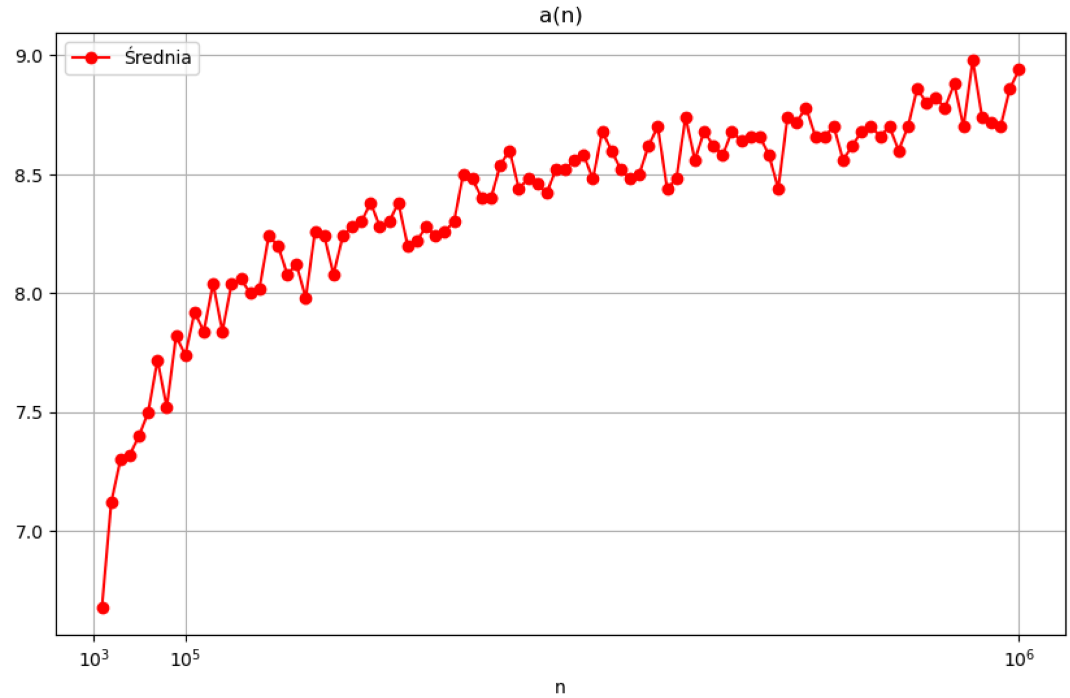
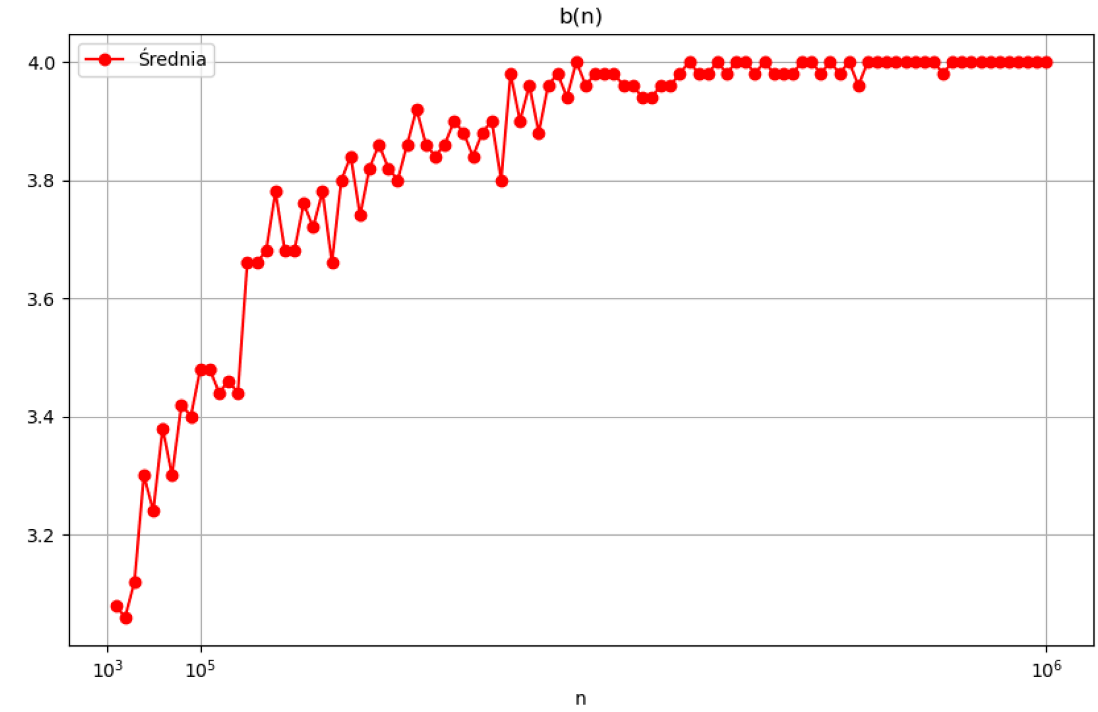
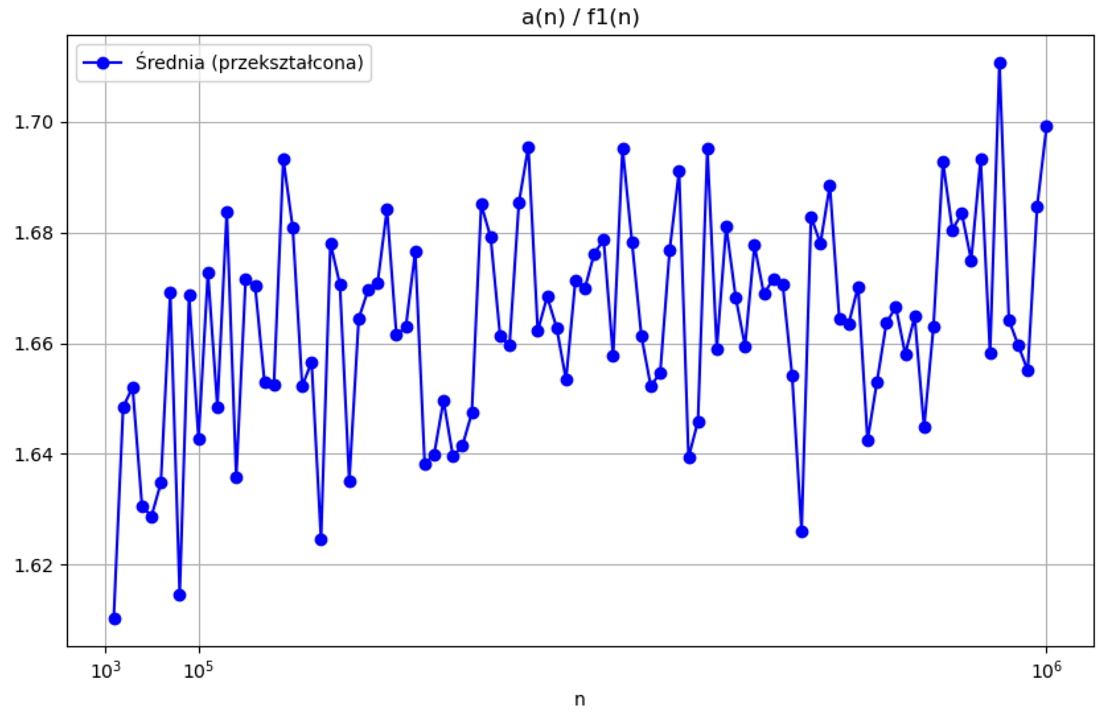
Wojciech Typer
279730

zadanie 1
Kody źródłowe zostały napisane w języku Rust. Wykresy zostały wygenerowane w Pythonie poprzez bibliotekę Matplotlib. Użyto generatora liczb pseudolosowych Mersenne Twister. Każdy kod źródłowy generuje dane liczbowe, a następnie zapisuje je do pliku tekstowego, z którego następnie generowany jest wykres bądź histogram.

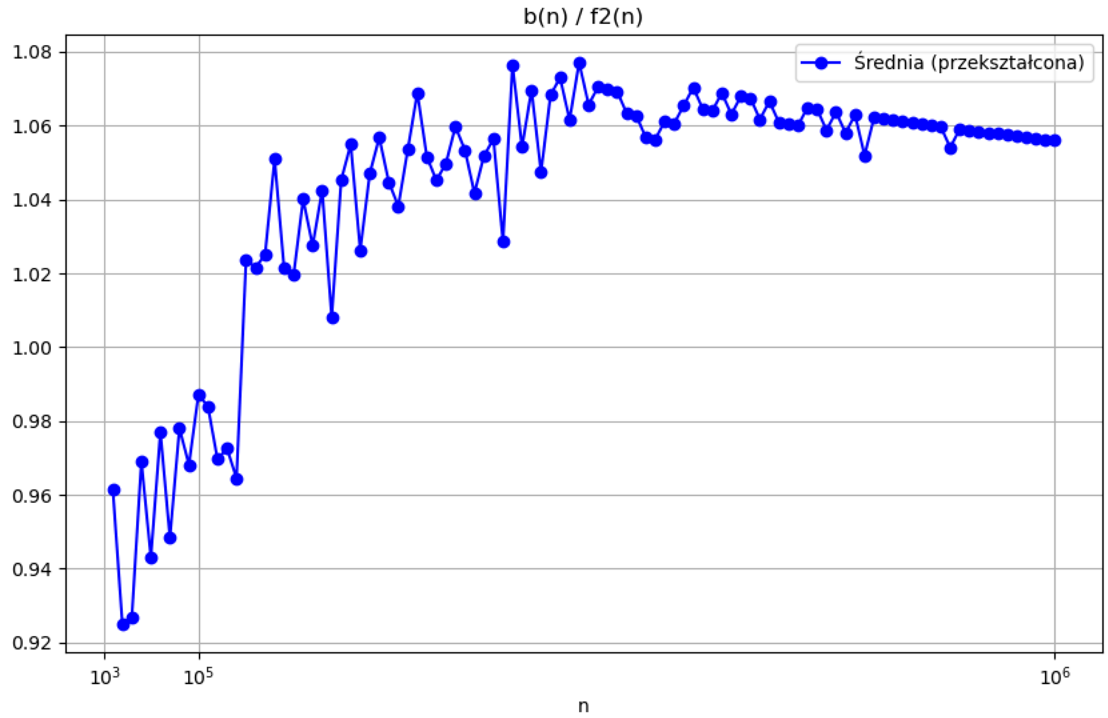




b(n)
Średnia
4.0
3.8
3.6
3.4
3.2
10^3
10^5
10^6
n



a(n) / f1(n)
Średnia (przekształcona)
1.70
1.68
1.66
1.64
1.62
10^3
10^5
10^6
n

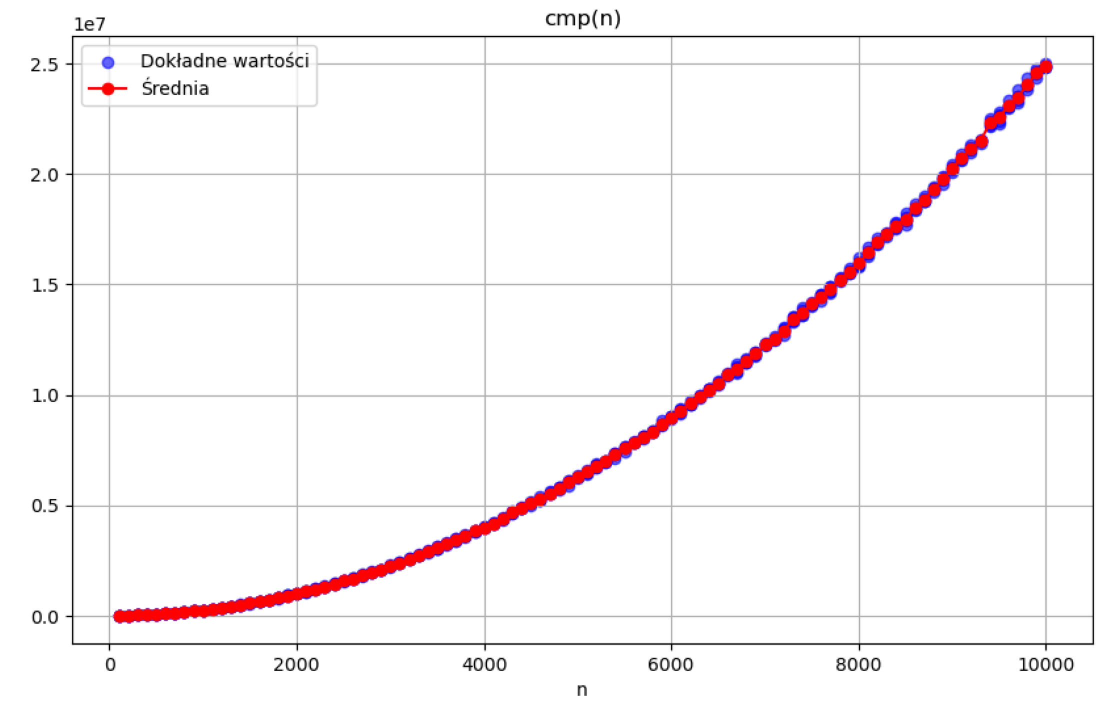


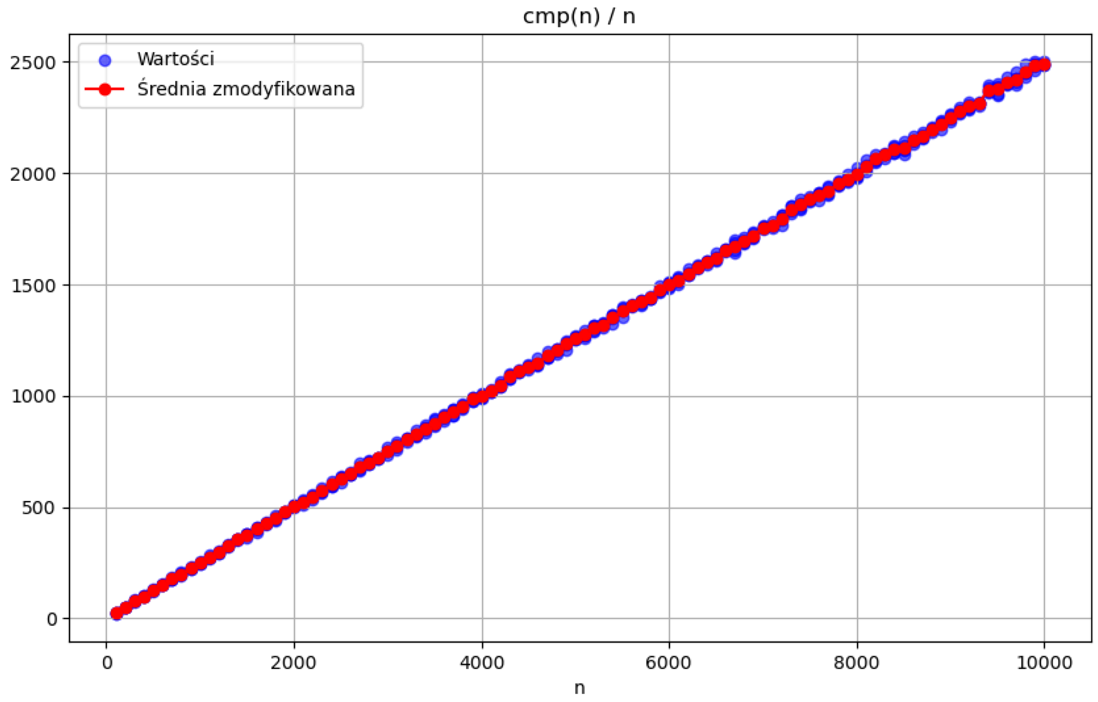
gdzie $f_1(n) = \dfrac{\ln n}{\ln \ln n}$ i $f_2(n) = \dfrac{\ln \ln n}{\ln 2}$

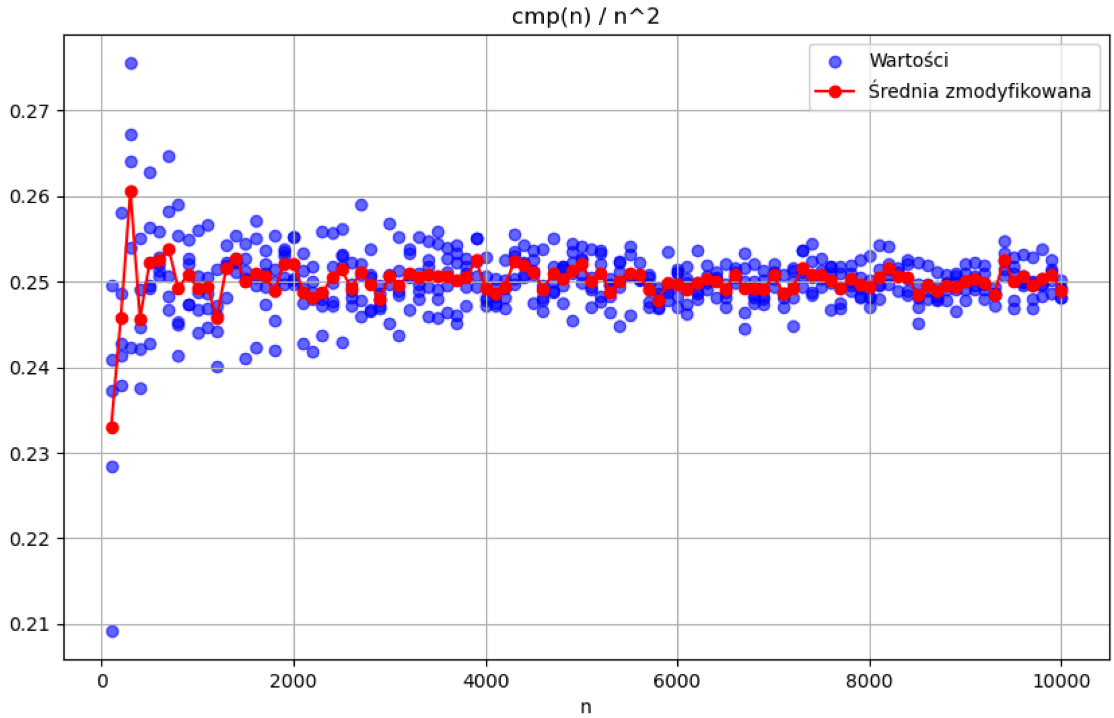
Z wykresu a(n), który pokazuje średnią wartość zapełnienia urny z największą ilością kulek, widzimy, że wraz ze wzrostem n - liczby urn i liczby losowań, rośnie średnia ilość kulek w urnie z największą ich ilością (z wykresu a(n) / f1(n) możemy odczytać, że asymptotyczne tempo wzrostu a(n) to w przybliżeniu ln n / (ln ln n)). Dla coraz większych wartości n, wartości średnie mniej się rozpraszają, ich fluktuacja jest większa.

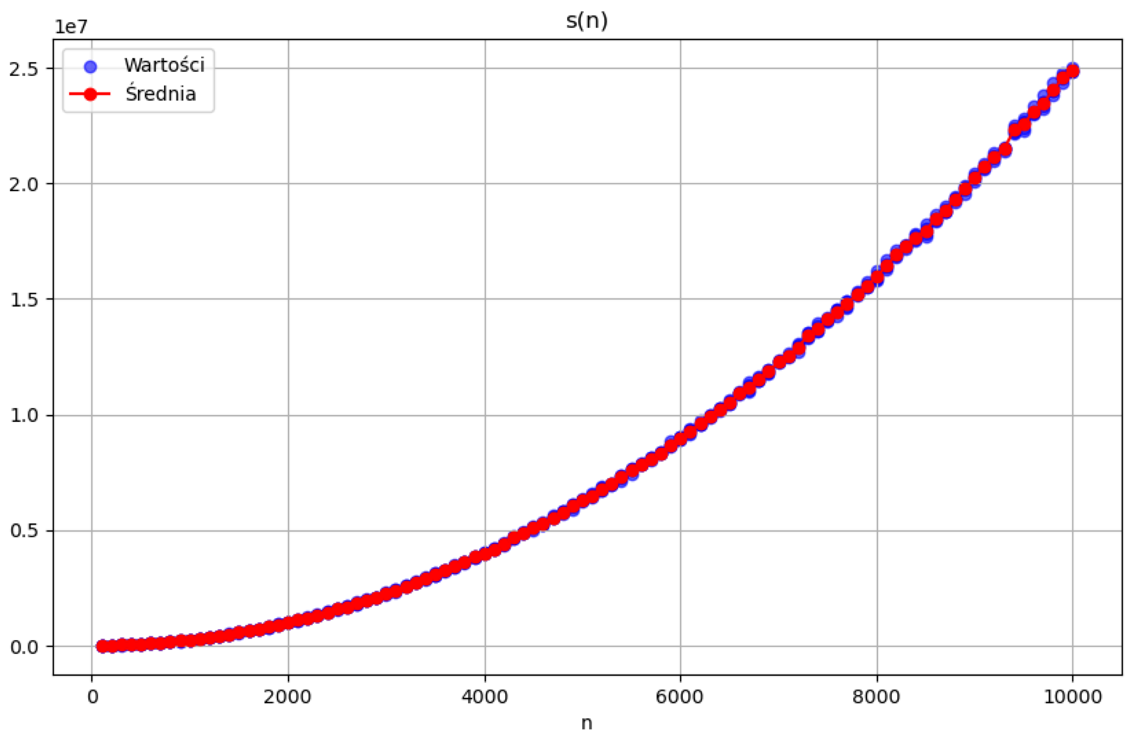
Z wykresu b(n) możemy odczytać, że dla coraz większych wartości n, średnia ilość kulek w urnie się zwiększa, lecz nie przekracza 4.0. Dla większych n, średnie wartości układają się bardzo blisko siebie, natomiast dla mniejszych n, średnie wartości rosną szybko (tak jak w przypadku a(n)). Z wykresu b(n) / f2(n) możemy odczytać, że asymptotyczne tempo wzrostu to w przybliżeniu (ln ln n) / ln 2.

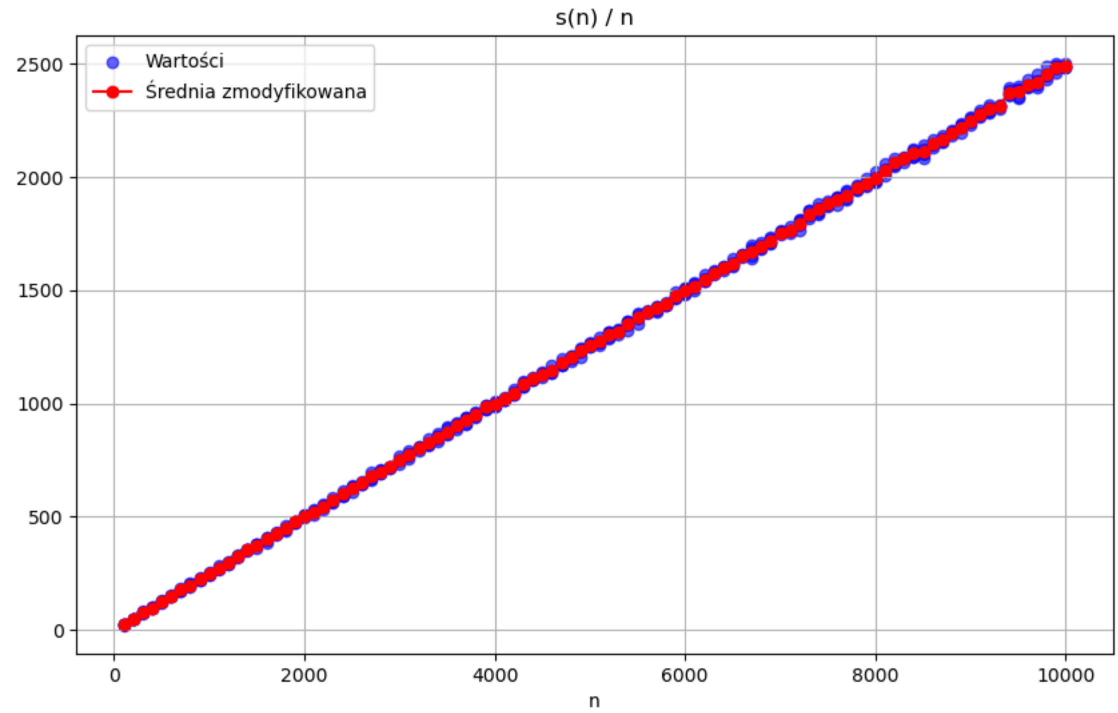
Zadanie 2

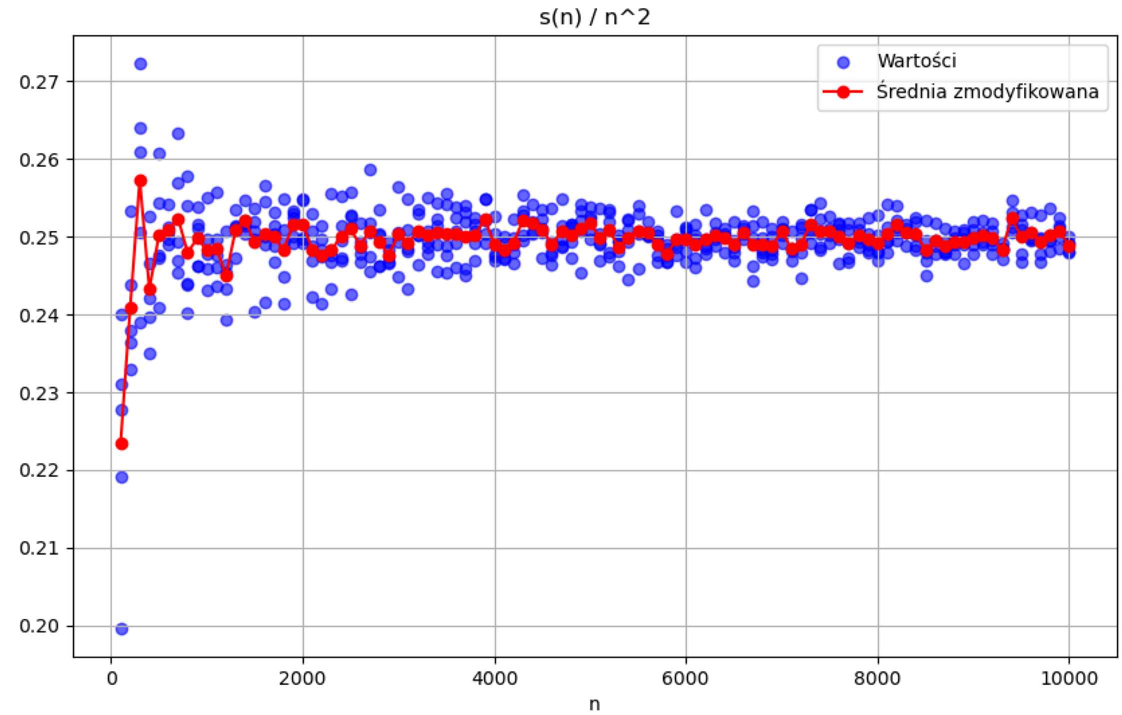

cmp(n)
1e7
2.5
2.0
1.5
1.0
0.5
0.0
Dokładne wartości
Średnia
0
2000
4000
6000
8000
10000
n


cmp(n) / n
2500
2000
1500
1000
500
0
Wartości
Średnia zmodyfikowana
0
2000
4000
6000
8000
10000
n


cmp(n) / n^2
Wartości
Średnia zmodyfikowana
0.27
0.26
0.25
0.24
0.23
0.22
0.21
0
2000
4000
6000
8000
10000
n


s(n)
1e7
Wartości
Średnia
2.5
2.0
1.5
1.0
0.5
0.0
0
2000
4000
6000
8000
10000
n

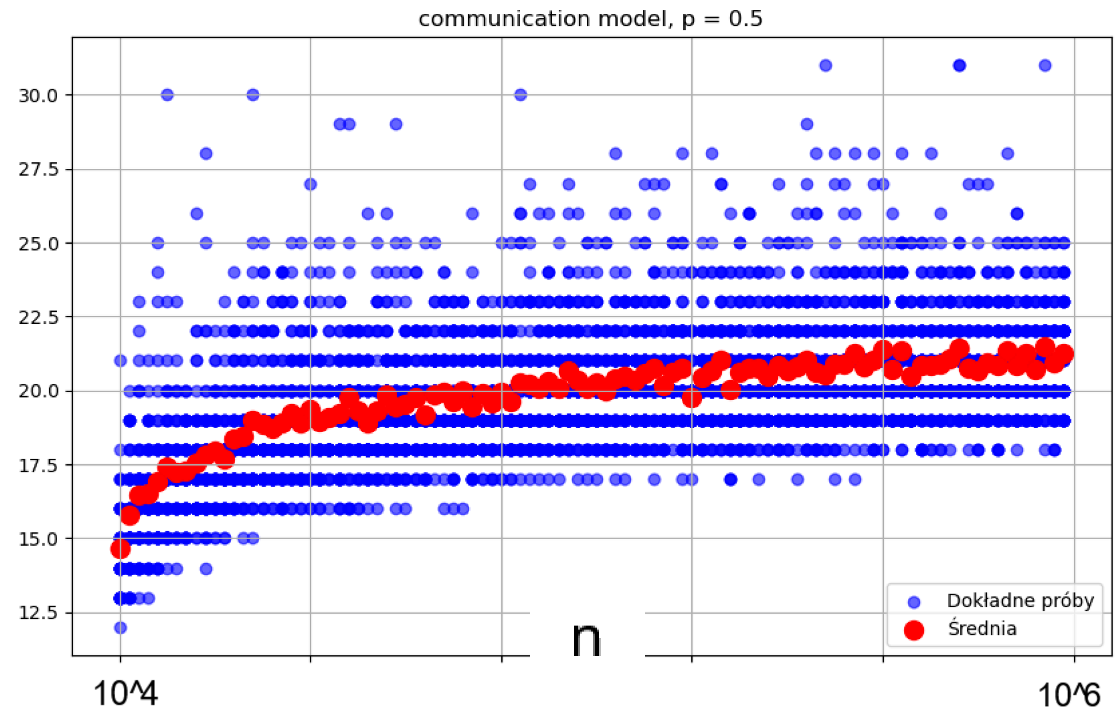

s(n) / n
Wartości
Średnia zmodyfikowana
2500
2000
1500
1000
500
0
0
2000
4000
6000
8000
10000
n

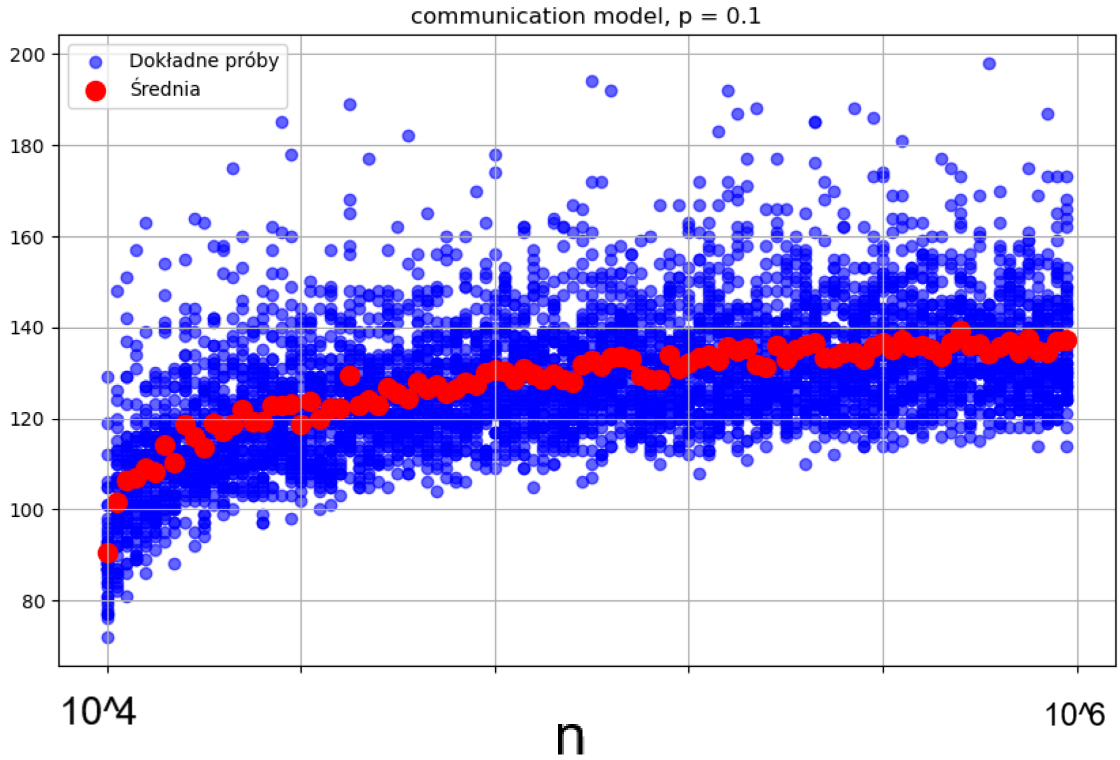

s(n) / n^2
Wartości
Średnia zmodyfikowana
0.27
0.26
0.25
0.24
0.23
0.22
0.21
0.20
0
2000
4000
6000
8000
10000
n

Cmp(n) - Czerwone kropki przedstawiają średnią liczbę porównań z 5 niezależnych symulacji dla danego n. Niebieskie kropki pokazują dokładną liczbę porównań dla pojedynczej próby. Z wykresów wynika, że wraz ze wzrostem n liczba porównań również rośnie, co jest zgodne z intuicją – dla tablicy zawierającej więcej elementów średnia liczba porównań musi być większa. Asymptotyczne tempo wzrostu cmp(n) wynosi n^2, co można zaobserwować na wykresie cmp(n) / n^2. Dodatkowo, dla większych wartości n fluktuacje i rozproszenie niebieskich kropek są wyraźniejsze niż dla mniejszych wartości n.

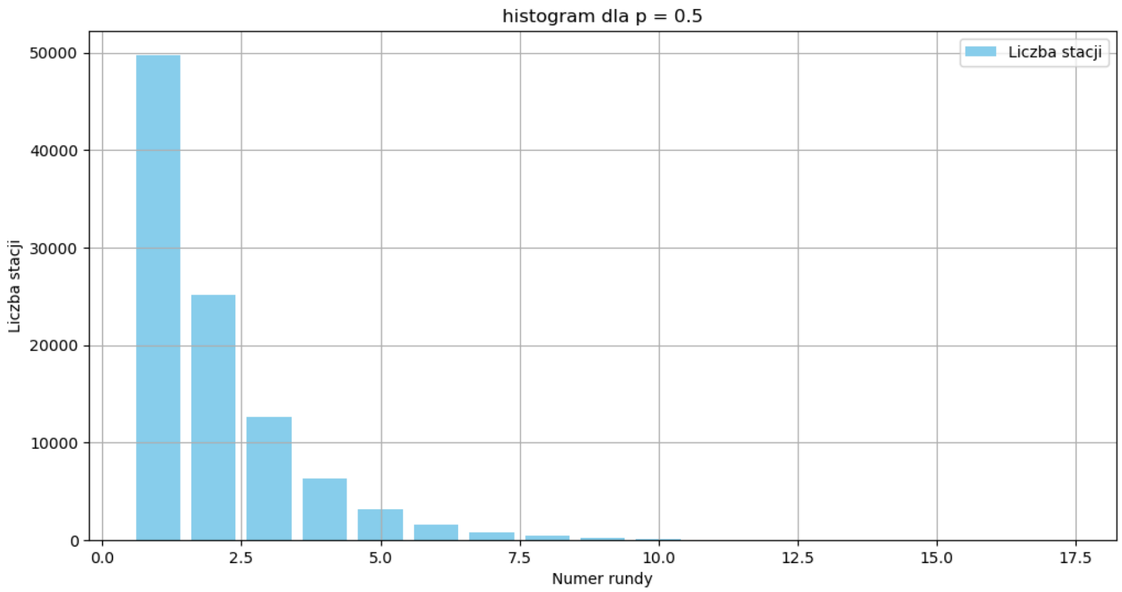
S(n) - Czerwone kropki przedstawiają średnią liczbę przestawień klucza z 5 niezależnych symulacji dla danego n. Niebieskie kropki pokazują dokładną liczbę przestawień dla pojedynczej próby. Z wykresów wynika, że wraz ze wzrostem n liczba potrzebnych przestawień klucza rośnie. Asymptotyczne tempo wzrostu s(n) wynosi w przybliżeniu n^2, co można zaobserwować na wykresie s(n) / n^2. Podobnie jak w przypadku cmp(n), dla większych wartości n fluktuacje i rozproszenie niebieskich kropek są większe niż dla mniejszych wartości n.

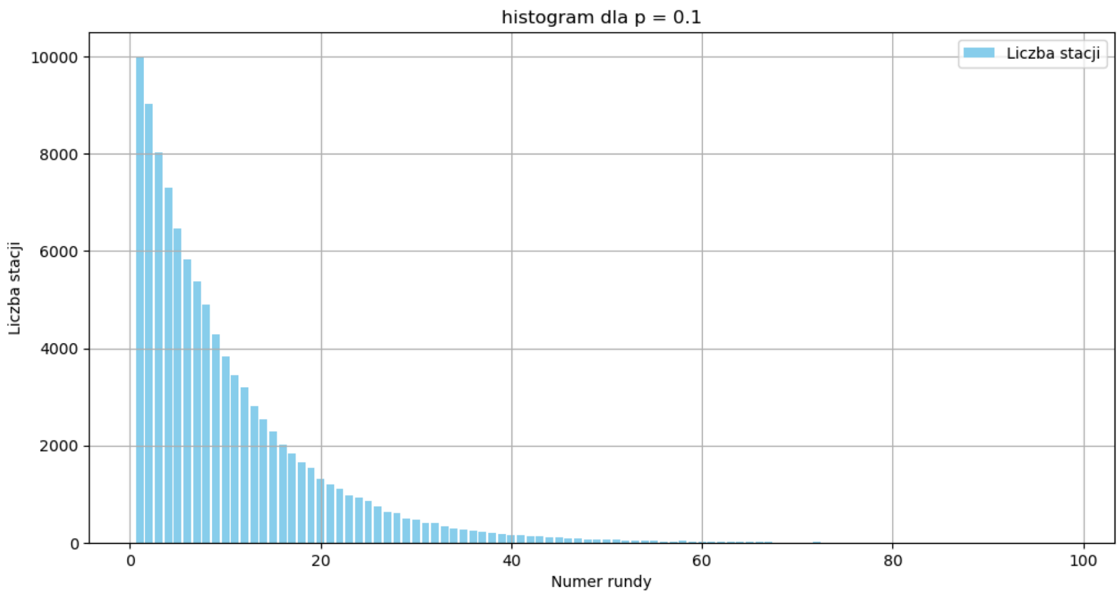
Zadanie 3





Wykresy przedstawiają wyniki symulacji uproszczonego modelu komunikacji. Dla każdego n w zbiorze (10000, 20000, …, 1000000) przeprowadzono 50 niezależnych pomiarów. Widzimy, że wraz ze wzrostem n rośnie ilość potrzebnych prób, aby każda stacja otrzymała informacje od stacji 0. Porównując oba wykresy, można zauważyć, że dla p = 0.1, średnia jest ponad 5 razy większa niż dla wykresu przedstawiającego wyniki dla p = 0.5 (dla większości n). Możemy zatem wysnuć wnioski, że średnia ilość potrzebnych prób, aby każda stacja otrzymała informacje nie maleje **liniowo** wraz ze wzrostem p - maleje szybciej.





Na powyższych histogramach przedstawiono ilość stacji, które dostały informację od stacji 0 w danej rundzie. Symulacje wykonano dla 100000 stacji. Z histogramów możemy odczytać, że rozkład zmiennych losowych $T_{n,i}$ jest rozkładem geometrycznym.

Teraz będziemy chcieli wyrazić T_n jako funkcję zmiennych losowych $T_{n,i}$;
Tn to runda, w której **najpóźniejsza** stacja otrzymała wiadomość: wynika to z faktu, że musimy czekać, aż każda ze stacji odbierze wiadomość - a najpóźniejsza z ich wyznacza czas zakończenia procesu. Zetem T_n możemy wyrazić jako:

$$\mathbf{T_n = \max(T_{n,1}, T_{n,2}, \ldots, T_{n,n})}$$