

Lista 5

Analiza przeżycia

Marta Stankiewicz (282244) Kacper Szmigielski (282255)

Spis treści

1	Lista 1	1
1.1	Zadanie 1	1
1.2	Zadanie 2	4
1.3	Zadanie 3	4

Spis rysunków

1	Wykres estymatorów Kaplana–Meiera dla leków A i B	2
2	Wykres estymatorów Fleminga-Harringtona dla leków A i B	3
3	Wykresy estymatorów Kaplana-Meiera dla leków A i B z ogonem	4

Spis tabel

1	Wyniki testu Shapiro–Wilka	7
---	--------------------------------------	---

1 Lista 1

1.1 Zadanie 1

Dane przedstawiają dwie grupy pacjentów, na których badano działanie leków A i B. Przedstawione wyniki prezentują czas do remisji choroby.

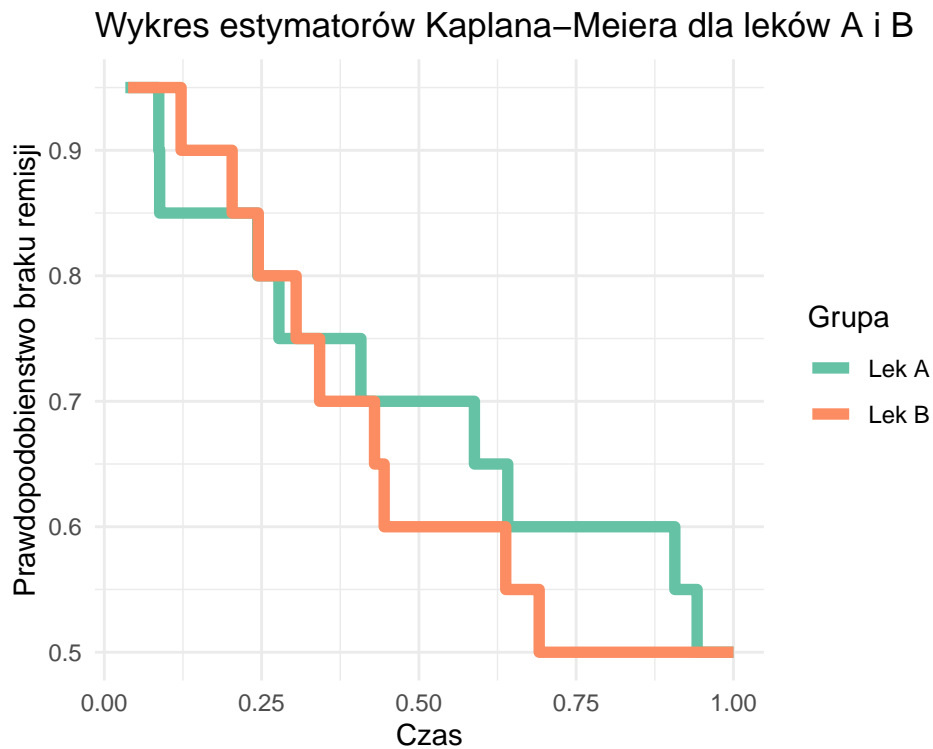
Dane opisane w zadaniu 3 z listy 2 :

Lek A : 0.03345514, 0.08656403, 0.08799947, 0.24385821, 0.27755032,
0.40787247, 0.58825664, 0.64125620, 0.90679161, 0.94222208

Lek B : 0.03788958, 0.12207257, 0.20319983, 0.24474299, 0.30492413, 0.34224462, 0.42950144, 0.44484582, 0.63805066, 0.69119721

W każdej z tych grup było jeszcze 10 osób u których nie zaobserwowano remisji choroby.

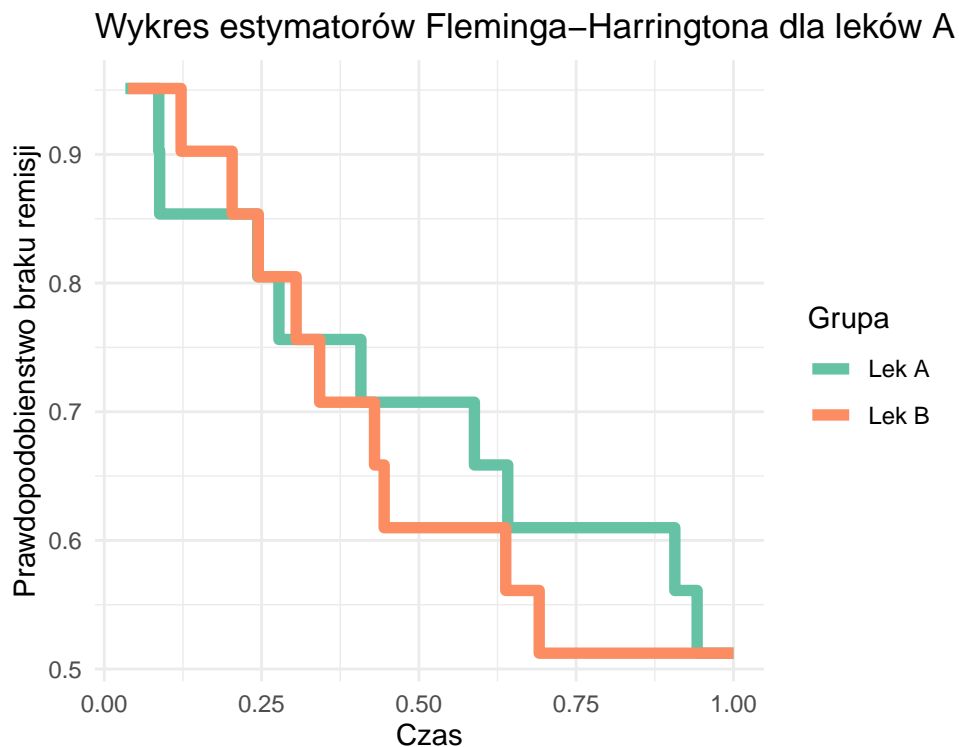
1.1.1 Wykresy Kaplana-Meiera



Rysunek 1: Wykres estymatorów Kaplana–Meiera dla leków A i B

Wykres estymatorów Kaplana–Meiera dla leków A oraz B wskazuje na szybsze osiągnięcie remisji w grupie leku B – prawdopodobieństwo pozostawania bez remisji spada do 0.5 wcześniej niż w grupie leku A. Krzywa dla leku A jest mniej stroma i przebiega wyżej, co oznacza, że po czasie $t = 0,75$ około 60% pacjentów wciąż nie osiągnęło remisji, podczas gdy w grupie leku B prawdopodobieństwo braku remisji jest już bliskie 50%

1.1.2 Wykresy Fleminga-Harringtona



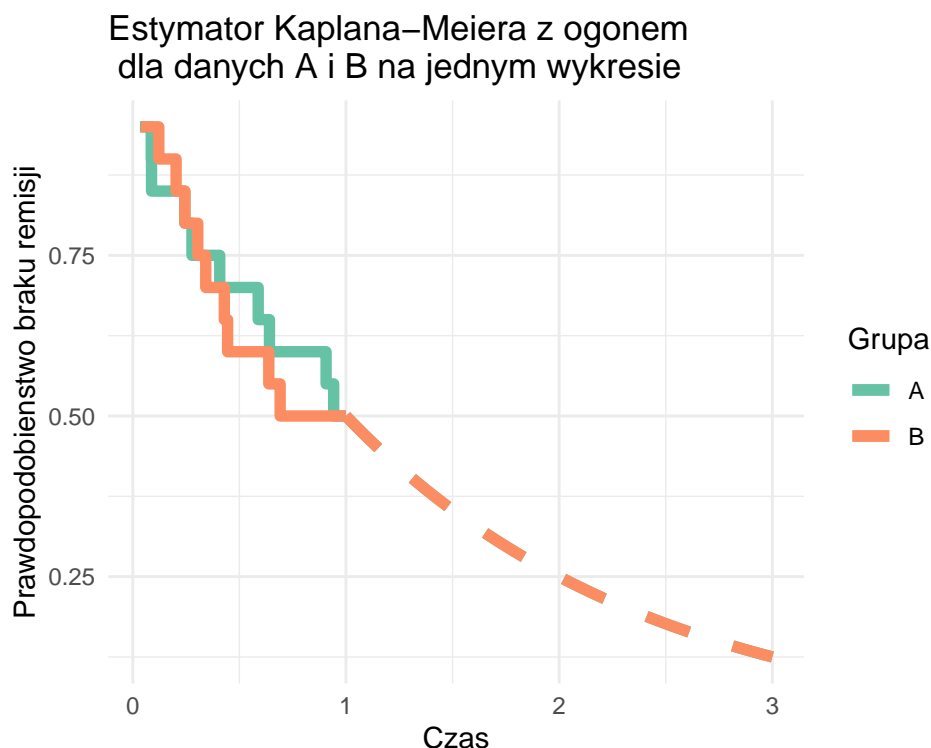
Rysunek 2: Wykres estymatorów Fleminga–Harringtona dla leków A i B

Z wykresu estymatorów Fleminga–Harringtona dla obu leków można wysnuć wnioski analogiczne jak na podstawie krzywych Kaplana–Meiera. Przebieg krzywych dla leków A i B jest bardzo zbliżony, co potwierdza szybsze osiąganie remisji w grupie leku B oraz wskazuje, że wybór estymatora (Kaplana–Meiera vs Fleminga–Harringtona) nie wpływa istotnie na interpretację wyników w tym przykładzie.

Na podstawie obu wykresów można przypuszczać, że leki A i B charakteryzują się zbliżonym profilem skuteczności, przy czym lek B wiąże się z szybszym uzyskaniem remisji. Krzywa przeżycia dla leku B spada szybciej i wcześniej osiąga niższe wartości prawdopodobieństwa braku remisji niż krzywa dla leku A, co sugeruje krótszy czas do remisji w tej grupie.

1.2 Zadanie 2

1.2.1 Wykresy estymatorów Kaplana-Meiera dla leków A i B z ogonem



Rysunek 3: Wykresy estymatorów Kaplana-Meiera dla leków A i B z ogonem

Ogon jest zdefiniowany względem definicji zaproponowanej przez Browna, Hollandera i Kowara. Na skutek czego wzór funkcji ogona w przypadku leku A jak i leku B jest taki sam i wynosi $e^{\frac{\ln \hat{S}(t^+)}{t^+} * t}$ gdzie w obu przypadkach $t^+ = \frac{1}{2}$. Ogon lepszy sposób odwzorowuje prawdę, że im dłużej pacjent jest chory tym szansa remisji jest coraz mniejsza

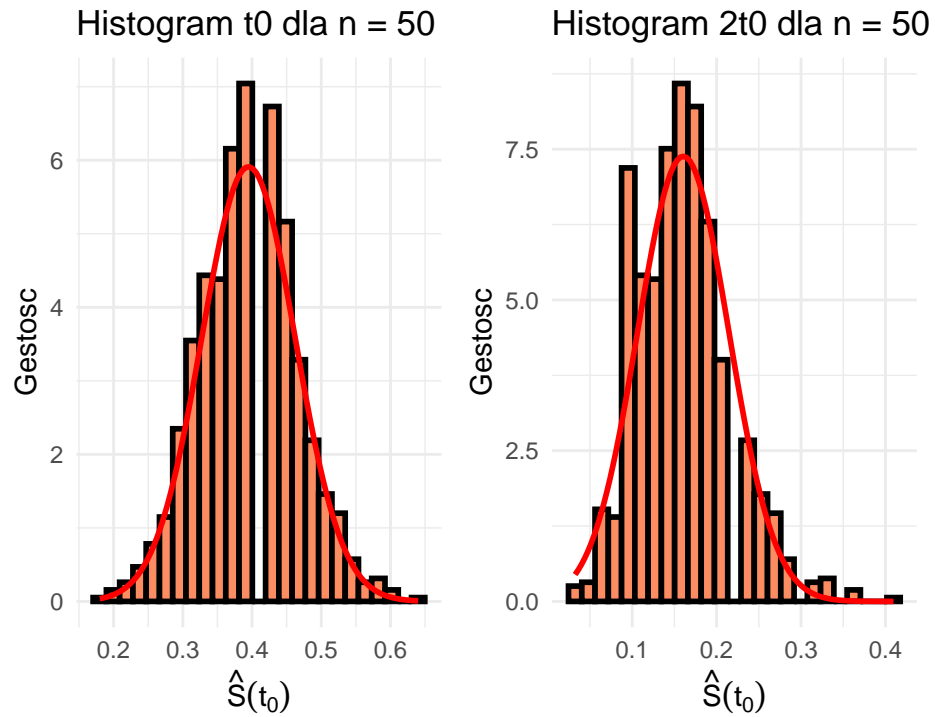
1.3 Zadanie 3

Wybrane parametry to $\alpha = 3, \lambda = 5$

Dla uogólnionego rozkładu wykładniczego

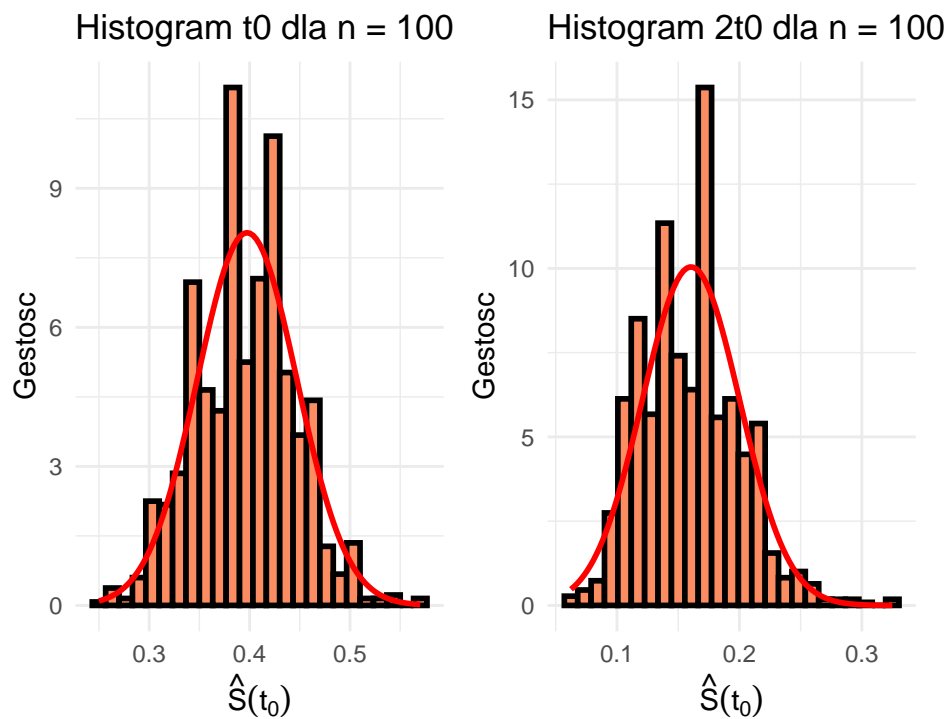
wartość oczekiwana jest równa $\mathbb{E}(X) = \frac{1}{\lambda} \sum_{k=1}^{\alpha} \frac{1}{k}$ U nas $t_0 = \mathbb{E}(X)$

1.3.1 Wykresy dla $n = 50$



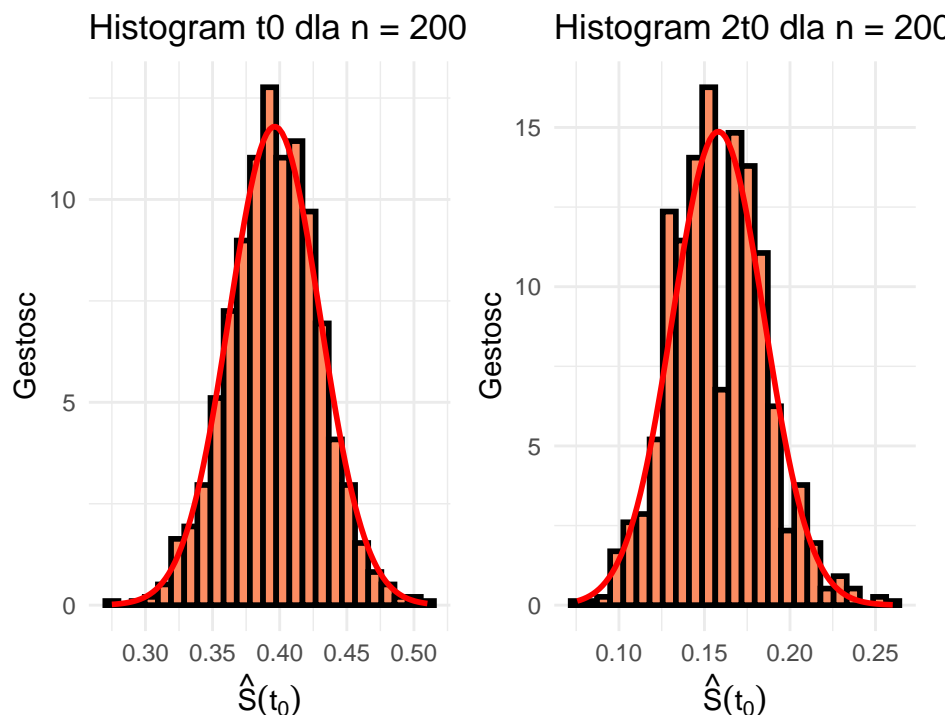
Histogramy dla t_0 i $2t_0$ ogólnie przypominają rozkład normalny, zwłaszcza w części centralnej. Widoczne są jednak lokalne odchylenia – pojedyncze słupki wyraźnie odbiegają od krzywej normalnej, co skłania do odrzucenia hipotezy o normalności rozkładów.

1.3.2 Wykresy dla $n = 100$



Tutaj analogiczne wnioski jak w przypadku próby o długości 50

1.3.3 Wykresy dla $n = 200$



Dla $n = 200$ rozkład t_0 bardzo dobrze odwzorowuje kształt rozkładu normalnego. W przypadku $2t_0$ podobieństwo również jest widoczne, choć nie tak wyraźne jak dla t_0 — pojawiają się większe lokalne odchylenia od krzywej normalnej.

Żeby potwierdzić wnioski z histrogramów, wykonamy jeszcze testy Shapiro-Wilka dla danych.

Tabela 1: Wyniki testu Shapiro-Wilka

	Próba	W	p-value
$n = 50, t_0.W$	$n = 50, t_0$	0.989640	0.000002
$n = 50, 2t_0.W$	$n = 50, 2t_0$	0.969206	0.000000
$n = 100, t_0.W$	$n = 100, t_0$	0.994940	0.002028
$n = 100, 2t_0.W$	$n = 100, 2t_0$	0.985059	0.000000
$n = 200, t_0.W$	$n = 200, t_0$	0.997119	0.069632
$n = 200, 2t_0.W$	$n = 200, 2t_0$	0.993677	0.000307

We wszystkich przypadkach statystyka $W > 0.95$, co wskazuje na bliskość z rozkładem normalnym. Jednak na podstawie p-value brak podstaw do odrzucenia normalności jedynie dla próby z $n = 200$ w punkcie t_0 .

Wnioski: Same histogramy nie dają jednoznacznej odpowiedzi co do normalności — przy mniejszych próbach ($n = 50, 100$) widoczne są lokalne odchylenia. Dla $n = 200$ w punkcie t_0 rozkład jest już bardzo zbliżony do normalnego.