

ANALIZA PRZEŻYCIA

Lista 9

Sprawozdanie dotyczy analizy danych *lung* dostępnych w pakiecie *survival*. Dane dotyczą pacjentów z zaawansowanym rakiem płuc. Zbiór zawiera następujące informacje o 228 pacjentach

- *inst*: kod instytucji,
- *time*: czas przeżycia w dniach,
- *status*: cenzura (1 – cenzura, 2 – śmierć)

oraz ich siedmiu charakterystykach

- *age*: wiek w latach,
- *sex*: płeć (1-mężczyzna, 2-kobieta),
- *ph.ecog*: skala sprawności ECOG wg. lekarza (0-sprawność prawidłowa, 5-zgon),
- *ph.karno*: skala sprawności Karnofsky'ego wg. lekarza (100-sprawność prawidłowa, 0-zgon),
- *pat.karno*: skala sprawności Karnofsky'ego wg. pacjenta,
- *meal.cal*: kalorie spożywane podczas posiłków,
- *wt.loss*: utrata masy ciała w ciągu ostatnich sześciu miesięcy.

W danych występują brakujące wartości (57), głównie w zmiennej *meal.cal* (47). Zakładamy, że te brakujące dane wystąpiły całkowicie losowo.

Na podstawie powyżej opisanych danych, przyjmując, że czas przeżycia ma rozkład Weibulla, wykonać poniższe zadania.

1. Oszacować parametry modelu przyspieszonego czasu awarii, przyjmując za zmienną zależną zmienną *time*, a za charakterystyki zmienne: *age*, *sex*, *ph.ecog*, *ph.karno*.
2. Podać interpretację współczynników modelu z zadania 1.
3. Wyznaczyć oszacowanie funkcji przeżycia (w dniach) odpowiadającej rozkładowi czasu życia kobiety w wieku 70 lat o charakterystyce *ph.ecog*=1 i *ph.karno*=90 i na podstawie uzyskanego oszacowania obliczyć szacowane prawdopodobieństwo, że czas życia tej kobiety będzie większy niż 300 dni.

4. Narysować wykres oszacowanej w zadaniu 3. funkcji przeżycia.
5. Oszacować parametry modelu proporcjonalnych hazardów, przyjmując za zmienną zależną zmienną *time*, a za charakterystyki zmienne: *age*, *sex*, *ph.ecog*, *ph.karno*.
6. Podać interpretację współczynników modelu z zadania 5.
7. Wyznaczyć oszacowanie funkcji hazardu odpowiadającej rozkładowi czasu życia
 - (a) kobiety w wieku 70 lat o charakterystyce *ph.ecog*=1 i *ph.karno*=90,
 - (b) kobiety w wieku 70 lat o charakterystyce *ph.ecog*=2 i *ph.karno*=90,narysować wykresy tych funkcji i wykresy logarytmów tych funkcji. Czy na podstawie tych wykresów możemy mieć wątpliwości co do przyjętego modelu proporcjonalnych hazardów?
8. Wyznaczyć oszacowanie funkcji przeżycia (w dniach) odpowiadającej rozkładowi czasu życia
 - (a) kobiety w wieku 70 lat o charakterystyce *ph.ecog*=1 i *ph.karno*=90,
 - (b) kobiety w wieku 70 lat o charakterystyce *ph.ecog*=2 i *ph.karno*=90i na podstawie tego oszacowania obliczyć szacowane prawdopodobieństwo, że czas życia kobiet o powyżej podanych charakterystykach będzie większy niż 300 dni. Oszacowane prawdopodobieństwo dla kobiety o charakterystykach opisanych w punkcie (a) porównać z prawdopodobieństwem uzyskanym w zadaniu 3.
9. Narysować wykres oszacowanej w zadaniu 8 punkt (a) funkcji przeżycia i porównać go z wykresem z zadania 4.

Alicja Jokiel-Rokita
9 grudnia 2025