

Analiza przeżycia

Laboratorium 6

Alicja Jokiel-Rokita

15 listopada 2020

1 Zadania do sprawozdania 2, część 2

Zadania dotyczą tych samych danych, które opisane były na poprzedniej liście. Poniżej, dla przypomnienia, przedstawiony jest ten opis.

Dane w pliku *CzasDoDializy.csv* dotyczą pacjentów leczonych na niewydolność nerek. Zmienna *Czas* informuje o czasie, w którym zaczęto dializować pacjenta od momentu rozpoznania choroby (w latach), jeżeli zmienna *Obcięcie* wynosi 1, lub czasie obserwacji pacjenta od momentu rozpoznania choroby (w latach), gdy zmienna *Obcięcie* wynosi 0. Pozostałe cztery zmienne dotyczą genotypu pacjenta. Interesuje nas czas od momentu rozpoznania choroby do momentu, gdy dializa jest konieczna do przeprowadzenia. W poniższych zadaniach czas od momentu rozpoznania choroby do dializy pełni rolę “czasu życia”. Oczywiście niektóre dane są cenzurowane i zakładamy, że czas cenzurowania jest niezależny od czasu do wystąpienia zdarzenia. Nie przyjmujemy dodatkowych założeń dotyczących “czasu życia”, tzn. nie przyjmujemy postaci rozkładu tego czasu.

1. Naszkicować wykres estymatora Kaplana-Meiera funkcji przeżycia w całej badanej grupie i w podgrupach ze względu na wartość zmiennej *Arg25Pro* wraz z realizacjami przedziałów ufności (asymptotycznych punktowo) wybierając różne typy tych przedziałów.
2. Zastanowić się nad wyborem typu przedziałów ufności i uzasadnieniem tego wyboru.
3. Sformułować wnioski wynikające z porównania uzyskanych wykresów w podgrupach *Arg25Pro*.

2 Zadanie dodatkowe

Wyprowadzić wzory na dolną i górną granicę asymptotycznych punktowo przedziałów ufności dla wartości $S(t)$ w oparciu o przekształcenie $g(x) = \log(x)$.