założeniu, że obserwowane liczby pseudolosowe są niezależne, statystyka testowa jest postaci:

$$\hat{\chi}^2 = \sum_{i=1}^{k} \frac{(Y_i - np_i)^2}{np_i},$$

gdzie Y_i - liczba obserwacji, które znalazły się w i-tej kategorii; n - liczba wszystkich wygenerowanych numerów pseudolosowych, p_i - teoretyczne. prawdopodobieństwo wpadnięcia pojedynczej obserwacji do i-tego koszyka. Statystyka $\hat{\chi}^2$ ma rozkład χ^2 z $(k-1)$ stopniami swobody.

Test Pokerowy

Test pokerowy w pierwszym kroku wymaga przyporządkowania obserwacji do k kubełków (jeśli generator zwraca liczby naturalne ze zbioru$\{0, 1, 2, ..., M-1\}$, za osobny kubełek możemy uznać każdą z możliwych do uzyskania wartości), które można interpretować jako karty. Następnie dany ciąg kart należy podzielić w piątki, a następnie każdej z nich przyporządkować układ analogiczny do tych z gry w pokera. Badane jest ile piątek spełnia jeden z poszczególnych układów pokerowych: 5 różnych kart, 4 różne karty (1 para), 3 różne karty (2 pary lub trójka), 2 różne karty (full lub kreta), 1 rodzaj karty.

Następnie uzyskane wyniki są porównywane z przewidywanymi częstotliwościami teretycznymi w teście χ^2.

Test Kolmogorowa-Smirnowa

Test Kolmogorowa-Smirnowa wiąże się z badaniem rozkładów ciągłych. Jest to test opierający się na porównywaniu dystrybuanty empirycznej z teoretyczną. Dystrybuantę empiryczną wyznaczyć można na podstawie obserwacji $X_1, X_2, ..., X_n$ korzystając ze wzoru $\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i < x)$. Statystyka testowa to

$$\hat{D}_n = \sqrt{n} \cdot \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)|.$$

Przy n dążącym do nieskończoności rozkład $\hat{D}_n$ zbiega do znanego tzw. rozkładu Kołogorowa - Smirnowa.

Generatory

LCG

LCG jest jednym z najprostszych generatorów liczb pseudolosowych ze zbioru liczb naturalnych $\{0, ..., M-1\}$. Generowanie ciągu liczb $x_1, x_2..., x_N$ z $LCG(M, a, c)$ przy zadanym ziarnie x_0 opiera się na algorytmie:

$$x_n = (a \cdot x_{n-1} + c) \ mod \ M.$$

Rozważone zostaną dwa przykłady tego generatora $LCG(2**10, 1, 5)$ oraz $LCG(2^{10}, 3, 7)$. Testy zostaną przeprowadzone na próbie 2^{20} liczb pseudolosowych. Za ziarno przyjęto $x_0 = 0$.

LCG(13, 1, 5)

Nie jest to dobry generator liczb pseudolosowych. Dla ziarna $x_0 = 0$ okres generatora wynosi $M = 13$, i symuluje powtarzającą się sekwencja liczb: 0, 5, 10, 2, 7, 12, 4, 9, 1, 6, 11, 3, 8.Każda z liczb występuje dokładnie 1 raz, co oznacza pewną jednorodność, jednak generator jest przewidywalny: kolejny wynik ściśle zależy od poprzeniego i nie spełnia założeń losowości.

Test χ^2 Najpierw wygenerowano $n = 2^{20}$ liczb pseudolosowych. W pierwszopoziomowym teście przy liczbie grup $k = M = 13$ generator $LCG(13, 1, 5)$ uzyskuje w teście χ^2 p-wartść równą 1, ponieważ każda z liczb ze zbioru $\{0, 1, ..., M - 1\}$ występuje w każdej sekwencji dokładnie raz (w dodatku w ustlonej kolejności).

Ze względu na ograniczenia pamięci doświadczenie testowania $n = 2^{17}$ powtórzono $R = 10^3$ razy, a następnie na uzyskanym wektorze p-wartości przeprowadzono test χ^2. Test wskazał finalną p-wartość poniżej $2.2 \cdot 10^{-16}$ zarówno dla liczby kubełków $k = 13$ jak i $k = 10$. Wynika to z faktu, że wszystkie p-wartości wskazują wartość 1, co zdecydowanie zaprzecza tezie o losowości i jednostajnym rozkładzie pierwszopoziomowych p-wartości na zbiorze $[0, 1]$. Według testu χ^2, $LCG(13, 1, 5)$ nie można uznać za dobry generator liczb pseudolosowych.

Test Pokerowy Ciąg liczb poddany testowi pokerowemu musi mieć długość podzielną przez 5. Pierwszopoziomowe testownie wykonano zatem na próbie $n = 2^{20} - 1 = 1.048575 \times 10^6$ liczb pseudolosowych. Już w testowaniu pierwszopoziomowym uzyskano p-wartość zbliżoną do zera. Wynika to z m.in. faktu, że generator $LCG(13, 1, 5)$ nie zwrócił żadnego układu 5 ani 4 takich samych "kart".

W testowaniu drugopoziomowym ze względu na ograniczenia pamięci doświadczenie testownia próby wielkości $n = 2^{15} - 3 = 3.2765 \times 10^4$ powtórzono $R = 10^3$ razy. Również w drugopoziomowym teście otrzymano p-wartość zbliżoną do zera. Praktycznie wszystkie układy p-wartości zostały zakwalifikowane jako "5 takich samych kart", co oznacza, że p-wartości były sobie bardzo bliskie (i wynosiły w przybliżeniu 0). Test pokerowy jednoznacznie stwierdza, że $LCG(13, 1, 5)$ nie jest dobrym generatorem liczb losowych.

Test Kolmogorowa-Smirnowa W teście Kolmogorowa-Smirnowa odpowiednią praktyką jest wykorzystanie niepowtarzających się liczb, dlatego wielkość próby w pierwszopoziomowym teście może wynieść maksymalnie $n = 13$. Będziemy badać czy liczby wygenerowane przez $LCG(13, 1, 5)$ po znormalizowaniu dzieleniem przez $M = 13$ można traktować jako realizacje rozkładu $\mathcal{U}(0, 1)$.

Wyniki pierwszopoziomowego testu przedstawiono w tabeli. Ponieważ każda z liczb wygenerowanych przez LCG występuje tyle samo razy, przy tak niewielu liczbach pseudolosowych otrzymano bardzo wysokie p-wartości.

Table 1: LCG(13,1,5) - first level testing

n	statystyka testowa	p-wartość
1	0.6153846	0.7692308
2	0.3846154	0.8550296
3	0.2820513	0.9262631
4	0.2307692	0.9519417
5	0.1846154	0.9833437
6	0.1538462	0.9943360
7	0.1648352	0.9739857
8	0.1153846	0.9993685
9	0.1282051	0.9935738
10	0.1076923	0.9987822
11	0.0979021	0.9994631
12	0.0769231	0.9999928
13	0.0769231	0.9999794

Natomiast w drugopoziomowym testowaniu finalne p-wartości w każdym przypadku były zbliżone do zera, co było spowodowane powtarzalnością sekwencji liczb, a w skutku otrzymywaniem podobnych p-wartości w *first level testing.*

LCG(2^{10}, 3, 7)

Okres tego generatora jest znacznie duższy niż w poprzednim przypadku. Dla $x_0 = 0$ wynosi on 511. Przy próbkach mniejszych niż 511 uzyskane liczby pseudolosowe mogą wyglądć na relizacje rozkładu jednostajnego na $\{0, 1, ..., M-1\}$, jednak nadal generator ten nie przejdzie wszystkich testów.

Test χ^2 Najpierw wygenerowano $n = 2^{20}$ liczb pseudolosowych. W pierwszopoziomowym teście przy liczbie grup $k = 10$ generator $LCG(2^{10}, 3, 7)$ uzyskuje p-wartść mniejszą niż $2 \cdot 10^{-16}$, podbnie dla $k = 32$. Wynika to z faktu, że generator mając okres długości 511 może wygenerować tylko tyle różnych wartości ze zbioru 2^{10}-elementowego.

Ze względu na ograniczenia pamięci doświadczenie testowania $n = 2^{13}$ powtórzono $R = 10^3$ razy, a następnie na uzyskanym wektorze p-wartości przeprowadzono test χ^2. Test wskazał finalną p-wartość poniżej $2 \cdot 10^{-16}$ zarówno dla liczby kubełków $k = 32$ jak i $k = 10$. Wynika to z faktu, że wszystkie p-wartości wskazują wartość bliską zeru, co zdecydowanie zaprzecza tezie o losowości i jednostajnym rozkładzie pierwszopoziomowych p-wartości na zbiorze $[0, 1]$. Według testu $LCG(2^{10}, 3, 7)$ nie można uznać za dobry generator liczb pseudolosowych.

Test Pokerowy Pierwszopoziomowe testownie wykonano na próbie $n = 2^{20} - 1 = 1.048575 \times 10^6$ liczb pseudolosowych. Już w testowaniu pierwszopoziomowym uzyskano p-wartość zbliżoną do zera. Wynika to z m.in. faktu, że generator $LCG(2^{10}, 3, 7)$ nie zwrócił żadnego układu 5 takich samych "kart".

W testowaniu drugopoziomowym ze względu na ograniczenia pamięci doświadczenie testownia próby wielkości $n = 2^{15} - 3$ powtórzono $R = 10^3$ razy. Również w drugopoziomowym teście otrzymano p-wartość mozno zbliżoną do zera. Praktycznie wszystkie układy p-wartości zostały zakwalifikowane jako "5 takich samych kart", co oznacza, że p-wartości z *first level testing* były sobie bardzo bliskie (i wynosiły w przybliżeniu 0). Test pokerowy jednoznacznie stwierdza, że $LCG(2^{10}, 3, 7)$ nie jest dobrym generatorem liczb losowych.

Test Kolmogorowa-Smirnowa W teście Kolmogorowa-Smirnowa odpowiednią praktyką jest wykorzystanie niepowtarzających się liczb, dlatego wielkość próby w pierwszopoziomowym teście może wynieść maksymalnie $n = 2^{10}$. Będziemy badać czy liczby wygenerowane przez $LCG(2^{10}, 3, 7)$ po znormalizowaniu dzieleniem przez $M = 2^{10}$ można traktować jako realizacje rozkładu $\mathcal{U}(0, 1)$.

Wyniki pierwszopoziomowego testu przedstawiono w tabeli.

Table 2: LCG(2^10,3,7) - first level testing

n	statystyka testowa	p-wartość
16	0.3867188	0.0115458
32	0.2304688	0.0561950
64	0.1992188	0.0105788
128	0.1601562	0.0028137
256	0.0947266	0.0202198
512	0.0019531	1.0000000
1024	0.0019531	1.0000000

Zauważmy, że gdy wielkość próby jest liczbą ze zbioru $\{2^9, 2^{10}\}$, $p - wartość$ wynosi 1, natomiast w pozostałych przypadkach przeważnie $p - wartość$ jest mniejsza niż poziom istotności $\alpha = 0.05$.

W drugopoziomowym testowaniu finalne p-wartości w każdym przypadku były zbliżone do zera. Zarówno *first level testing* jak i *second level testing* odrzucają tezę o losowości i jednostajnym rozkładzie pseudolosowych livczb pochodzących z generatora $LCG(2^{10}, 3, 7)$.

GLCG

Generator $GLCG$ jest pewnym rozszerzeniem LCG. Przyjmuje ziarno $x_0, x_1, ..., x_{k-1}$ i przy zadanych M i $\{a_i\}_{i=1}^k$polega na wyznaczeniu kolejnuch liczb pseudolosowych na podstawie rekurecji:

$$x_n = \left(\sum_{i=1}^{k} a_i \cdot x_{n-i} \right) \ mod\ M$$

Badaniu poddano generator $GLCG\left(M = 2^{10}, (a_1, a_2, a_3) = (3, 7, 68)\right)$ z ziarnem $(x_0, x_1, x_2) = (1, 2, 3)$. Ma on okres 287.

Test χ^2 Najpierw wygenerowano $n = 2^{20}$ liczb pseudolosowych. W pierwszopoziomowym teście gdy liczba grup k jest dzielnikiem liczby $M = 2^{10}$ generator uzyskuje w teście χ^2 p-wartść równą 1. *First level testing* wydaje się potwierdzać hipotezę o losowości wygenerowanych liczb (jest nawet podejrzanie za dobry).

Ze względu na ograniczenia pamięci w *second level testing* doświadczenie testowania $n = 2^{17}$ powtórzono $R = 10^3$ razy, a następnie na uzyskanym wektorze p-wartości przeprowadzono test χ^2. Test wskazał finalną p-wartość poniżej $2.2 \cdot 10^{-16}$. Wynika to z faktu, że wszystkie p-wartości wskazują wartość 1 lub bliską 1, co zdecydowanie zaprzecza tezie o losowości i jednostajnym rozkładzie pierwszopoziomowych p-wartości na zbiorze $[0, 1]$. Według testu χ^2, rozważanego $GLCG$ nie można uznać za dobry generator liczb pseudolosowych.

Test Pokerowy Ciąg liczb poddany testowi pokerowemu musi mieć długość podzielną przez 5. Pierwszopoziomowe testownie wykonano zatem na próbie $n = 2^{20} - 1$ liczb pseudolosowych. Już w testowaniu pierwszopoziomowym uzyskano p-wartość zbliżoną do zera. Jedną z przyczyn tego wyniku może być fakt, że w pseudolosowej sekwencji ułożenie w sekwencję 5 takich samych "kart" nie wystąpiło ani razu natomiast.

W testowaniu drugopoziomowym ze względu na ograniczenia pamięci doświadczenie testownia próby wielkości $n = 2^{15} - 3$ powtórzono $R = 10^3$ razy. Również w drugopoziomowym teście otrzymano ostateczną p-wartość zbliżoną do zera. Żaden z układów p-wartości nie został zakwalifikowane jako "5 różnych kart" ani "4 różne karty" ("1 para"), co oznacza, że pierwszopoziomowych p-wartości nie można uznać za realizację rozkładu $\mathcal{U}(0, 1)$. Test pokerowy jednoznacznie stwierdza, że rozważany $GLCG$ nie jest dobrym generatorem liczb losowych.

Test Kolmogorowa-Smirnowa W teście Kolmogorowa-Smirnowa odpowiednią praktyką jest wykorzystanie niepowtarzających się liczb, dlatego wielkość próby w pierwszopoziomowym teście może wynieść maksymalnie $n = 2^{10}$. Będziemy badać czy liczby wygenerowane przez $GLCG\left(2^{10}, (3, 7, 68)\right)$ po znormalizowaniu dzieleniem przez $M = 2^{10}$ można traktować jako realizacje rozkładu $\mathcal{U}(0, 1)$.

Wyniki pierwszopoziomowego testu przedstawiono w tabeliponiższej. Uzyskane p-wartości są wysokie. Nie dają podstaw do odrzucenia hipotezy o pochodzeniu wygenerowanych liczb z rozkładu jednostajnego.

Table 3: GLCG(2^10,(3,7,68)) - first level testing

n	statystyka testowa	p-wartość
8	0.2626953	0.5536684
16	0.1220703	0.9474059
32	0.1748047	0.2821799
64	0.1230469	0.2871538
128	0.0888672	0.2642377
256	0.0498047	0.5492352
512	0.0205078	0.9824477

n	statystyka testowa	p-wartość
1024	0.0136719	0.9909033

Natomiast wyniki testowania drugopoziomowego w większości przypadków prowadzą do odrzucenia hipotezy zerowej,choć nie zawsze, o czym świadczy drugopoziomowa p-wartość uzyskana dla $n = 2^9$.

Table 4: GLCG(2^10,(3,7,68)) - second level testing

n	statystyka testowa	p-wartość
8	0.0111406	0.9996570
16	0.0450365	0.0346160
32	0.0695999	0.0001240
64	0.0616161	0.0010079
128	0.0443008	0.0394789
256	0.0728420	0.0000492
512	0.0324818	0.2420117
1024	0.0944355	0.0000000

RC4(32)

Algorytm RC4(m) jest związany z kryptografią Polega on na zainicjalizoaniu kluczem pewnej permutacji liczb ze zbioru $\{0, 1, 2, ..., m-1\}$, a następnie powtarzania sekwecji dalszego permutowania, podczas którego jednocześnie wybierane są konkretne liczby pseudolosowe.

Second level testing można rozważać na 2 sposoby:

i) generując ciąg liczb długości $R \cdot n$

ii) powtarzając R - krotnie losowanie n liczb, ze zmienionym kluczem; kluczami będą kolejne podzbiory $[m]$.

Test χ^2 Najpierw wygenerowano $n = 2^{20}$ liczb pseudolosowych. W pierwszopoziomowym teście przy liczbie grup $k = M = 32$ generator $RC4(32)$ uzyskuje w teście χ^2 przyzwoitą p-wartść równą 0.2604485.

Ze względu na ograniczenia pamięci doświadczenie testowania $n = 2^{17}$ powtórzono $R = 10^3$ razy, a następnie na uzyskanym wektorze p-wartości przeprowadzono test χ^2. Test dla metody i) wskazał finalną p-wartość 0.9060686 , natomiast dla sposobu ii) wyniosła ona poniżej $2.2 \cdot 10^{-16}$. Sposób i) wydaje się lepszym na *second level testing*. Przy wybraniu go, test χ^2 nie daje powodów do odrzucenia hipotezy o poprawności generatora $RC4(32)$.

Test Pokerowy Ciąg liczb poddany testowi pokerowemu musi mieć długość podzielną przez 5. Pierwszopoziomowe testownie wykonano zatem na próbie $n = 2^{20} - 1 = 1.048575 \times 10^6$ liczb pseudolosowych. W pierwszopoziomowym testowaniu uzyskana p-wartość jest malutka: 1.8177377×10^{-4}. Może ona być podstawą do odrzucenia hipotezy zerowej.

W testowaniu drugopoziomowym ze względu na ograniczenia pamięci doświadczenie testownia próby wielkości $n = 2^{15} - 3 = 3.2765 \times 10^4$ powtórzono $R = 10^3$ razy.

Dla podejścia i) rówineż drugopoziomowy test zwraca małą p-wartość - wynoszącą 0.026564. Wynik ten osttecznie może prowadzić do odrzucenia hipotezy zerowej.

Natomiast dla podejścia drugiego uzyskana p-wartość była zaskakująco wysoka, bo aż 0.9614294.

Generator $RC4(32)$ niestetynie przeszedł testu pokerowego.

Test Kolmogorowa-Smirnowa W teście Kolmogorowa-Smirnowa odpowiednią praktyką jest wykorzystanie niepowtarzających się liczb, dlatego wielkość próby w pierwszopoziomowym teście może wynieść maksymalnie $n = 32$. Będziemy badać czy liczby wygenerowane przez $RC4(32)$ po znormalizowaniu dzieleniem przez $M = 32$ można traktować jako realizacje rozkładu $\mathcal{U}(0, 1)$.

Wyniki pierwszopoziomowego testu przedstawiono w tabeli. Dla każdej wielkości próby n generator $RC4(32)$ przechodzi test Kolmogorowa-Smirnowa.

Table 5: RC4(32) - KS - first level testing

n	statystyka testowa	p-wartość
5	0.2250000	0.9129625
6	0.2500000	0.7694830
7	0.1785714	0.9506613
8	0.1875000	0.8953831
9	0.2083333	0.8295531
10	0.2500000	0.5595598
11	0.2045455	0.7468258
12	0.2500000	0.4413066
13	0.2115385	0.6058127
14	0.1785714	0.7633831
15	0.1500000	0.8884604
16	0.1562500	0.8295531
17	0.1893382	0.5759022
18	0.2187500	0.3551549
19	0.2450658	0.2039069
20	0.2687500	0.1112375
21	0.2425595	0.1688860
22	0.2642045	0.0927052
23	0.2459239	0.1237980
24	0.2604167	0.0771411
25	0.2387500	0.1156572
26	0.2187500	0.1660100
27	0.2372685	0.0956590
28	0.2544643	0.0532377
29	0.2359914	0.0790997
30	0.2187500	0.1132533
31	0.2348790	0.0653947
32	0.2500000	0.0366311

Następnie przeprowadzono *second level testing* korzstając ze sposobu ii). Finalne p-wartości w każdym przypadku wyszły niesatysfakcjonująco małe. Nie jest to jednak słabość generatora, lecz problem ten został najprawdopodobniej spowodowany faktem, że prawdziwa dystrybuanta $RC4(32)$ nawet po zrzutowaniu liczb na przedział [0,1] nie dąży do dystrybuanty rozkładu $\mathcal{U}(0, 1)$, lecz jest schodkowa.

Table 6: RC4(32) - KS - second level testing

n	statystyka testowa	p-wartość
1	0.0620000	0.0009166
2	0.0890000	0.0000003
3	0.0395720	0.0872629
4	0.0585106	0.0021253

n	statystyka testowa	p-wartość
5	0.0450000	0.0348446
6	0.0534830	0.0065537
7	0.0518025	0.0093362
8	0.0953742	0.0000000
9	0.0590761	0.0018607
10	0.0861806	0.0000007
11	0.0937917	0.0000000
12	0.0820488	0.0000028
13	0.0933408	0.0000001
14	0.0805064	0.0000047
15	0.0767131	0.0000155
16	0.1701671	0.0000000
17	0.0701392	0.0001067
18	0.0487658	0.0171965
19	0.0697950	0.0001174
20	0.0824129	0.0000025
21	0.0610161	0.0011676
22	0.0633631	0.0006512
23	0.0503513	0.0125584
24	0.0872043	0.0000005
25	0.0664030	0.0002959
26	0.0467185	0.0254231
27	0.0670148	0.0002513
28	0.0682216	0.0001813
29	0.0399265	0.0824820
30	0.0556083	0.0041220
31	0.0659323	0.0003351
32	0.1903742	0.0000000

Mersenne Twister

Algorytm ten jest domyślnym generatorem liczb losowych w Pythonie lub R. Opiera się na przyjęciu za okres tzw. liczby pierwszej Mersenne'a. Generator został zaprojektowany z myślą o metodach Monte Carlo i innych symulacjach statystycznych. Jest on uważany za szybki i skuteczny generator liczb pseudolosowych.

Test χ^2 Najpierw wygenerowano $n = 2^{20}$ liczb pseudolosowych. W pierwszopoziomowym teście dla $k = 16$ uzyskano p-wartość 0.1379828. Dla $k \in \{32, 64\}$ uzyskano p-wartość równą 1. Wyniki te nie prowadzą do odrzucenia hipotezy o losowości i jednostajnym rozkładzie uzyskanych liczb. Jednkże p-wartość uzyskana dla $k = 10$ wyniosła 0.843412, co jest bardzo małą liczbą.

Ze względu na ograniczenia pamięci w *second level testing* doświadczenie testowania $n = 2^{17}$ powtórzono $R = 10^3$ razy, a następnie na uzyskanym wektorze p-wartości przeprowadzono test χ^2 z $k = 10$ i $k = 32$ kubełkami. Niestety test wskazał, że finalne p-wartość wyniosły poniżej $2.2 \cdot 10^{-16}$. Problem ten wynika prawdopodobnie z nieoptyalnego doboru kubełków.

Test Pokerowy Ciąg liczb poddany testowi pokerowemu musi mieć długość podzielną przez 5. Pierwszopoziomowe testownie wykonano zatem na próbie $n = 2^{20} - 1$ liczb pseudolosowych. W *first level testing* otrzymano p-wartość 0.3132981, kóra nie nie daje podstaw do zaprzeczenia hipotezie zerowej o losowości i jednostajnym rozkładzie wygenerowanych liczb.

W testowaniu drugopoziomowym ze względu na ograniczenia pamięci doświadczenie testownia próby wielkości $n = 2^{15} - 3$ powtórzono $R = 10^3$ razy. Tym razem uzyskano dużą p-wartość wynoszącą0.6775469, co nie odrzuca hipotezy zerowej. Jest to spodziewany wynik.

Test pokerowy zatem nie wykrywa żadnych nieprawidłowości w generatorze MT.

Test Kolmogorowa-Smirnowa Algorytm Mersenne Twister bardzo dobrze poradził sobie z tym testem. Wyniki pierwszopoziomowego testu przedstawiono w tabeli poniższej. Uzyskane p-wartości są dostatecznie wysokie i dają podstaw do odrzucenia hipotezy o pochodzeniu wygenerowanych liczb z rozkładu jednostajnego.

Table 7: MT - KS test - first level testing

n	statystyka testowa	p-wartość
8	0.3819492	0.1473685
16	0.3063701	0.0787623
32	0.1103569	0.7906369
64	0.0813731	0.7597068
128	0.0799734	0.3861470
256	0.0502713	0.5370970
512	0.0482675	0.1839129
1024	0.0225276	0.6762480

Natomiast wyniki testowania drugopoziomowego w większości przypadków również są satysfakcjonujące, żadna z nich nie prowadzi do odrzucenia hipotezy zerowej. Szczegóły są widoczne poniżej.

Table 8: MT - KS test - second level testing

n	statystyka testowa	p-wartość
8	0.0348931	0.1750687
16	0.0260830	0.5043443
32	0.0265483	0.4813591
64	0.0271980	0.4501448
128	0.0276999	0.4267822
256	0.0219859	0.7191173
512	0.0278749	0.4188004
1024	0.0377958	0.1148549

Liczby niewymierne

W tej części pracy zbadane zostanie czy kolejne bity rozwinięcia dwójkowego liczb niewymiernych: π, e oraz $\sqrt{2}$ można traktować jako pewien generator liczb pseudolosowych. Wykorzystantego zostanie *Frequency Monobit* test służący do badania losowości ciągów bitów.

Bity π - Frequency Monobit Test

Dla pierwszych $n = 1004882$ wyrazów liczby π w zapisie bitowym p-wartość uzyskana we *Frequency Monobit* Teście wyniosła 0.6212069. Wynik ten sugeruje, że nie ma podstaw by odrzucać hipotezę o losowości badanych bitów.

Podobnie przy rozpatrzeniu mniejszej próby pierwszych $n = 1004$ bitów z rozwinięcia liczby π otrzymano p-wartość 0.4874854, która również nie zaprzecza hipotezie o losowości i jednostajnym rozkładzie bitów.

Ponieważ plik źródłowy zawiera jedynie 1004882 bitów, podczas przeprowadzania *second-level testing* doświadczenie testowania ciągu długości $n = 1004$ powrórzono $R = 10^3$ razy. Następnie aby na uzyskanych w ten sposób p-wartościach wykonać *Frequency Monobit Test*, każdą z nich zapisano jako szereg bitów i wybrano pierwsze k z nich (przekształcając funkcją f_k):

$$f_k : [0,1] \to \{0,1\}^k \qquad f_k(p_{val}) = f_k\left(\sum_{i=1}^{\infty} b_i \cdot 2^{-i}\right) = (b_1, b_2, .., b_k),$$

a następnie uzyskane w ten sposób wektory połączono jeden, który poddano testowaniu. Wyniki dla wybranych k przedstawiono w poniższej tabeli.

Table 9: second level testiong - pi

k - długość 1 wektora bitów	statystyka testowa	p-wartość
1	1.420000	0.155608
2	-6.477098	9.35034e-11
3	-6.916656	4.6243e-12
4	7.470000	8.01581e-14
5	9.606148	0.0000000
6	12.361758	0.0000000
7	2.736463	0.00621036
8	-1.979899	0.0477149
9	-22.800000	0.0000000
10	-2.536147	0.011208
11	20.080655	0.0000000

Uzyskane p-wartości w testowniu drugopoziomowym są bardzo małe, w niektórych przypadkach autmatycznie zaokrąglone do zera. Przy ustalonym poziomie istotności $\alpha = 0.05$ hipoteza zerowa zakładająca, że p-wartości pochodzące z pierwszopoziomowego testowania pochodzą z rozkładu $\mathcal{U}(0,1)$ zostanie odrzucona dla wszystkich k.

Również po poddaniu testowi χ^2 z 10 jednakowymi kubełkami badanych p-wartości pochodzących z pierwszopoziomowego testowania *Frequency Monobit* testem, otrzymano ostateczną p-wartość wynoszącą mniej niż $2.2 \cdot 10^{-16}$.

Można zatem stwierdzić, że również w tym przypadku, mimo dobrych wyników uzyskanych podczas *first level testing*, ostatecznie należy uznać, że bitów liczby π nie można traktować jako dobry generator liczb z rozkładu jednostajnego na zbiorze {0,1}, co wynika jednoznacznie z *second-level testing*.

Bity e - Frequency Monobit Test

Dla pierwszych $n = 1004882$ wyrazów liczby e w zapisie bitowym p-wartość uzyskana we *Frequency Monobit* Teście była wysoka - wyniosła 0.9284114. Wynik ten nie daje podstaw do stwierdzenia, że bity pochodzące z rozwnięcia dwójkowego e nie są losowymi bitami rozłożonymi jednostajnie na {0,1}.

Przy rozpatrzeniu mniejszej próby pierwszych $n = 1004$ bitów z rozwinięcia liczby e otrzymano znacznie mniejszą p-wartość 0.1145688, która jednak przy istotności $\alpha = 0.05$ nie prowadzi do odrzucenia hipotezy o pochodzeniu rozważanych bitów z rozkładu jednostajnego na {0,1}.

Ponieważ plik źródłowy zawiera jedynie 1004882 bitów, podczas przeprowadzania *second-level testing* doświadczenie testowania ciągu długości $n = 1004$ powrórzono $R = 10^3$. Następnie aby na uzyskanych w ten

sposób p-wartościach wykonać *Frequency Monobit Test*, każdą z nich zapisano jako szereg bitów i wybrano pierwsze k z nich (jak wcześniej przekształcając funkcją f_k),a następnie uzyskane w ten sposób wektory połączono jeden, na którym przeprowadzono test. Wyniki dla wybranych k przedstawiono w poniższej tabeli.

Table 10: second level testiong - e

k - długość 1 wektora bitów	statystyka testowa	p-wartość
1	3.800000	0.000144696
2	-5.388154	7.11852e-08
3	-6.974391	3.07199e-12
4	6.880000	5.98521e-12
5	9.105269	0.0000000
6	12.810831	0.0000000
7	2.910326	0.00361051
8	-2.114249	0.034494
9	-23.806667	0.0000000
10	-2.706910	0.00679127
11	19.761053	0.0000000

Ostatecznie uzyskane p-wartości są małe, w niektórych przypadkach tak małe, że zostały autmatycznie zaokrąglone do zera. Przy ustalonym poziomie istotności $\alpha = 0.05$ Hipoteza zerowa o losowości p-wartości pochodzących z pierwszopoziomowego testowania zostanie odrzucona we wszystkich przypadkach poza $k = 1$.

Ponadto gdy p-wartości otrzymane podczas pierwszopoziomowego testowania poddano testowi χ^2 z 10 jednakowymi kubełkami otrzymano ostateczną p-wartość wynoszącą mniej niż $2.2 \cdot 10^{-16}$.

Można zatem stwierdzić, że mimo dobrych wyników uzyskanych podczas *first level testing*, ostatecznie należy uznać, że bitów liczby e nie można traktować jako dobry generator liczb z rozkładu jednostajnego na zbiorze $\{0,1\}$.

Bity $\sqrt{2}$ - Frequency Monobit Test

Dla pierwszych $n = 1004883$ wyrazów liczby $\sqrt{2}$ w zapisie bitowym po przeprowadzeniu *Frequency Monobit* Tesu, uzyskana p-wartość wyniosła aż 0.8215235 co sugeruje, że nie ma podstaw do odrzucenia hipotezy o pochodzeniu badanych bitów z rozkładu jednostajnego na $\{0,1\}$.

Podobnie przy rozpatrzeniu mniejszej próby pierwszych $n = 1004$ bitów z rozwinięcia liczby $\sqrt{2}$ otrzymano p-wartość 0.6586078, która również nie zaprzecza hipotezie o losowości bitów.

Ponieważ plik źródłowy zawiera jedynie 1004883 bitów, podczas przeprowadzania *second-level testing* doświadczenie testowania ciągu długości $n = 1004$ powrórzono $R = 10^3$. Następnie każdą p-wartości z testowania pierwszopoziomowego przekształcono funkcją f_k, a następnie uzyskane w ten sposób wektory połączono jeden, na którym przeprowadzono *Frequency Monobit Test.* Wyniki dla wybranych k przedstawiono w poniższej tabeli.

Table 11: second level testiong - sqrt 2

k - długość 1 wektora bitów	statystyka testowa	p-wartość
1	2.9600000	0.00307639
2	-5.5578593	2.73103e-08
3	-6.4085880	1.46873e-10
4	7.3400000	2.13607e-13

k - długość 1 wektora bitów	statystyka testowa	p-wartość
5	10.1696372	0.0000000
6	13.4803586	0.0000000
7	4.4070658	1.0478e-05
8	-0.8202439	0.412077
9	-21.4866667	0.0000000
10	-0.3478505	0.727952
11	22.5711393	0.0000000
12	14.4222097	0.0000000
13	37.6807843	0.0000000
14	27.6936099	0.0000000
15	33.0649498	0.0000000

Ostatecznie dla większości k uzyskane p-wartości przeważnie są bardzo małe, a często autmatycznie zaokrąglone przez komputer do zera, choć dla $k = 8$ p-wartość przekroczyła 0.4, a dla $k = 10$ wyniosła aż ponad 0.7. Przy ustalonym poziomie istotności $\alpha = 0.05$ w większości przypadków hipoteza zerowa zostanie jednak odrzucona.

Ponadto gdy p-wartości otrzymane podczas pierwszopoziomowego testowania poddano testowi χ^2 z 10 jednakowymi kubełkami, finalna p-wartość wyniosła mniej niż $2.2 \cdot 10^{-16}$.

Można zatem stwierdzić, że mimo dobrych wyników uzyskanych podczas *first level testing*, ostatecznie należy uznać, że bitów liczby $\sqrt{2}$ nie można traktować jako dobry generator liczb z rozkładu jednostajnego na zbiorze {0,1}.

Źródła

https://ipsec.pl/files/ipsec/ving-krypto.pdf str 7

"Theory and Practice of Monte Carlo Methods" Paweł Lorek Tomasz Rolski

https://pl.wikipedia.org/wiki/RC4

dokumentacja R

https://pl.wikipedia.org/wiki/Mersenne_Twister