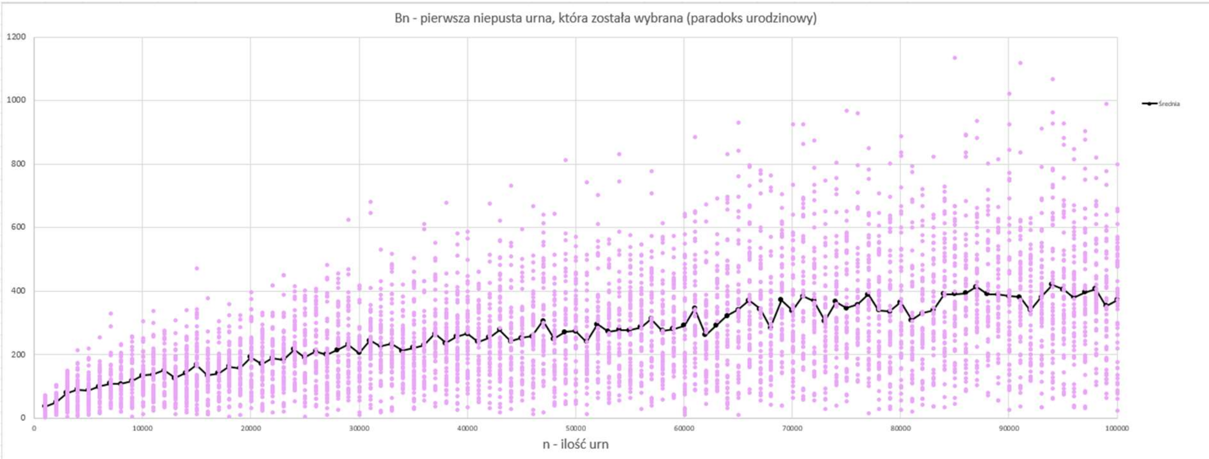
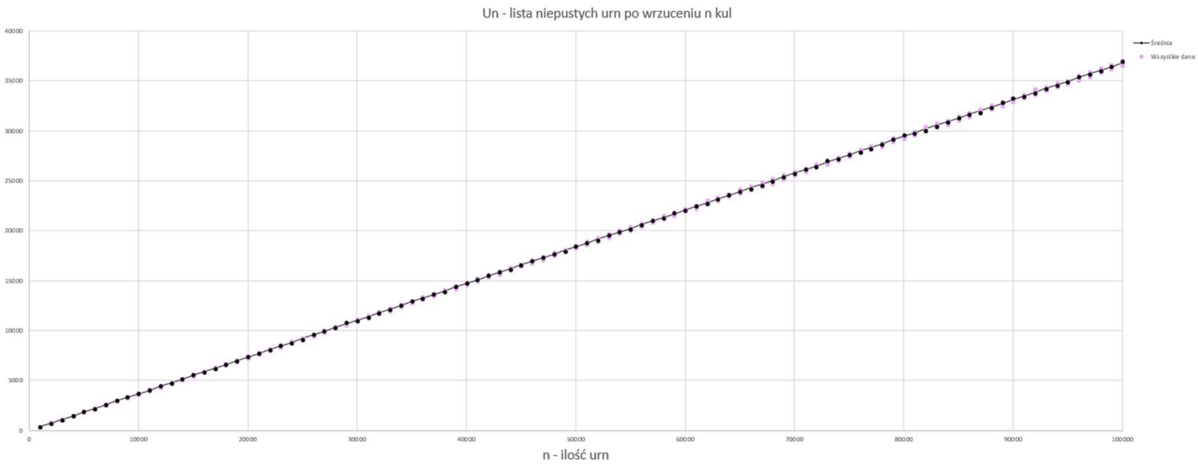
JOANNA KULIG
nr indeksu 261738

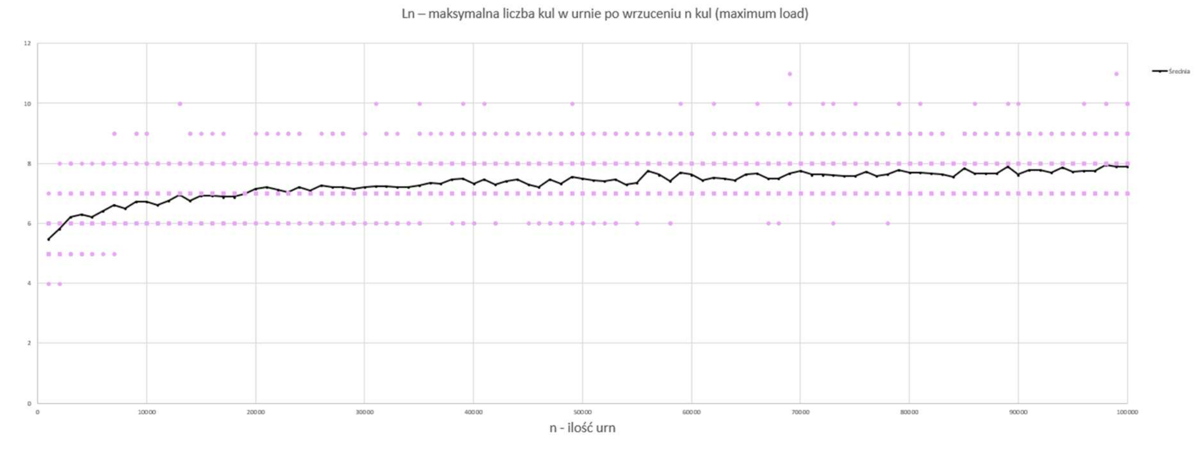
METODY PROBABILISTYCZNE I STATYSTYKA

LISTA 7, ZADANIE 1

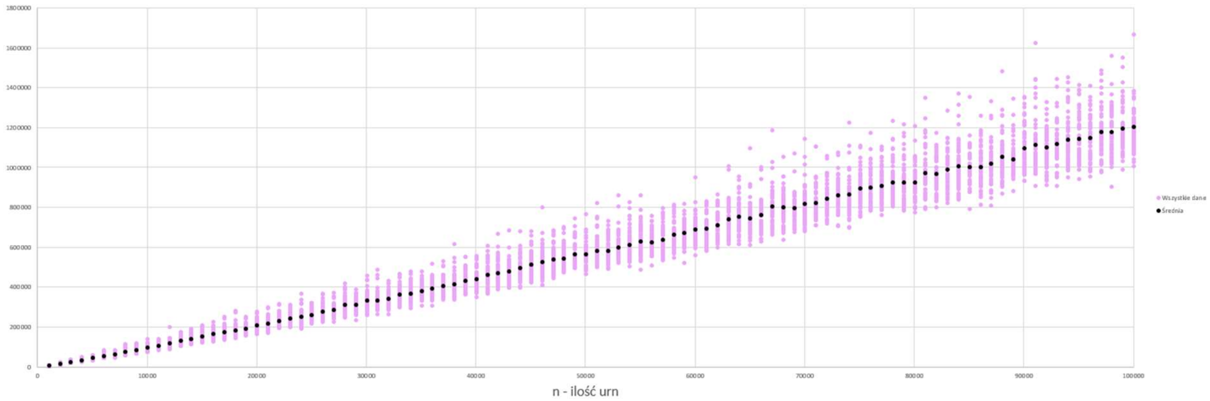
Wykresy statystyk dla $Bn, Un, Ln, Cn, Dn - Cn$ wraz ze średnimi wartościami:



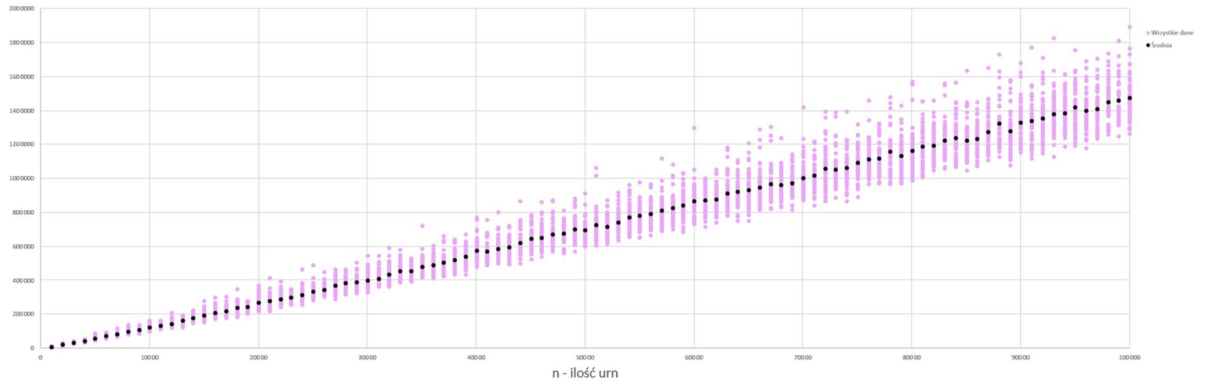




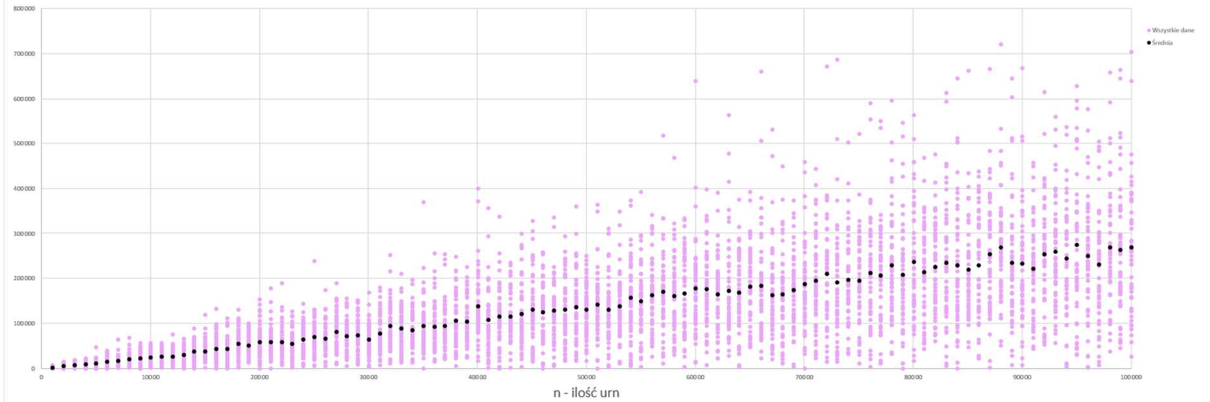
Cn – minimalna liczba rzutów, po której w kazdej z urn jest co najmniej jedna kula
(problem kolekcjonera kuponów)



Dn – minimalna liczba rzutów, po której w kazdej z urn są co najmniej dwie kule (coupon collector's brother)

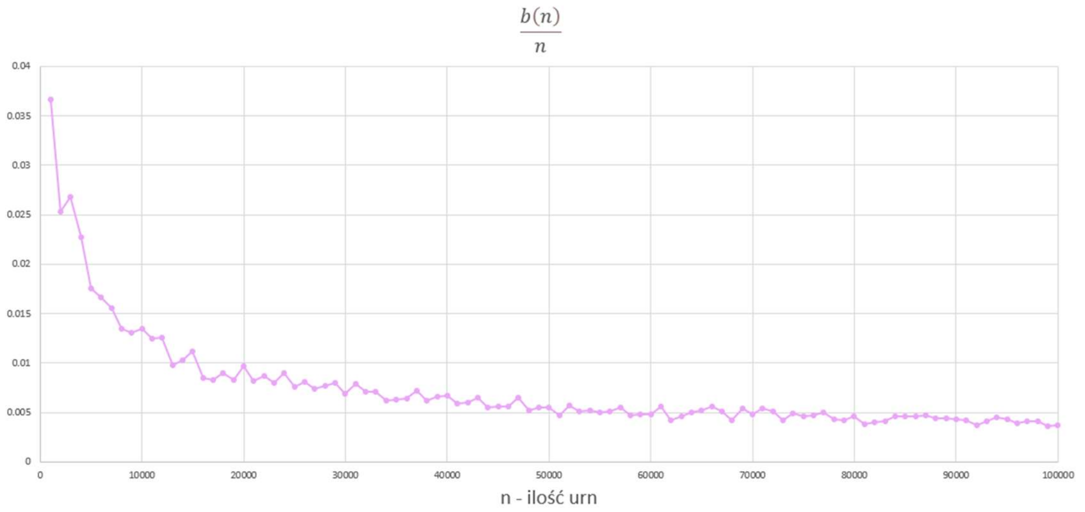


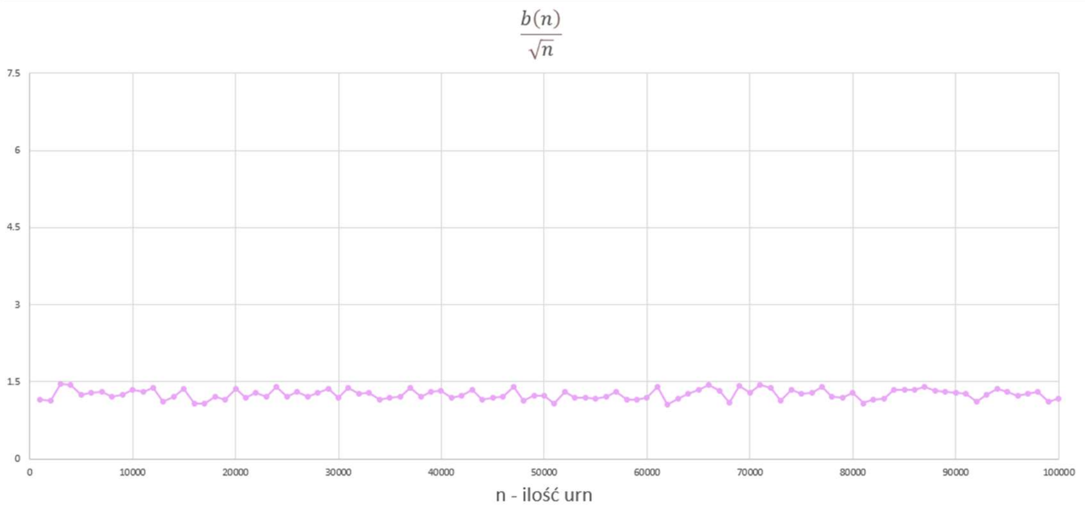
Dn –Cn – liczba rzutów od momentu Cn potrzeba do tego, zeby w kazdej urnie były co najmniej dwie kule



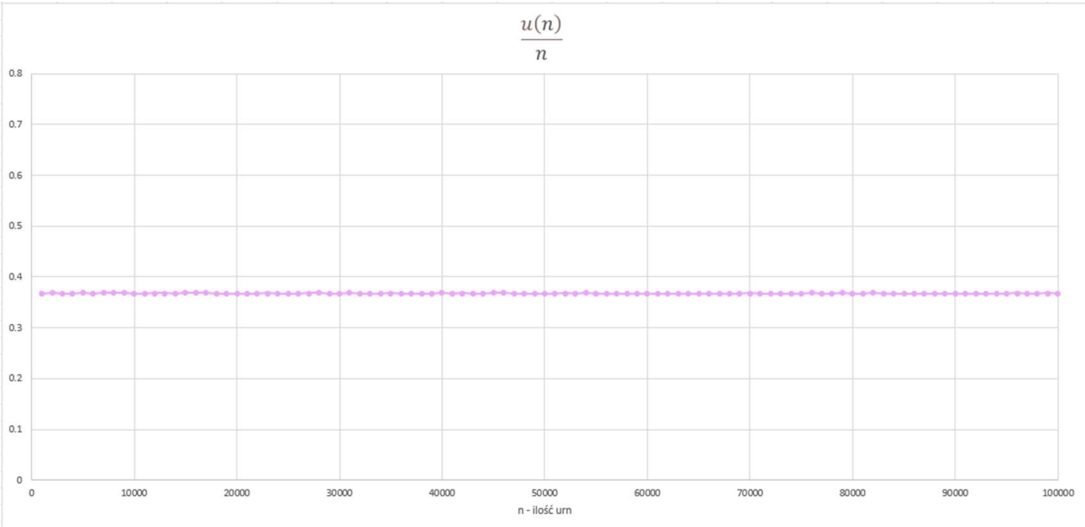
Dodatkowe wykresy:

$\frac{b(n)}{n}$ i $\frac{b(n)}{\sqrt{n}}$ $jako\ funkcja\ n$:

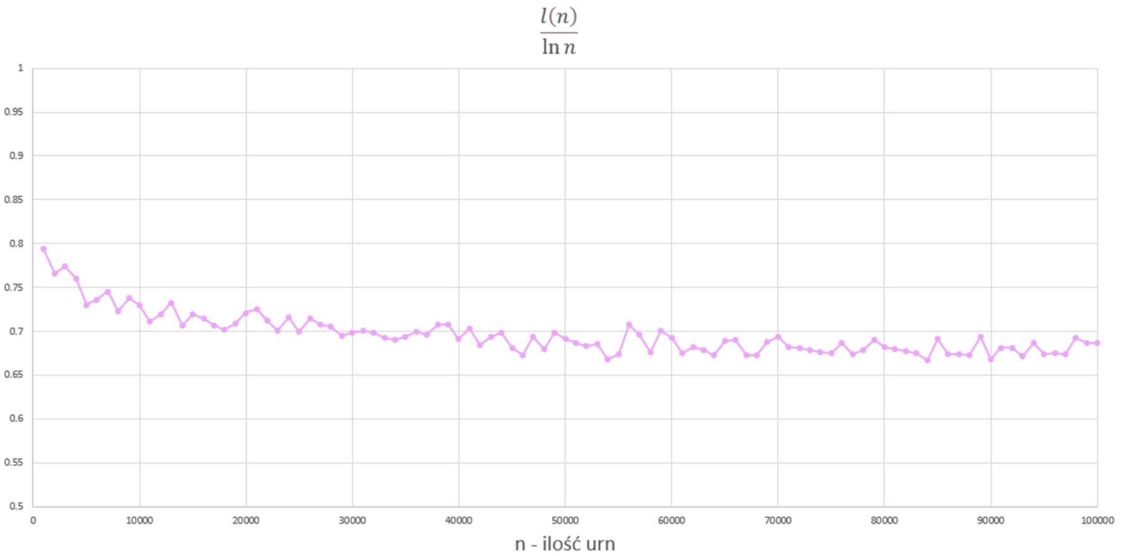


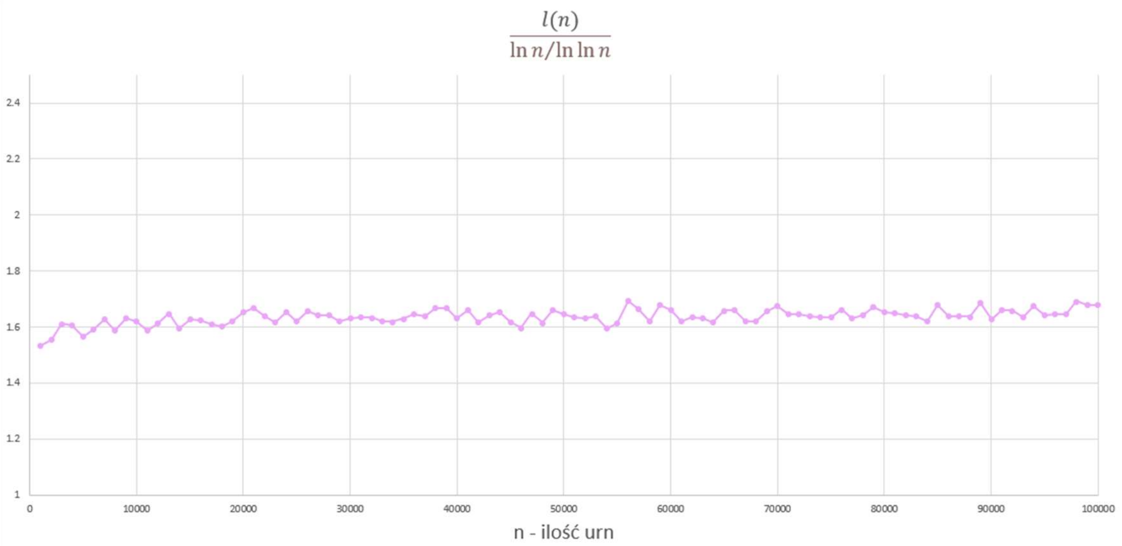


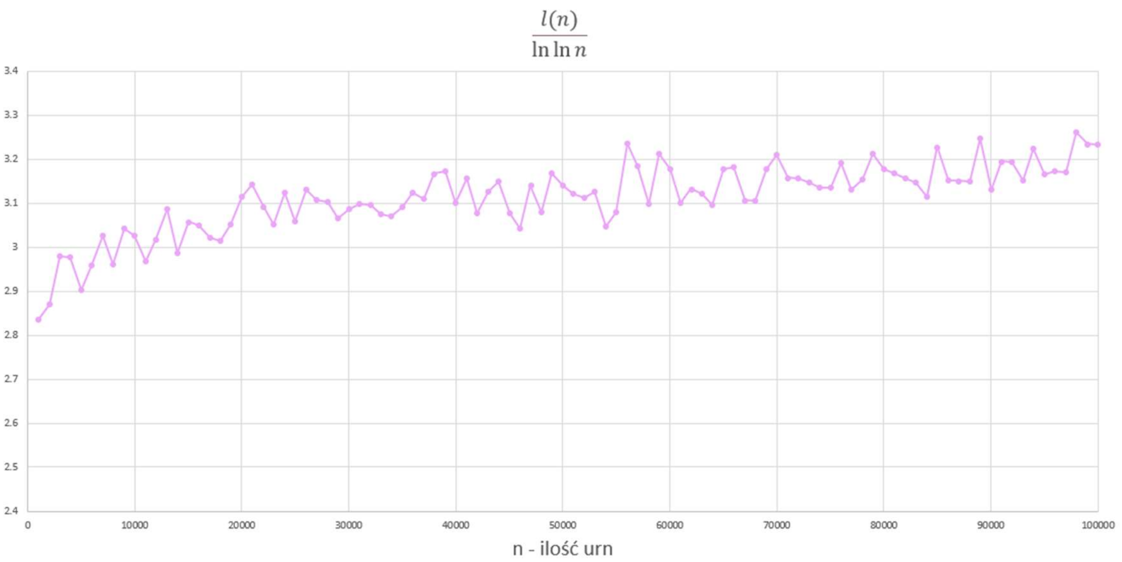
$\frac{u(n)}{n}$ $jako\ funkcja\ n$



$$\frac{l(n)}{\ln n}, \frac{l(n)}{(\ln n)\ /\ \ln \ln n}\ oraz\ \frac{l(n)}{\ln \ln n}\ jako\ funkcja\ n$$

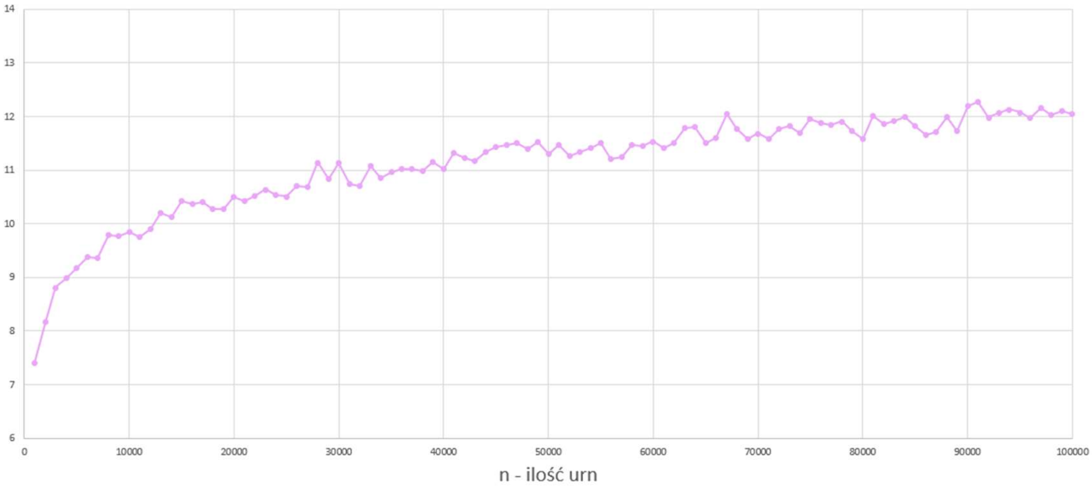




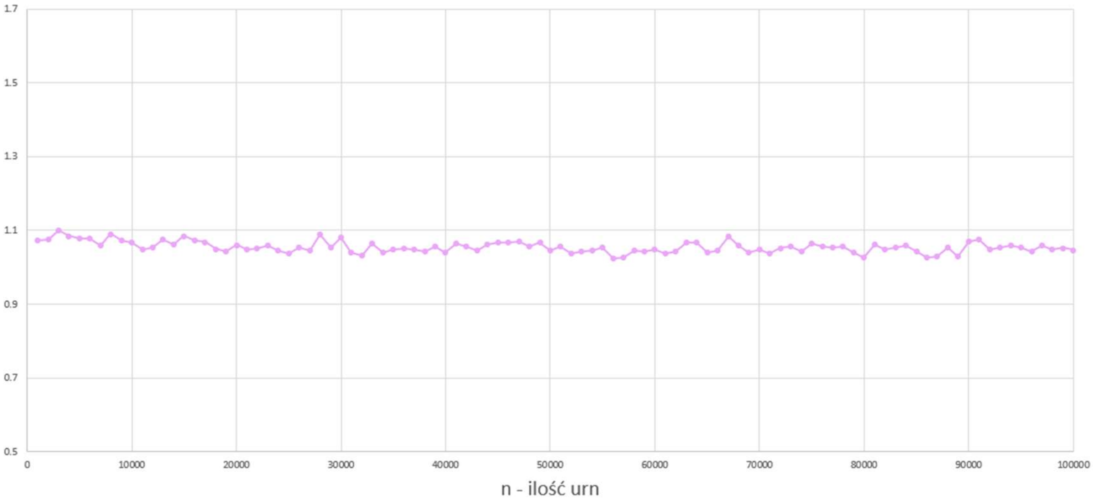


$$\frac{c(n)}{n}, \frac{c(n)}{n \ln n} \text{ oraz } \frac{c(n)}{n^2} \text{ jako funkcja } n$$

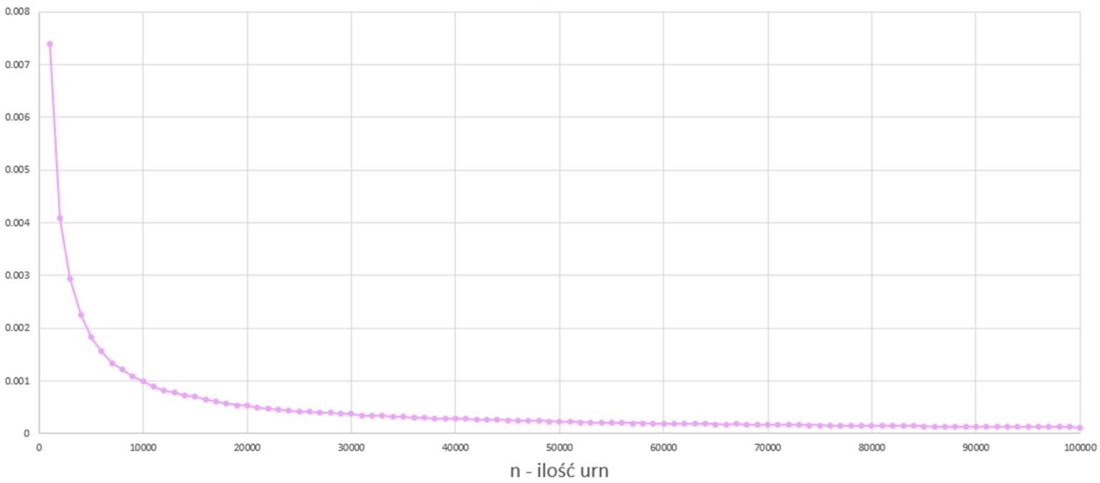
$$\frac{c(n)}{n}$$



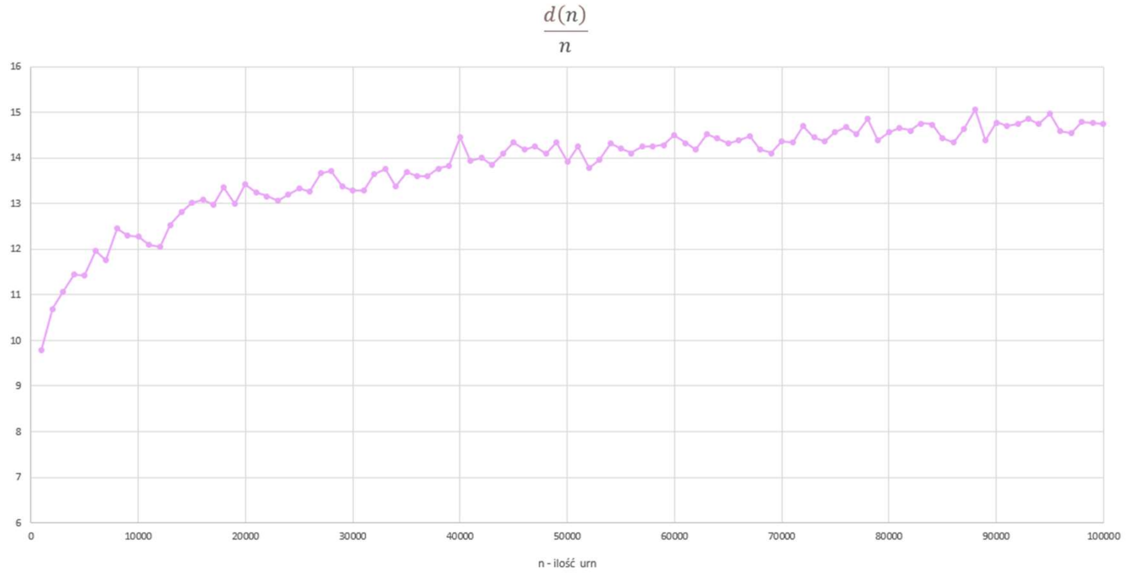
$$\frac{c(n)}{n \ln n}$$

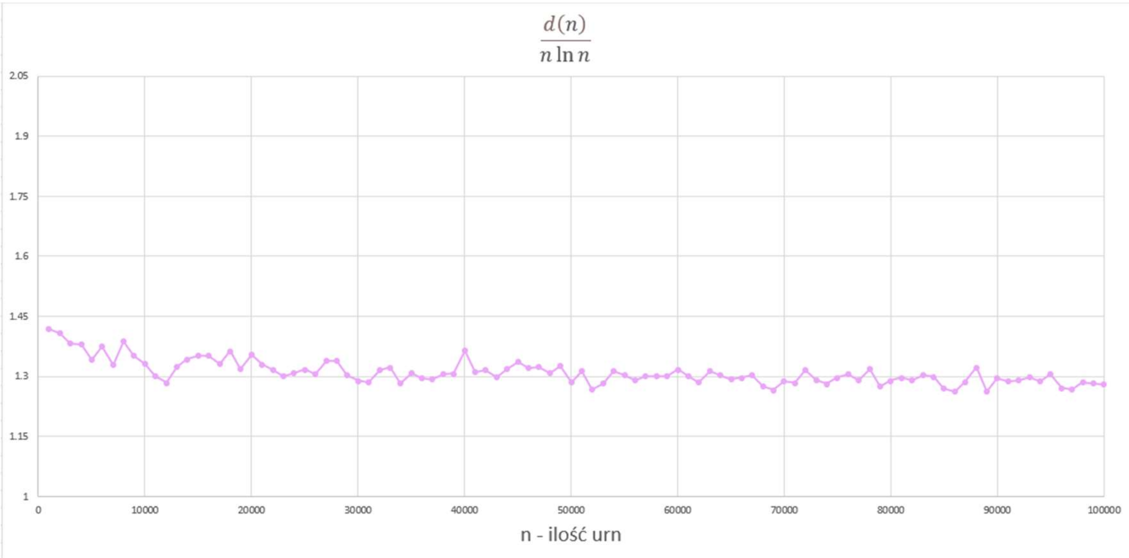


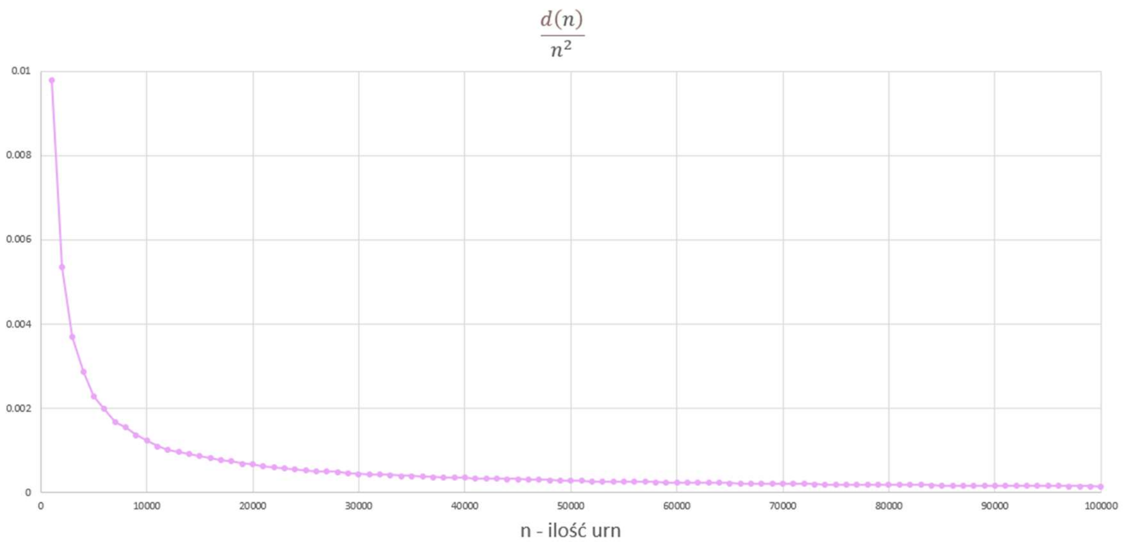
$$\frac{c(n)}{n^2}$$



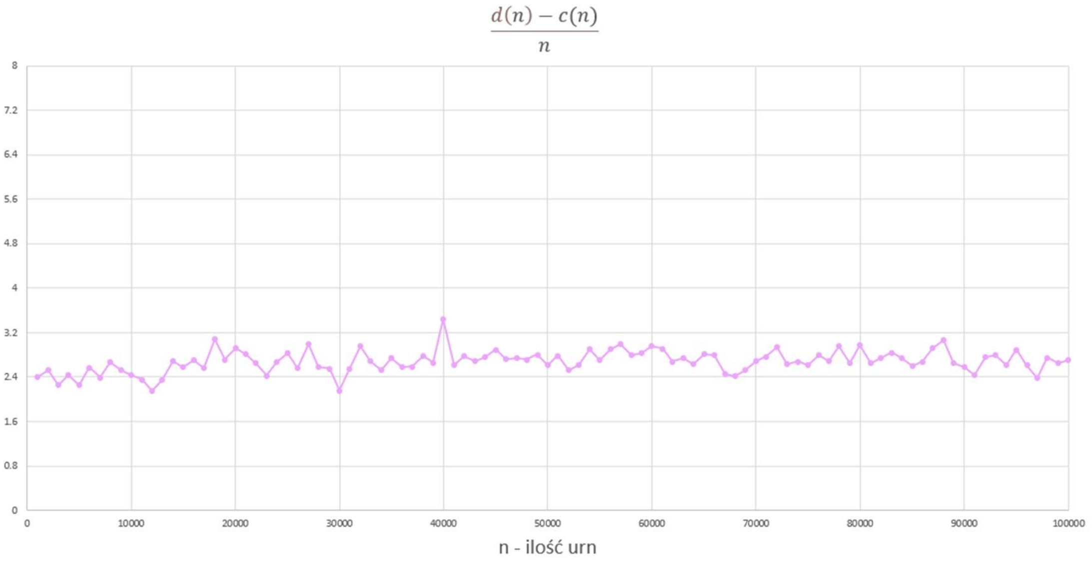
$$\frac{d(n)}{n}, \frac{d(n)}{n\,ln\,n}\ oraz\ \frac{d(n)}{n^2}\ jako\ funkcja\ n$$

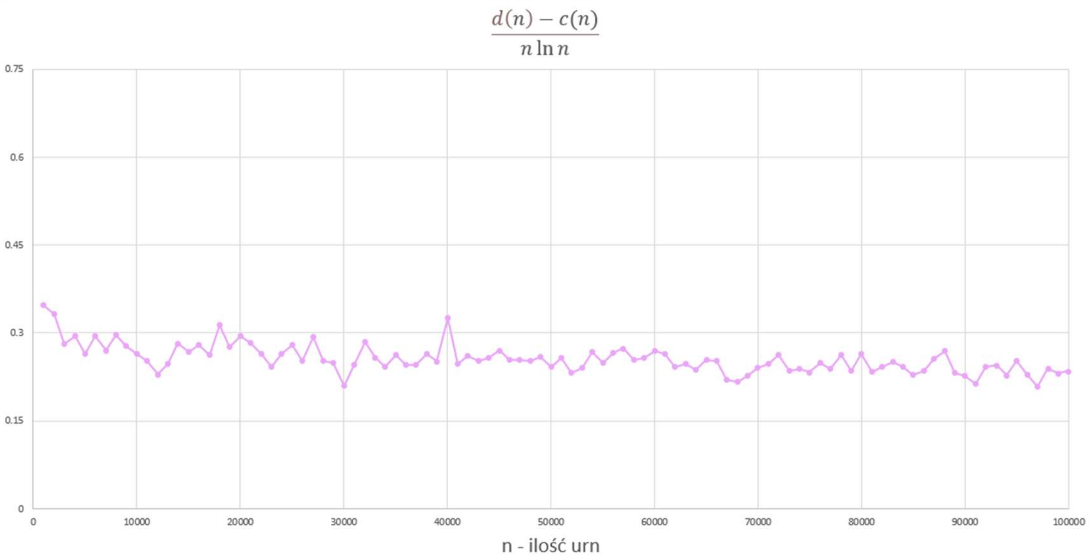


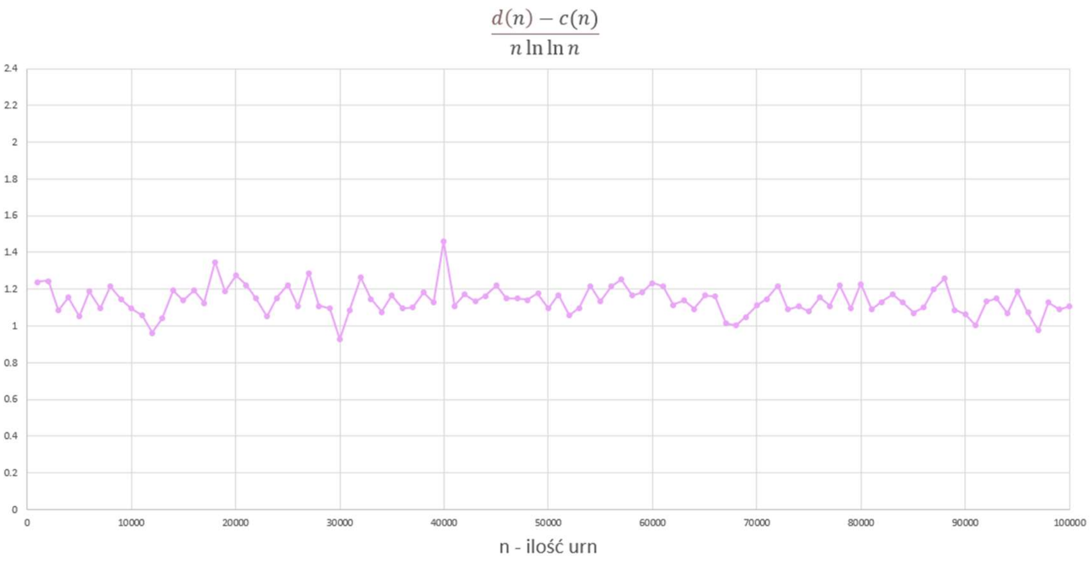




$$\frac{d(n)-c(n)}{n}, \frac{d(n)-c(n)}{n \ln n} \ oraz \ \frac{d(n)-c(n)}{n \ln \ln n} \ jako \ funckja \ n$$







Na podstawie powyższych wykresów postaw rozsądne hipotezy odnośnie asymptotyki średnich wartości badanych statystyk.

Zabierając się za analizę asymptotyki średnich wartości statystyk potrzebujemy w tym celu pojęć związanych z notacjami, dzięki którym możemy je określić.

Notacja „duże O”:

Funkcja $f(n)$ należy do $O\big(g(n)\big)$, gdy $\exists k > 0\ \exists n_0\ \exists n > n_0\colon\ |f(n)| \leq k \cdot g(n)$, co można zapisać jako granica $\underset{n\to\infty}{lim}\ sup \frac{|f(n)|}{g(n)} < \infty$

Notacja „Ω”:

Funkcja $f(n)$ należy do $\Omega\big(g(n)\big)$, gdy $\exists k > 0\ \exists n_0\ \exists n > n_0\colon\ |f(n)| \geq k \cdot g(n)$, co można zapisać jako granica $\underset{n\to\infty}{lim}\ inf \frac{|f(n)|}{g(n)} > 0$.

Notacja „Θ”:

Funkcja $f(n)$ należy do $\Theta\big(g(n)\big)$, gdy $\exists k_1 > 0 \exists k_2 > 0 \exists n_0 \forall n > n_0 \colon k_1 \cdot g(n) \leq f(n) \leq k_2 \cdot g(n)$, co można zapisać jako granice $\underset{n\to\infty}{lim}\ sup \frac{|f(n)|}{g(n)} < \infty$ oraz $\underset{n\to\infty}{lim}\ inf \frac{|f(n)|}{g(n)} > 0$.
Oznacza to, że jeżeli funkcja $f(n)$ należy do $\Theta\big(g(n)\big)$, to należy ona do $O\big(g(n)\big)$, oraz $\Omega\big(g(n)\big)$.

$b(n)$:

- $\frac{b(n)}{n}$ jest malejąca, więc $\underset{n\to\infty}{lim}\ inf \frac{b(n)}{n} = 0$
- $\frac{b(n)}{\sqrt{n}}$ osiąga wartości ≈ 1.3
- Zatem możemy stwierdzić, że $b(n) = \Theta\big(\sqrt{n}\big)$

$u(n)$:

- $\frac{u(n)}{n}$ osiąga wartości ≈ 0.37
- Zatem możemy stwierdzić, że $u(n) = \Theta(n)$

$l(n)$:

- $\frac{l(n)}{\ln n}$ jest malejąca, więc $\underset{n\to\infty}{lim}\ inf \frac{l(n)}{\ln n} = 0$
- $\frac{l(n)}{\ln \ln n}$ jest rosnąca, więc $\underset{n\to\infty}{lim}\ \sup \frac{l(n)}{\ln \ln n} = \infty$
- $\frac{l(n)}{\frac{\ln n}{\ln \ln n}}$ osiąga wartości ≈ 1.6
- Zatem możemy stwierdzić, że $l(n) = \Theta\left(\frac{\ln n}{\ln \ln n}\right)$

$c(n)$:

- $\frac{c(n)}{n}$ jest rosnąca, więc $\underset{n\to\infty}{lim}\ sup \frac{c(n)}{n} = \infty$

- $\frac{c(n)}{n^2}$ jest malejąca, więc $\liminf_{n\to\infty} \frac{c(n)}{n^2} = 0$
- $\frac{c(n)}{n \ln}$ osiąga wartości ≈ 1.1
- Zatem można powiedzieć, że $c(n) = \Theta(n \ln n)$

$d(n)$:

- $\frac{d(n)}{n}$ jest rosnąca, więc $\limsup_{n\to\infty} \frac{d(n)}{n} = \infty$
- $\frac{d(n)}{n^2}$ jest malejąca, więc $\liminf_{n\to\infty} \frac{d(n)}{n^2} = 0$
- $\frac{d(n)}{n \ln}$ osiąga wartości ≈ 1.3
- Zatem można powiedzieć, że $c(n) = \Theta(n \ln n)$

Trudno jest jednoznacznie określić asymptotykę dla $d(n) - c(n)$, gdyż ciężko wywnioskować z wykresów monotoniczność funkcji. Można by przypuszczać, że $d(n) - c(n) = \Theta(n\, ln\, ln\, n)$, gdyż wartości oscylują na niewielkim zakresie wartości, podczas gdy funkcja $\frac{d(n)-c(n)}{n}$ delikatnie rośnie, a wartości dla funkcji $\frac{d(n)-c(n)}{n\, ln\, n}$ trochę zmalały.