

Lista 5

Analiza przeżycia

Olga Foriasz, Tomasz Warzecha

2025-10-08

Spis treści

1	Lista 5	1
1.1	Zadanie 1 - Analiza graficzna estymatorów funkcji przeżycia	1
1.2	Zadanie 2 - tworzenie wykresu estymatora Kaplana-Meiera z “ogonem”	3
1.3	Zadanie 3 - oszacowanie wartości funkcji przeżycia	3

1 Lista 5

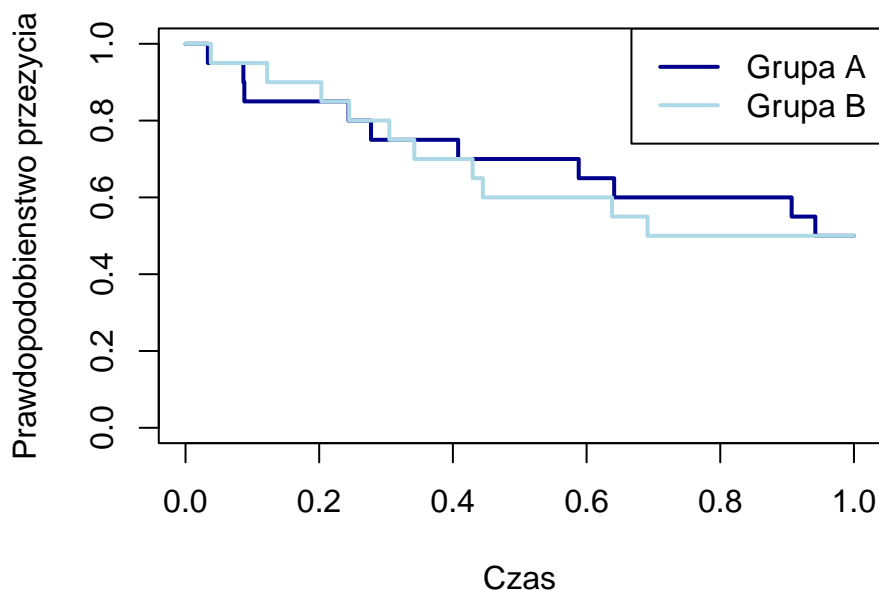
1.1 Zadanie 1 - Analiza graficzna estymatorów funkcji przeżycia

Celem zadania było graficzne przedstawienie i porównanie estymatorów funkcji przeżycia dla czasu do remisji choroby w grupach pacjentów leczonych lekiem A oraz lekiem B. Dane pobrano z zadania 3 z listy 2. Wykorzystano dwie metody nieparametryczne:

- a) Estymator Kaplana-Meiera (KM)
- b) Estymator Fleminga-Harringtona (FH)

W tym celu skorzystamy z funkcji *survfit()* z pakietu *survival*.

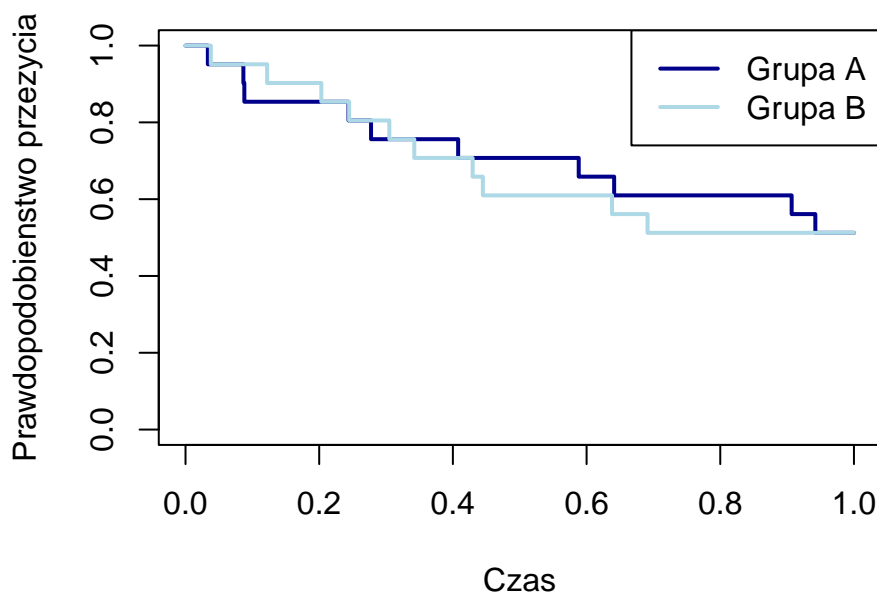
Estymator Kaplana–Meiera dla grup A i B



Na powyższym wykresie widzimy estymator Kaplana-Meiera dla grup A i B.

Teraz wyznaczmy sobie estymator Fleminga-Harringtona:

Estymator Fleminga–Harringtona dla grup A i B



Na powyższym wykresie widzimy estymator Fleminga-Harringtona dla grup A i B. Zarówno wykres estymatora Kaplana-Meiera, jak i Fleminga-Harringtona, są wizualnie niemal identyczne i prowadzą do tych samych wniosków.

Na podstawie wykresów estymatorów Kaplana-Meiera można zauważyć, że krzywe przeżycia dla grup A i B mają zbliżony przebieg. Początkowo wartości funkcji przeżycia są prawie identyczne, natomiast w środkowym zakresie czasów (około 0.3-0.7) krzywa dla grupy B przebiega nieco niżej, co sugeruje,

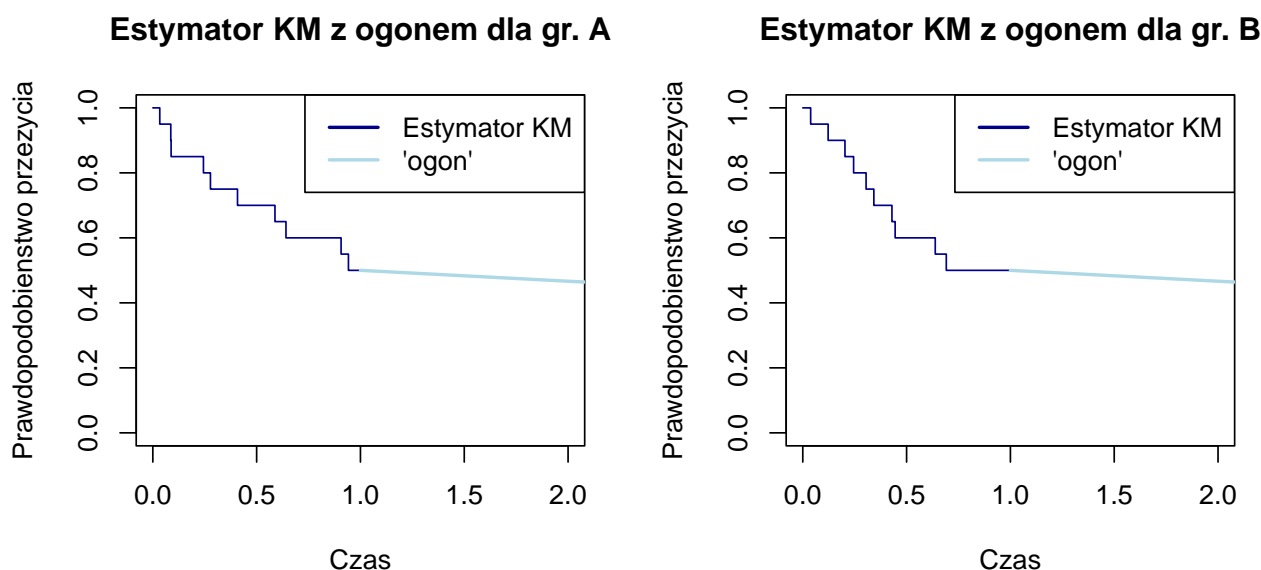
że pacjenci leczeni lekiem B częściej osiągnęli remisję wcześniej niż pacjenci z grupy A. Ostatecznie jednak obie krzywe schodzą do podobnego poziomu, co wskazuje, że końcowy odsetek remisji w obu grupach jest zbliżony. Zatem na podstawie otrzymanych wykresów można przypuszczać, że leki A i B wykazują podobne działanie, przy czym lek B może powodować nieco szybsze wystąpienie remisji we wcześniejszym okresie obserwacji.

1.2 Zadanie 2 - tworzenie wykresu estymatora Kaplana-Meiera z “ogonem”

Na początku stworzymy sobie funkcję do tworzenia wykresu estymatora Kaplana-Meiera z “ogonem” estymowanym zgodnie z propozycją Browna, Hollandera i Kowara. W tym celu na początku wyznaczmy sobie estymator Kaplana-Meiera za pomocą funkcji *survfit()* dla $t \in (0, t^+]$, następnie dla $t \in (t^+, 2t^+)$ wyznaczmy sobie “ogon” naszego estymatora za pomocą wzoru:

$$\hat{S}(t) = \exp \left[\frac{\ln \hat{S}(t^+)}{t^+} t \right],$$

Teraz mając funkcję tworzącą wykresy, narysujemy sobie wykres naszego estymatora dla danych z zad.3 z listy 2.



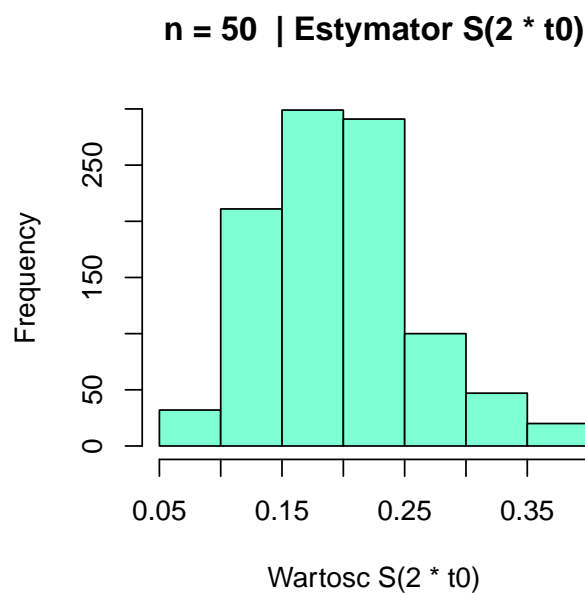
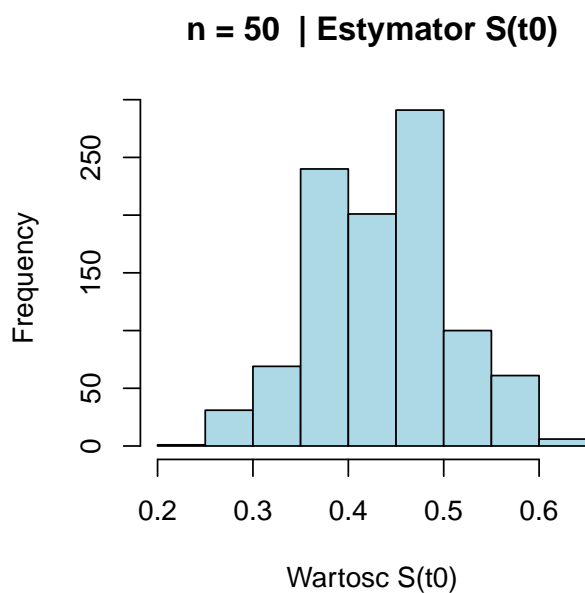
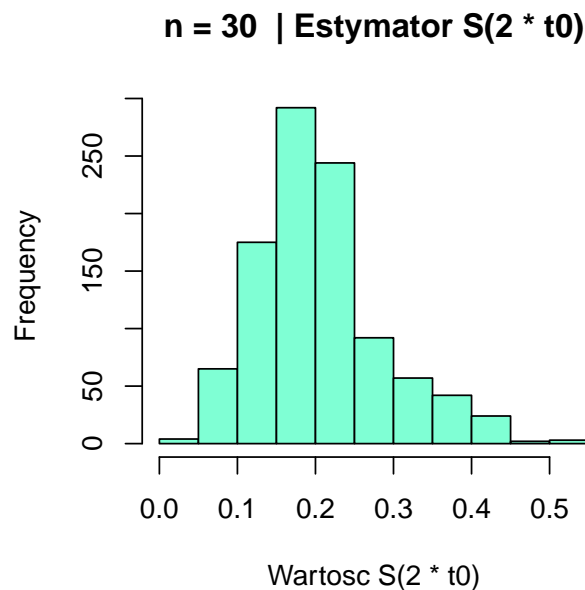
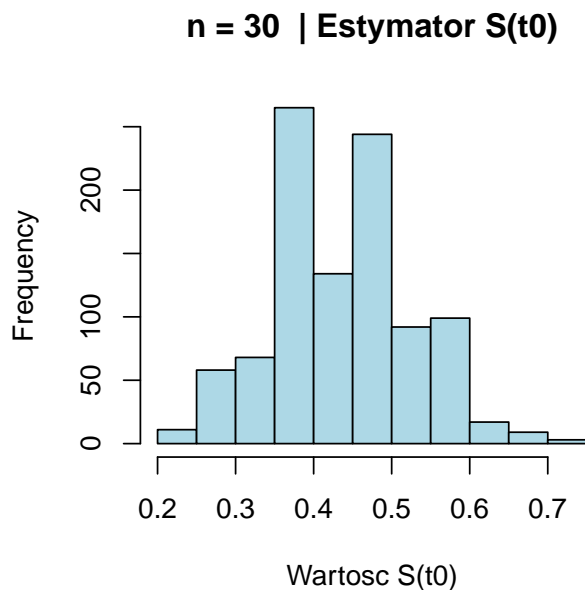
Na powyższych wykresach możemy zobaczyć wartości estymatora K-M z ogonem dla grupy A i B. Jak widać powyższe wykresy znacznie się nie różnią, ogon estymatora Kaplana-Meiera pokazuje dalszy spadek funkcji przeżycia po ostatnich zaobserwowanych zdarzeniach, co oznacza, że prawdopodobieństwo przeżycia (braku remisji) nadal maleje w miarę upływu czasu, mimo braku nowych obserwacji zdarzeń.

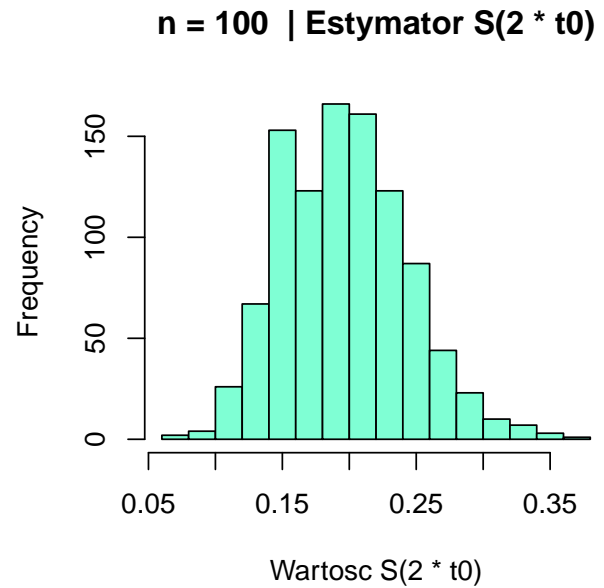
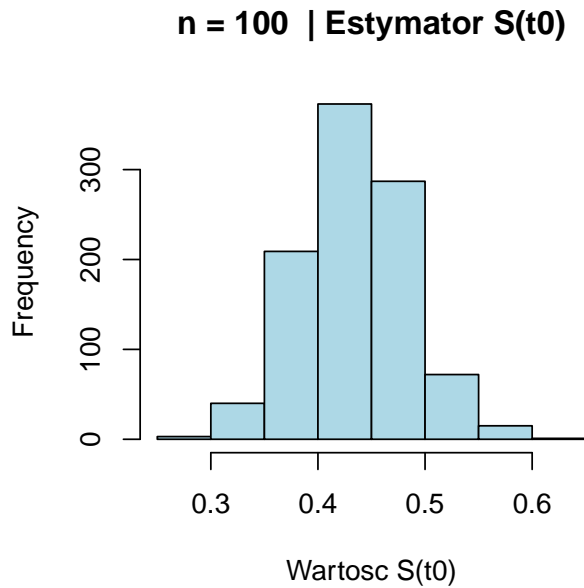
1.3 Zadanie 3 - oszacowanie wartości funkcji przeżycia

W ramach zadania przeprowadzono symulację Monte Carlo, generując $M = 1000$ prób cenzurowanych I-go typu z uogólnionego rozkładu wykładniczego $GE(\lambda, \alpha)$ dla wybranych wartości parametrów $\alpha = 4$ oraz $\lambda = 1$. Rozmiary prób wynosiły odpowiednio $n = 30, 50, 100$. Wartość t_0 przyjęto w przybliżeniu równą wartości oczekiwanej rozkładu, a następnie wyznaczono estymatory

funkcji przeżycia w punktach t_0 oraz $2t_0$, korzystając z estymatora Kaplana–Meiera z tzw. ogonem estymowanym metodą Browna, Hollandera i Kowara.

Dla każdego n uzyskano 1000 estymacji wartości $\hat{S}(t_0)$ i $\hat{S}(2t_0)$, a ich rozkłady przedstawiono na histogramach poniżej.





Na podstawie uzyskanych histogramów zauważono, że wraz ze wzrostem liczby obserwacji rozkłady estymatora Kaplana–Meiera stają się coraz bardziej symetryczne i skupione wokół wartości średniej. Dla małych prób ($n = 30$) rozkłady są bardziej rozproszone i lekko skośne, natomiast dla większych prób ($n = 100$) przypominają rozkład normalny. Oznacza to, że wraz ze wzrostem liczby danych zmienność estymatora maleje, a jego rozkład zbliża się do rozkładu normalnego.

Dodatkowo można zauważyć, że wartości estymatora w punkcie $2t_0$ są niższe niż w punkcie t_0 , co jest zgodne z interpretacją funkcji przeżycia — prawdopodobieństwo „przeżycia” (braku zdarzenia) maleje wraz z upływem czasu.

Wyniki te potwierdzają przypuszczenie, że estymator Kaplana–Meiera ma własność asymptotycznej normalności zarówno w punkcie t_0 , jak i $2t_0$.