Wojciech Typer

279730

Kody źródłowe, wykresy i histogramy zostały napisane w języku Python. Do obliczeń wykorzystano biblioteki NumPy i SciPy.

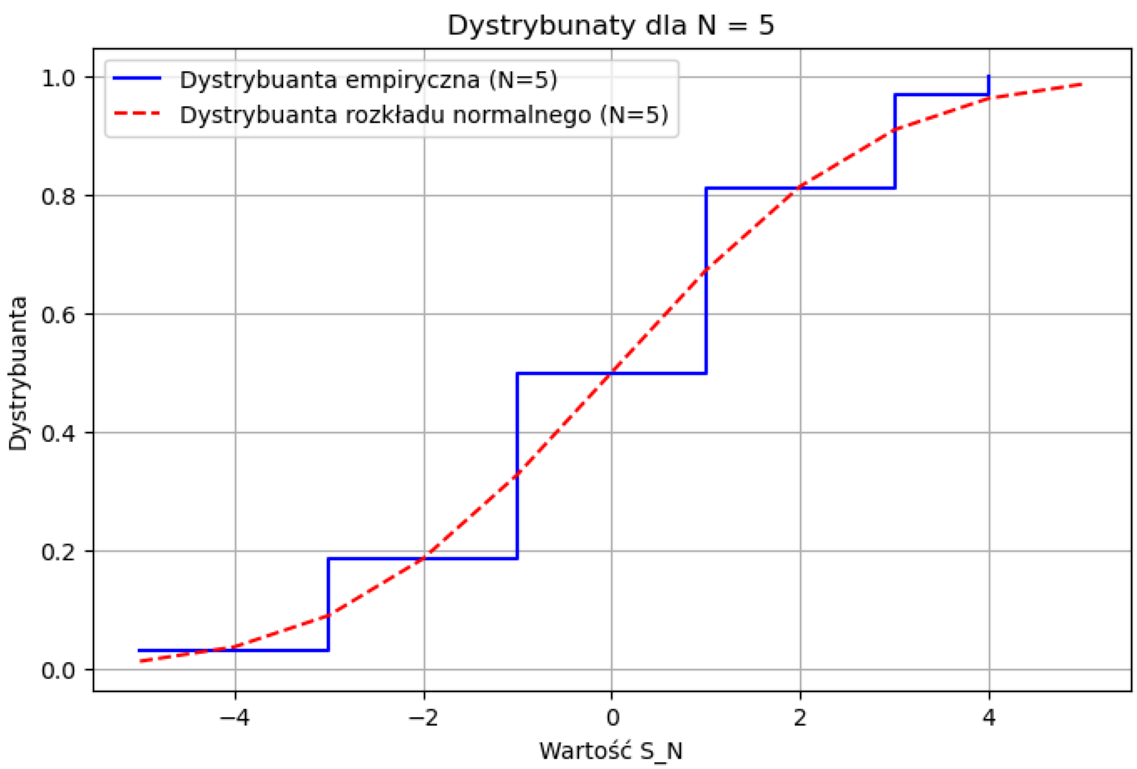
zadanie 1

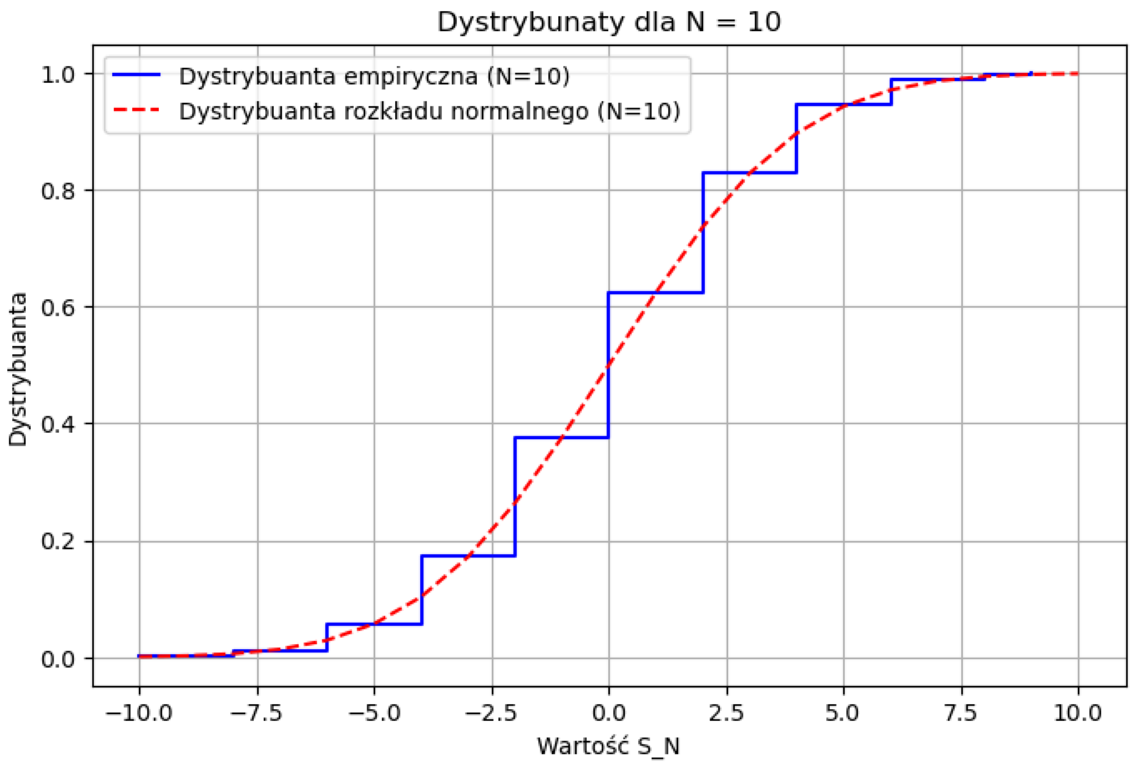
n	Markov Bound	P(X >= 6/5E(X)	Chebyshev Bound	P(|X - E(X)| >= 1/10E(X)
100	0.8333...	2.84E-02	1	3.68E-01
1000	0.8333...	1.36E-10	0.1	1.73E-03
10000	0.8333...	0	0.01	1.55E-23 ~ 0.0

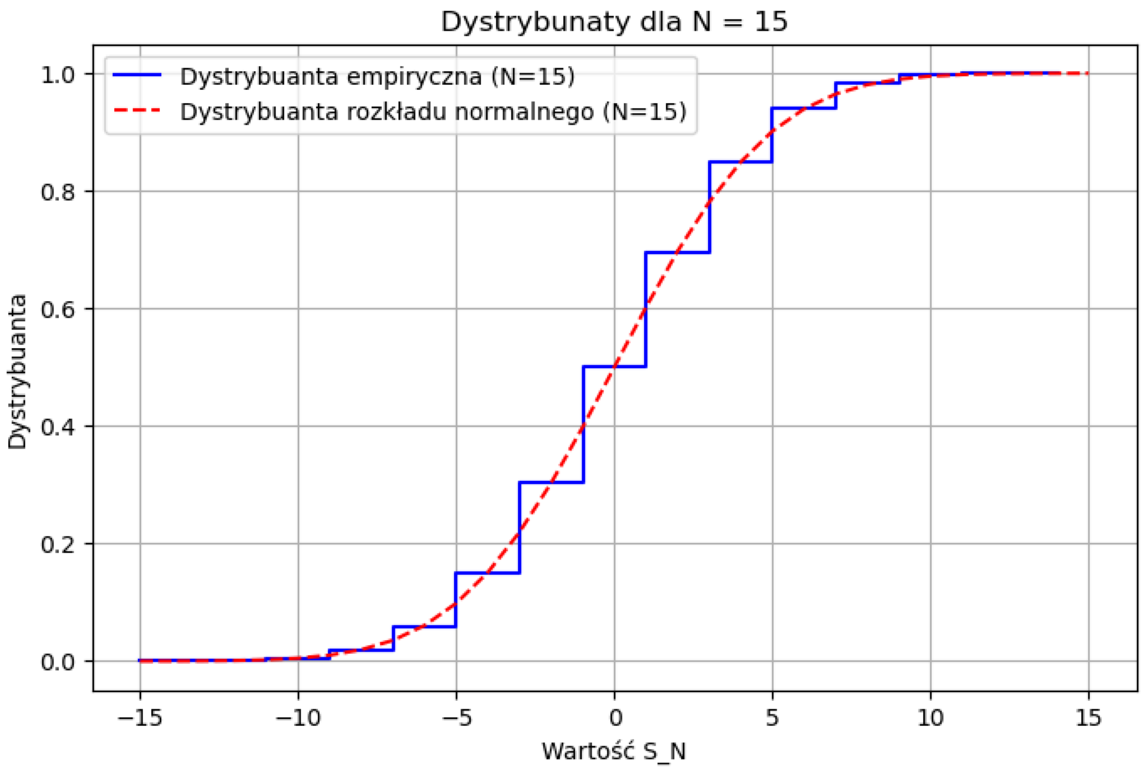
Powyższa tabela przedstawia wynik pierwszego zadania. Markov Bound to prawa strona nierówności Markowa - E(X) / a - czyli w kontekście zadania E(X) / (6/5 * E(X)) = ⅚ ~ 0.8333. Chebyshev Bound to analogicznie, prawa strona nierówności Czebyszewa - var(x) / a^2, czyli w kontekście naszego zadania, var(x) / (0.1E(x)) ^2 = np(1-p) / (0.01(np)^2) = 1-p / 0.01np = 100/n.

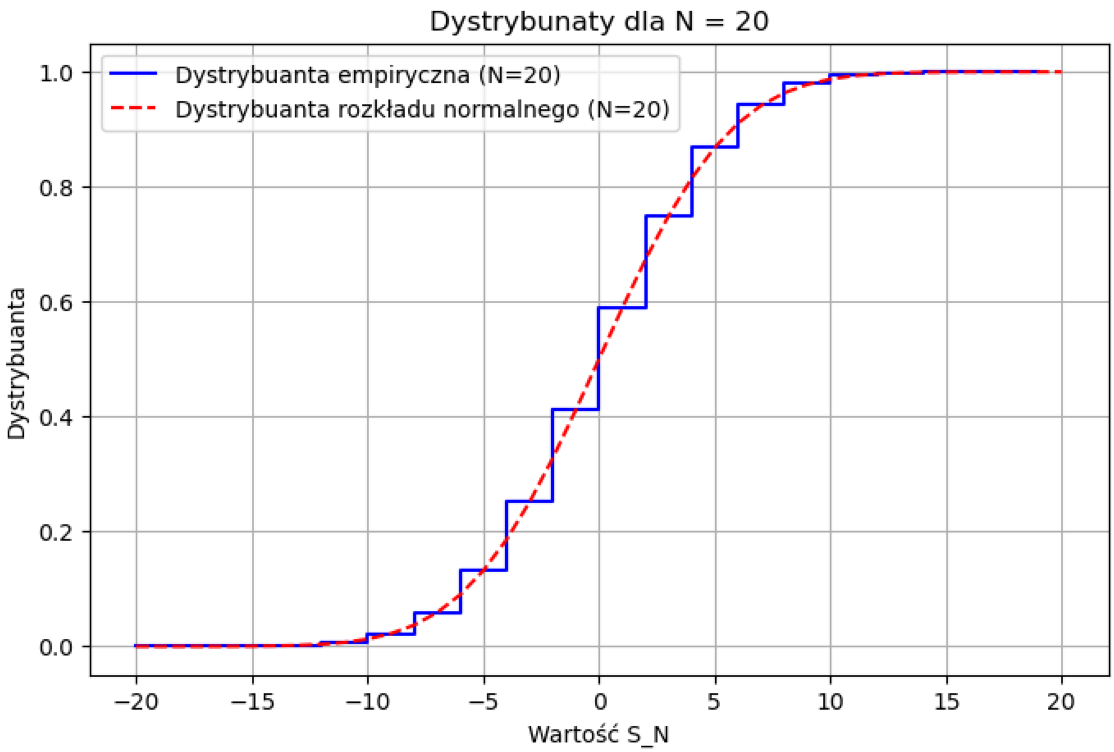
Z tabeli wyników możemy odczytać, że nierówności Markowa i Czebyszewa zostały zachowane, gdyż prawe ich strony (Markov Bound i Chebyshev Bound) są większe niż rzeczywiste prawdopodobieństwo wystąpienia danych zdarzeń. Widzimy również, że dokładniejsze oszacowanie daje nierówność Czebyszewa.

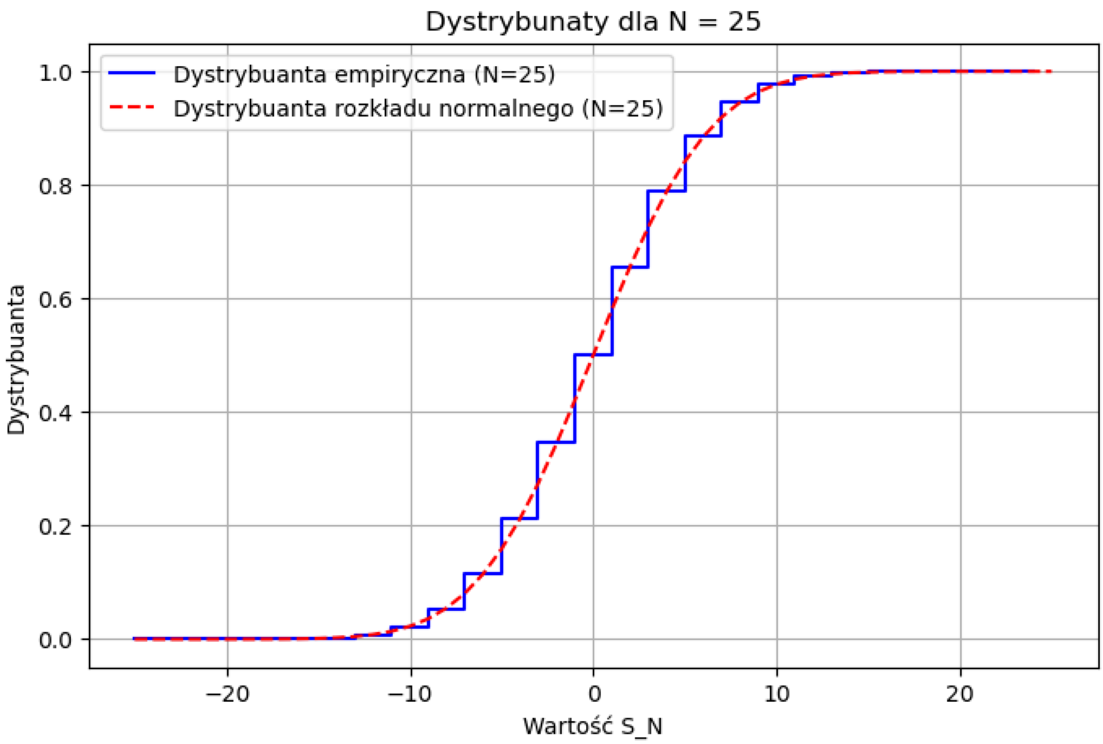
zadanie 2

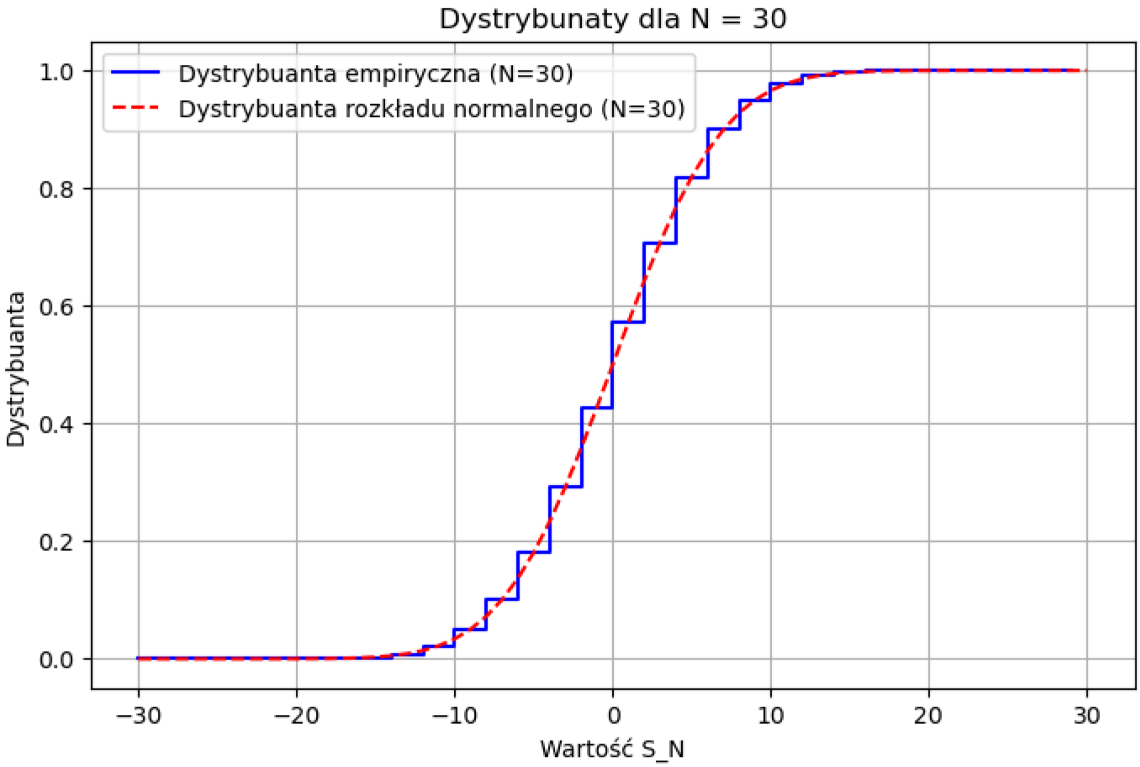



Dystrybunaty dla N = 10
Dystrybuanta empiryczna (N=10)
Dystrybuanta rozkładu normalnego (N=10)
Dystrybuanta
1.0
0.8
0.6
0.4
0.2
0.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Wartość S_N

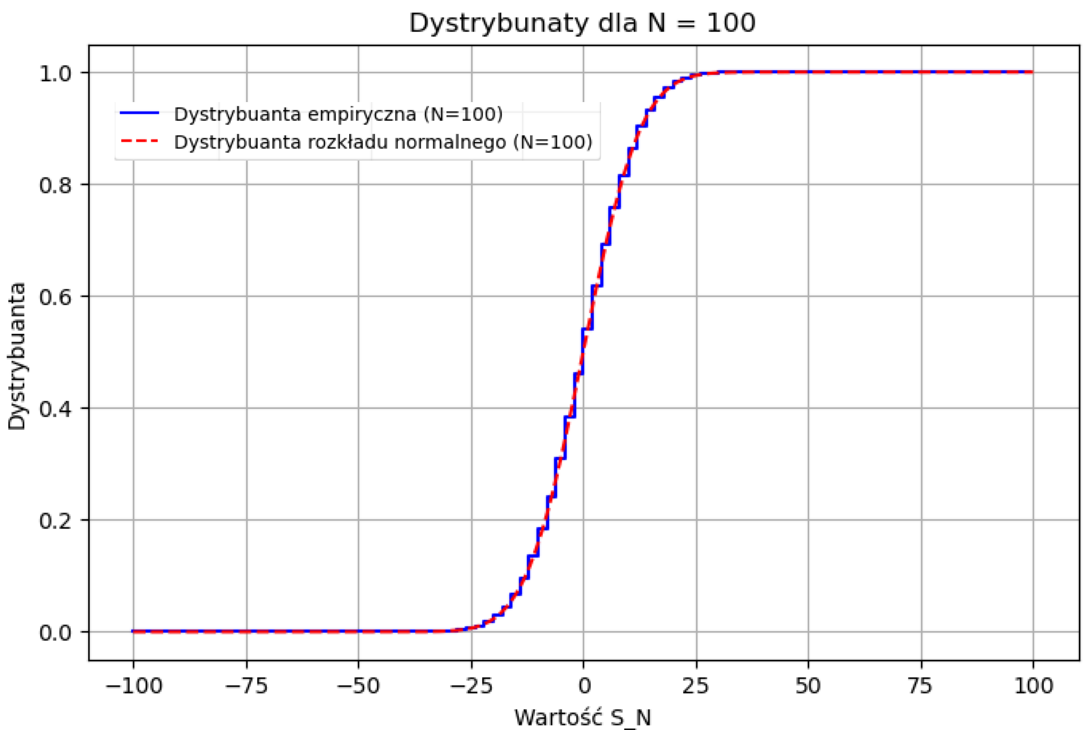

Dystrybunaty dla N = 15
Dystrybuanta empiryczna (N=15)
Dystrybuanta rozkładu normalnego (N=15)
Dystrybuanta
1.0
0.8
0.6
0.4
0.2
0.0
−15
−10
−5
0
5
10
15
Wartość S_N


Dystrybunaty dla N = 20
Dystrybuanta empiryczna (N=20)
Dystrybuanta rozkładu normalnego (N=20)
Dystrybuanta
Wartość S_N
1.0
0.8
0.6
0.4
0.2
0.0
−20
−15
−10
−5
0
5
10
15
20


Dystrybunaty dla N = 25
Dystrybuanta empiryczna (N=25)
Dystrybuanta rozkładu normalnego (N=25)
Dystrybuanta
Wartość S_N
1.0
0.8
0.6
0.4
0.2
0.0
−20
−10
0
10
20



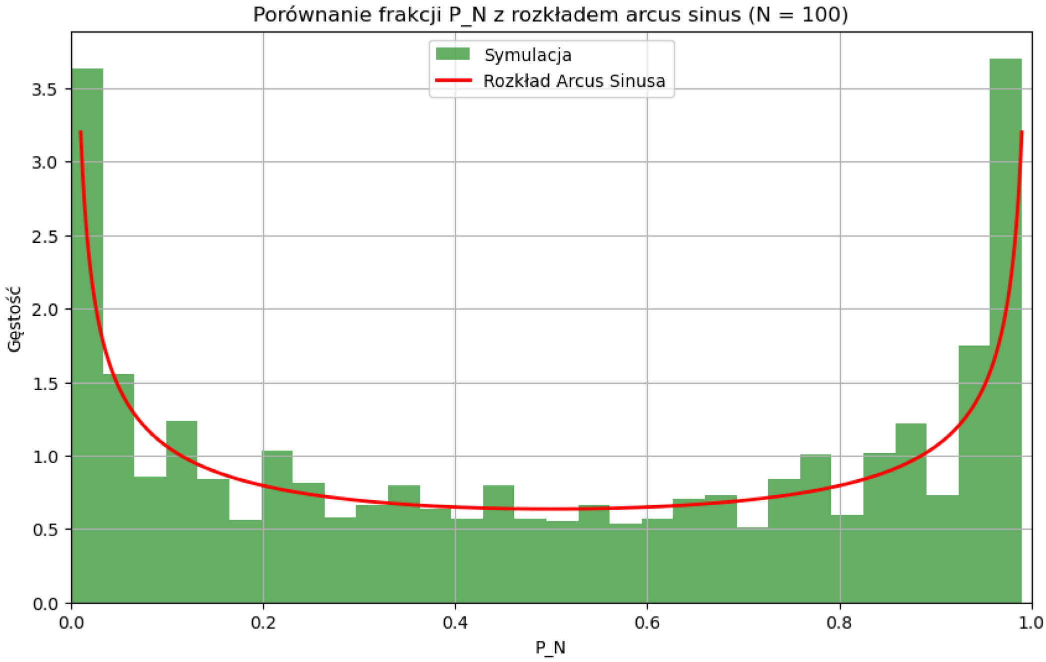
Dystrybunaty dla N = 30
Dystrybuanta empiryczna (N=30)
Dystrybuanta rozkładu normalnego (N=30)
Dystrybuanta
1.0
0.8
0.6
0.4
0.2
0.0
−30
−20
−10
0
10
20
30
Wartość S_N

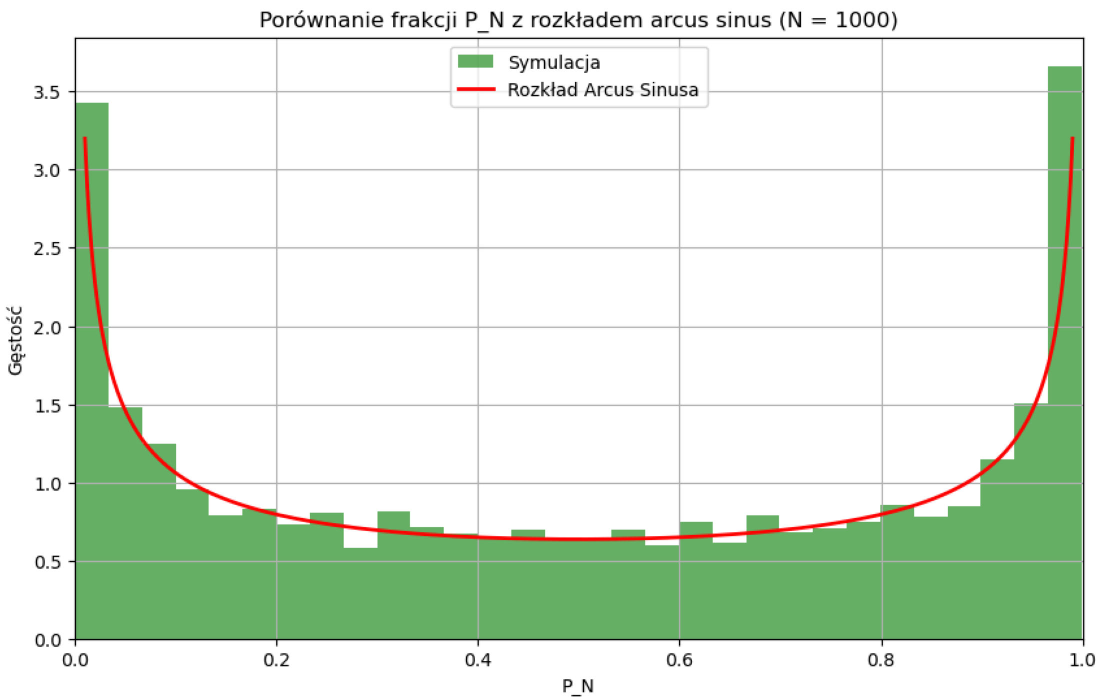


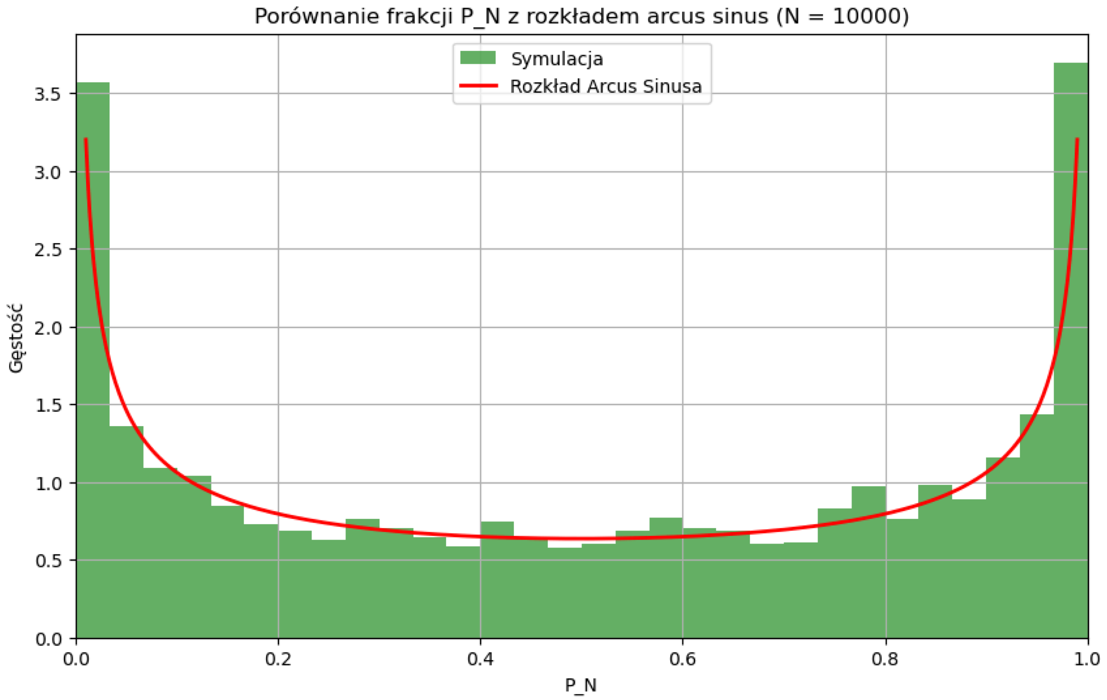
Dystrybunaty dla N = 100
Dystrybuanta empiryczna (N=100)
Dystrybuanta rozkładu normalnego (N=100)
Dystrybuanta
1.0
0.8
0.6
0.4
0.2
0.0
−100
−75
−50
−25
0
25
50
75
100
Wartość S_N

Z wykresów możemy odczytać, że dla coraz większych N, dystrybuanta empiryczna coraz bardziej przypomina dystrybuantę rozkładu normalnego. Jest to zgodne z Centralnym Twierdzeniem Granicznym, które mówi, że suma niezależnych zmiennych losowych o skończonych wariancjach zbiega w rozkładzie do rozkładu normalnego. W naszym przypadku Sn (suma niezależnych zmiennych losowych Xi o wartościach +/- 1) coraz lepiej aproksymuje rozkład normalny, co szczególnie widać dla dużych wartości N.

zadanie 3







Z histogramów możemy odczytać, że proces błądzenia losowego ma tendencję do znajdowania się większość czasu po jednej ze stron od osi OX - pod nią lub nad nią. Świadczy o tym fakt, że im bardziej skrajna wartość, tym słupek jest wyższy - niskie słupki w środkowej części histogramu wskazują, że proces rzadko kiedy spędzał równą część czasu po dwóch stronach osi OX. Zauważyć możemy również, że Pn = 0.5 jest swego rodzaju osią symetrii histogramu, co jest zgodne z wcześniejszymi wnioskami. Dodatkowo, im większa liczba N, tym bardziej zauważalna jest koncentracja wartości Pn przy krańcach histogramu. Jest to zgodne z obserwacjami z wcześniejszych list zadań i przewiduje większą stabilność wyników dla coraz to większych liczb przeprowadzenia symulacji.

Dla dużych N wykres coraz bardziej przybliża wykres gęstości rozkładu arcusa sinusa (o parametrach a = b = 0.5). Gęstość rozkładu arcusa sinusa wyrażona jest wzorem:

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

i odpowiada wartościom przedstawionym na wykresie: lim(x → 0) = lim (x → 1) = ∞