Metody Probabilistyczne i Statystyka

Sprawozdanie Homework 4

Anna Grelewska

Zadanie 1

	n	Markov $P(X \geq 1.2E(X))$	Chebyshev $P(X \geq 1.2E(X))$	Exact $P(X \geq 1.2E(X))$	Markov $P(\lvert X - E(X)\rvert \geq 0.1E(X))$	Chebyshev $P(\lvert X - E(X)\rvert \geq 0.1E(X))$	Exact $P(\lvert X - E(X)\rvert \geq 0.1E(X))$
0	100	0.8333333333	0.2500000000	0.0284439668	0.9090909091	1.0000000000	0.3682016173
1	1000	0.8333333333	0.0250000000	0.0000000001	0.9090909091	0.1000000000	0.0017305361
2	10000	0.8333333333	0.0025000000	0.0000000000	0.9090909091	0.0100000000	0.0000000000

Wnioski:

Tabela przedstawia wyniki dla różnych wartości nn (100, 1000, 10000) i pozwala porównać oszacowania nierówności Markowa i Czebyszewa z dokładnymi wartościami prawdopodobieństwa.

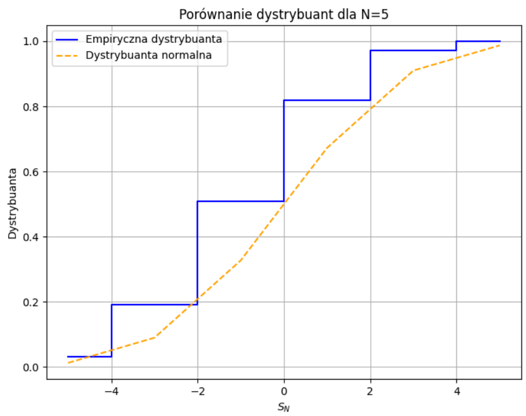
Nierówność Markowa
- Bardzo słabe oszacowania, znacznie zawyża prawdopodobieństwo.
- Stałe wartości (0.833333 i 0.909091) niezależnie od n.
- Najgorsza metoda spośród analizowanych.

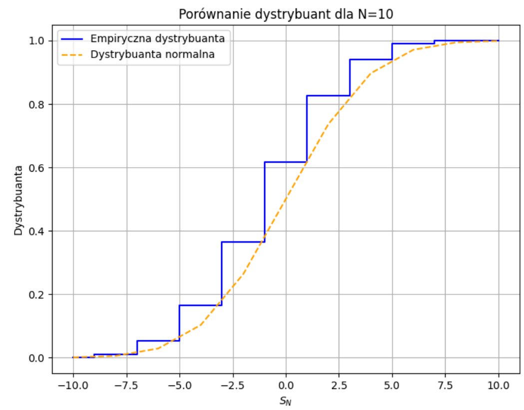
Nierówność Czebyszewa
- Lepsza od Markowa, ale nadal przeszacowuje rzeczywiste wartości.
- Maleje wraz ze wzrostem n, co czyni ją bardziej dokładną dla dużych próbek.
- Dla dużych n jest bliżej rzeczywistych wartości.

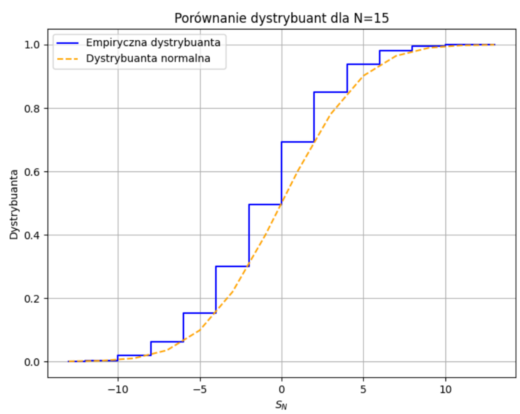
Dokładne wartości prawdopodobieństwa
- Są znacznie mniejsze niż oszacowania Markowa i Czebyszewa.
- Dla dużych n szybko dążą do 0, co pokazuje, jak konserwatywne są te nierówności.

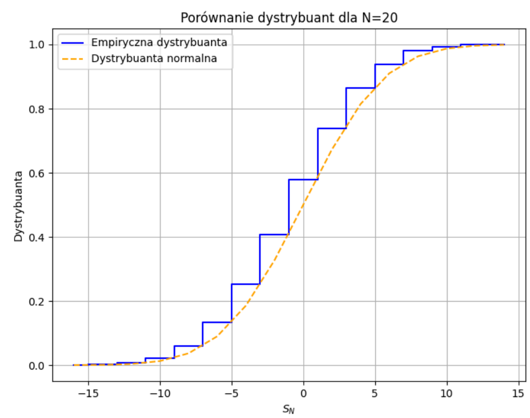
- Czebyszew jest dokładniejszy niż Markow, ale nadal znacznie zawyża wartości.
- Dokładne wartości są znacznie niższe, co pokazuje, że nierówności te dają bardzo luźne ograniczenia.

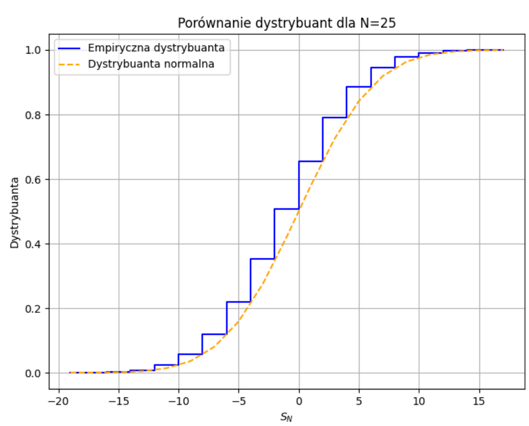
Zadanie 2

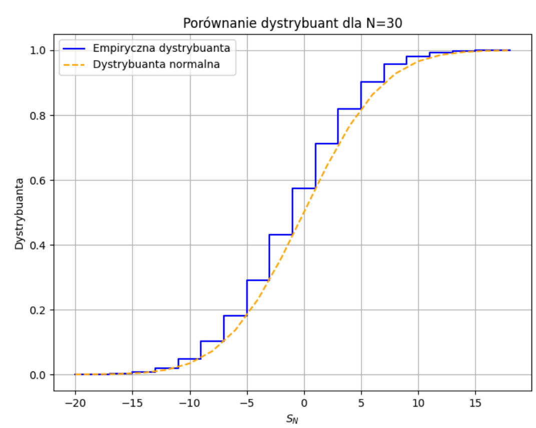

Porównanie dystrybuant dla N=5
Empiryczna dystrybuanta
Dystrybuanta normalna
Dystrybuanta
S_N

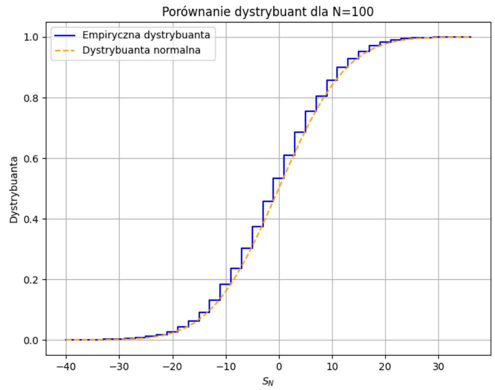

Porównanie dystrybuant dla N=10
Empiryczna dystrybuanta
Dystrybuanta normalna
Dystrybuanta
S_N


Porównanie dystrybuant dla N=15
Empiryczna dystrybuanta
Dystrybuanta normalna
Dystrybuanta
S_N


Porównanie dystrybuant dla N=20
Empiryczna dystrybuanta
Dystrybuanta normalna
Dystrybuanta
S_N


Porównanie dystrybuant dla N=25
Empiryczna dystrybuanta
Dystrybuanta normalna
Dystrybuanta
S_N


Porównanie dystrybuant dla N=30
Empiryczna dystrybuanta
Dystrybuanta normalna
Dystrybuanta
S_N



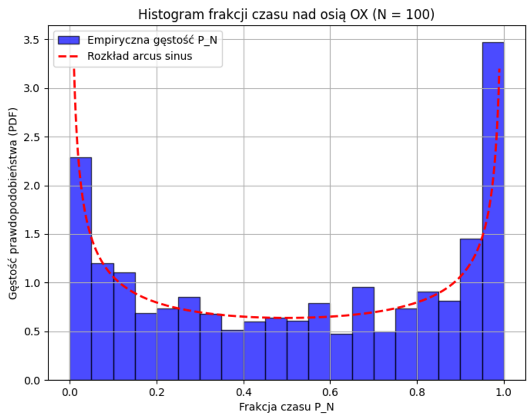
Wnioski:

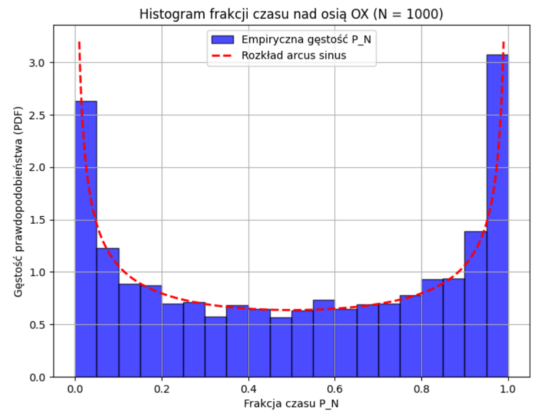
Dla małych N (np. 5, 10) empiryczna dystrybuanta jest silnie skokowa, co wynika z dyskretnego charakteru procesu. Wartości zmiennej losowej S_N przyjmują tylko określone liczby całkowite, co powoduje widoczne "stopnie" w dystrybuancie. Rozkład normalny nie pasuje jeszcze dobrze, ponieważ rozkład S_N jest daleki od ciągłości. Dla większych N (np. 15, 20, 25, 30) empiryczna dystrybuanta staje się bardziej płynna, a skoki są coraz mniejsze. To efekt zwiększającej się liczby możliwych wartości S_N, co zmniejsza różnice między kolejnymi poziomami. Aproksymacja normalna zaczyna lepiej pasować do danych. Dla N = 100 empiryczna dystrybuanta niemal całkowicie pokrywa się z dystrybuantą normalną. Skoki są praktycznie niewidoczne, a dystrybuanta przyjmuje niemal ciągły charakter.

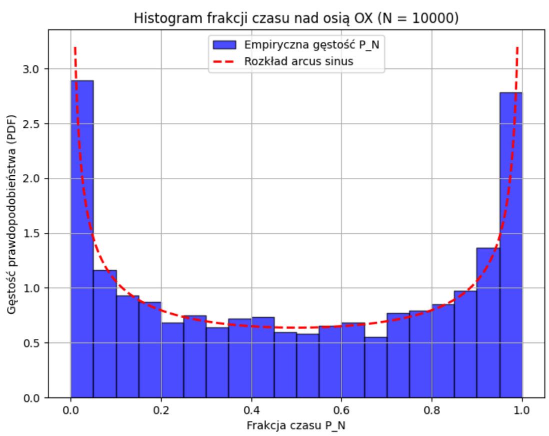
Błądzenie losowe dla dużego N jest bardzo dobrze aproksymowane przez rozkład normalny, co jest zgodne z Centralnym Twierdzeniem Granicznym. Dla małych wartości N, proces pozostaje wyraźnie dyskretny, przez co rozkład normalny nie jest jeszcze dokładnym modelem.
Dla N = 100 suma kroków losowych zachowuje się niemal jak zmienna o rozkładzie normalnym N(0, N), co oznacza, że wartości empiryczne są praktycznie zgodne z wartościami teoretycznymi wynikającymi z rozkładu normalnego.

Wyniki pokazują, że wraz ze wzrostem N dystrybuanta empiryczna coraz lepiej dopasowuje się do rozkładu normalnego. Dla dużych wartości N błądzenie losowe można skutecznie przybliżyć rozkładem Gaussa, co potwierdza teoretyczne przewidywania.

Zadanie 3







Wnioski:

Dla N = 100 histogram jest nieregularny, ale widoczna jest tendencja do większej liczby wartości blisko 0 i 1.
Dla N = 1000 kształt histogramu coraz lepiej pasuje do rozkładu arcus sinus.
Dla N = 10000 histogram niemal idealnie pokrywa się z funkcją gęstości arcus sinus.

Empiryczna gęstość P_N dobrze aproksymuje rozkład arcus sinus, zwłaszcza dla dużych N.
Najwięcej wartości P_N znajduje się blisko 0 i 1, co oznacza, że błądzenie często przebywa nad lub pod osią OX przez długi czas.
Średnia E(P_N) ≈ 0.5, a wariancja maleje wraz ze wzrostem N.
Dla dużych wartości N rozkład P_N jest zgodny z przewidywaniami teoretycznymi.