Zadanie domowe 4

Nel Skowronek

January 30, 2025

Zadanie 1: Nierówności ogonowe dla rozkładu dwumianowego Bin $\left(n, \frac{1}{2}\right)$

Przybliżymy wartości następujących prawdopodobieństw za pomocą nierówności Markowa i Czebyszewa:

- $P\left(X \geq \frac{6}{5} \cdot \mathbb{E}(X)\right)$
- $P\left(|X - \mathbb{E}(X)| \geq \frac{1}{10} \cdot \mathbb{E}(X)\right)$

Aby móc zastosować obie nierówności do obu prawdopodobieństw, trzeba skożystać z faktu, że rozkład Bin $\left(n, \frac{1}{2}\right)$ jest symetryczny:

- $P\left(X \geq \frac{6}{5} \cdot \mathbb{E}(X)\right) = P\left(X - \mathbb{E}(X) \geq \frac{1}{5} \cdot \mathbb{E}(X)\right) = \frac{1}{2} P\left(|X - \mathbb{E}(X)| \geq \frac{1}{5} \cdot \mathbb{E}(X)\right)$
- $P\left(|X - \mathbb{E}(X)| \geq \frac{1}{10} \cdot \mathbb{E}(X)\right) = 2P\left(X - \mathbb{E}(X) \geq \frac{1}{10} \cdot \mathbb{E}(X)\right) = 2P\left(X \geq \frac{11}{10} \cdot \mathbb{E}(X)\right)$

Wyniki przybliżeń oraz dokładne wartości prawdopodobieństw:

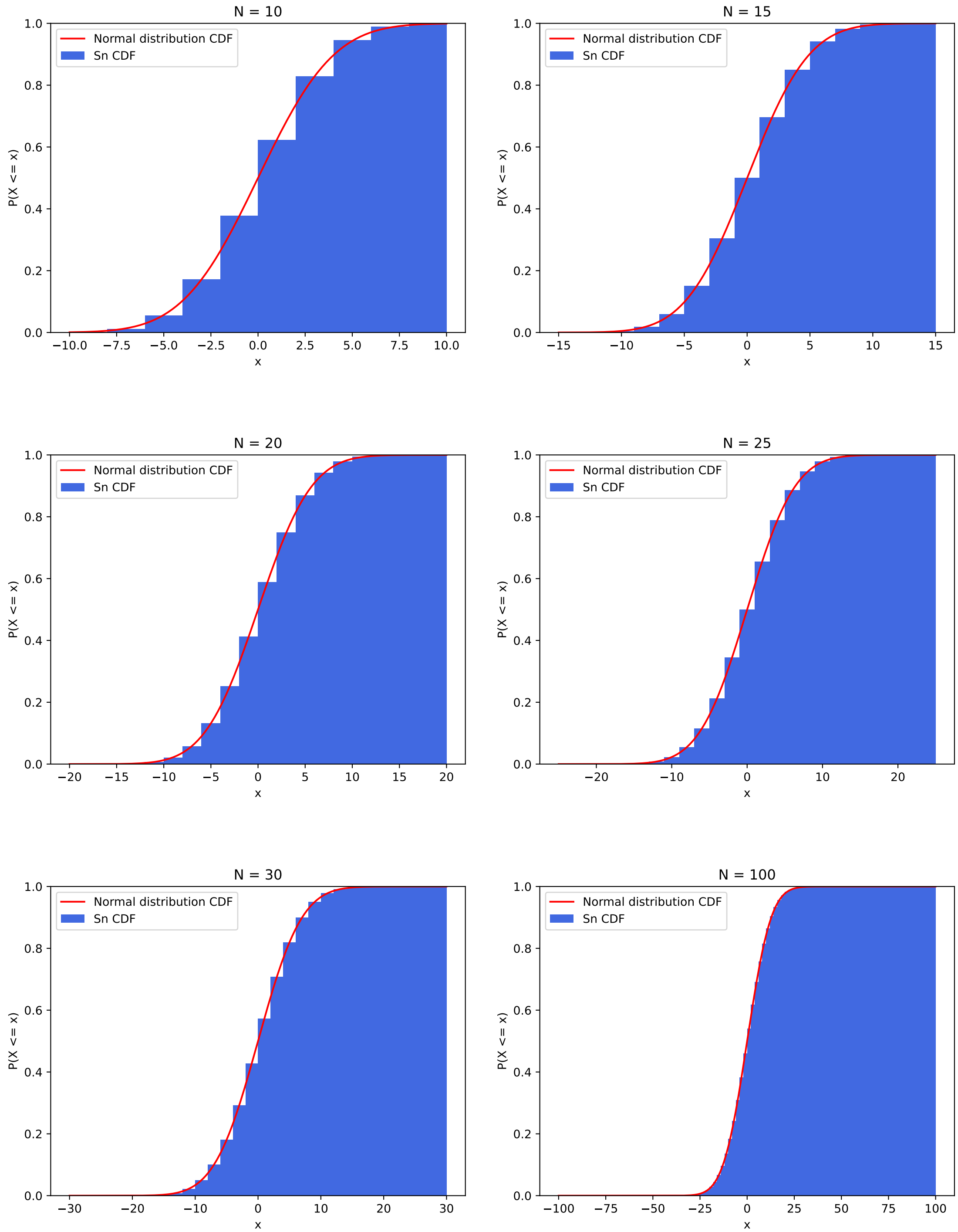
	n	P(X >= 6/5E(X))	P(|X - E(X)| >= 1/10E(X))
Markow	100	0.8333	1.818
Czebyszew	100	0.125	1.0
Exact	100	0.02844	0.3682
Markow	1000	0.8333	1.818
Czebyszew	1000	0.0125	0.1
Exact	1000	1.364e-10	0.001731
Markow	10000	0.8333	1.818
Czebyszew	10000	0.00125	0.01
Exact	10000	0.0	0.0

Jak widać, obie nierówności zostawiają wiele do życzenia, jednak nierówność Czebyszewa jest zdecydowanie dokładniejsza. Daje bliższe wyniki, oraz zależy od n, co nie ma miejsca w nierówności Markowa w tych przypadkach.

Zadanie 2: Błądzenie losowe na liczbach całkowitych

W celu ułatwienia zadania należy zauważyć, że zmienna S_N jest transformacją liniową rozkładu Bin $\left(N, \frac{1}{2}\right)$.
Dokładniej:
$S_N = 2 \cdot \text{Bin}\left(n, \frac{1}{2}\right) - N$
A więc:
$F_{S_N}(t) = P(S_N \leq t) = P\left(2 \cdot \text{Bin}\left(n, \frac{1}{2}\right) - N \leq t\right) = P\left(\text{Bin}\left(n, \frac{1}{2}\right) \leq \frac{t+N}{2}\right) = F_{\text{Bin}\left(n, \frac{1}{2}\right)}\left(\frac{t+N}{2}\right)$
Wystarczy więc wyznaczyć dystrybuantę Bin $\left(N, \frac{1}{2}\right)$ i odpowiednio przeskalować pod nią oś OX.

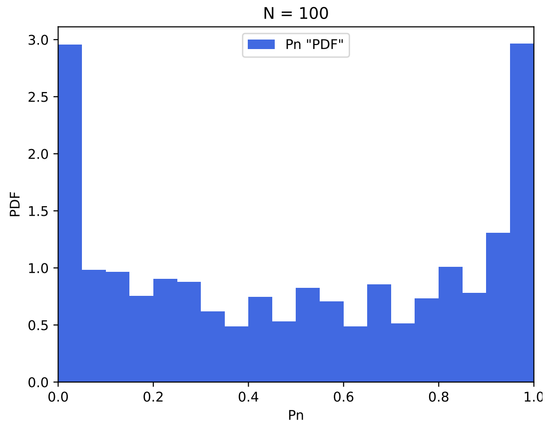
Wykresy dystrybuant S_N oraz przybliżających ich rozkładów normalnych $\mathcal{N}\left(E(S_N), \sqrt{\text{var}(S_N)}\right)$:

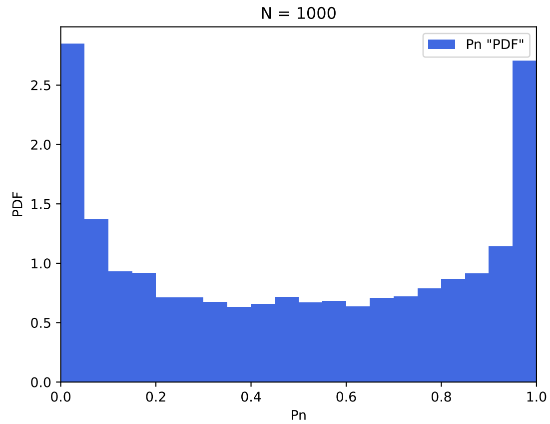


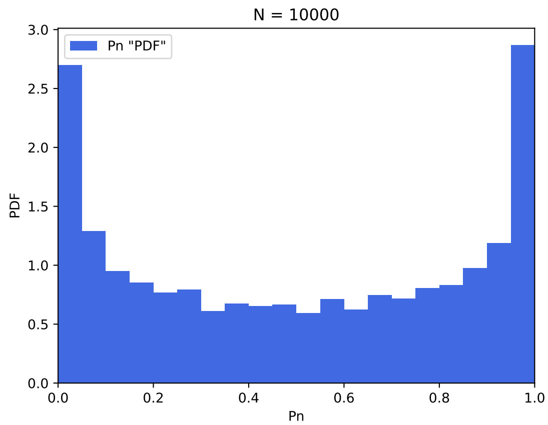
Wyraźnie widać że razem z rosnącym N, S_N zbiega według rozkładu do $\mathcal{N}\left(E\left(S_N\right), \operatorname{var}\left(S_N\right)\right)$ co jest zgodne z **CLT**.

Zadanie 3: Błądzenie losowe na $\mathbb{Z}$ - rozkład *„czasu spędzonego nad osią OX”*

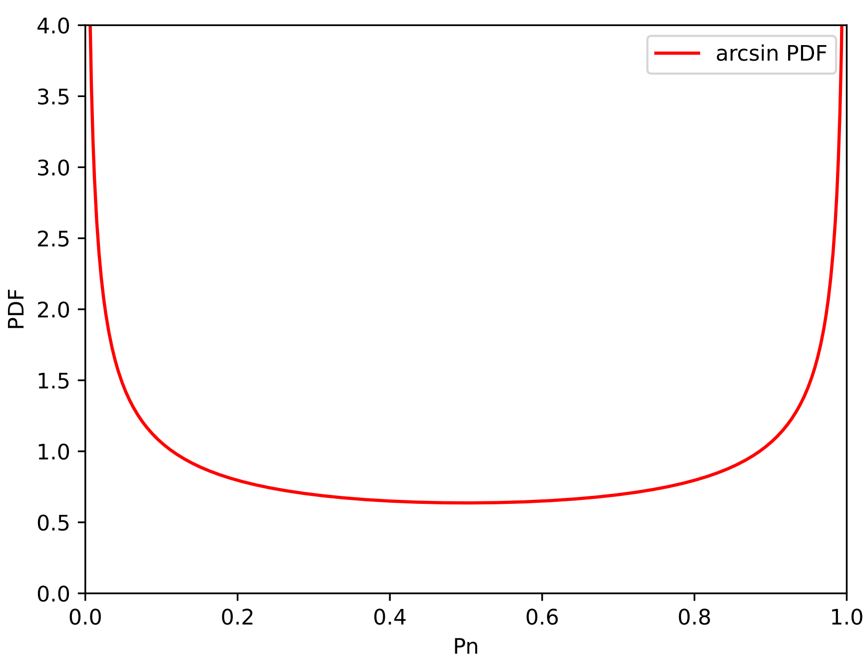
Wyniki przeprowadzonych symulacji:







Dla porównania - funkcja gęstości rozkładu z zadania 5 listy 8 $f_X(x) = \frac{1}{\pi\sqrt{x-x^2}}$:



Widać mocne podobieństwo pomiędzy wykresami, co sugerowałoby dążenie rozkładu P_N do rozkładu arcusa sinusa z zadania 5. Jedną z rzeczy, które najmocniej rzucają się w oczy jest jednak tendencja do przyjjmowania przez P_N granicznych wartości (0 lub 1). Interpretacja tego jest taka, że prawie cały spacer spędziliśmy albo nad osią OX albo pod. Jest to zgodne z intuicją, z powodu *braku pamięci* procesu S_N - po oddaleniu się o kilka kroków od osi OX proces *nie pamięta* jakie było jego położenie początkowe. Zachowuje się jakby zaczynał zupełnie od nowa na tej wysokości. Nic go nie ciągnie spowrotem do osi OX.

Kod źródłowy implementacji: Github repository