Lista 11

Analiza przeżycia

Marta Stankiewicz (282244) Kacper Szmigielski (282255)

Spis treści

Spis rysunków

Spis tabel

1 Zadanie 1

1.1 Treść

Oszacować parametry modelu proporcjonalnych szans, przyjmując za zmienną zależną zmienną time, a za charakterystyki zmienne: age, sex, ph.ecog, ph.karno. Uwaga. Zadanie to można wykonać korzystając np. z funkcji prop.odds biblioteki timereg lub funkcji nltm biblioteki o tej samej nazwie pakietu R. Należy jednak zwrócić uwagę na inną definicję modelu proporcjonalnych szans niż była podana na wykładzie wykorzystywaną w tej drugiej funkcji.

1.2 Rozwiązanie

```
model<-prop.odds(Event(time, cause = status) ~ age + factor(sex) + factor(ph.ecog) + ph
           data = dane,n.sim=500,profile=1)
```

2 Zadanie 2

2.1 Treść

Podać interpretację współczynników modelu z zadania 1

2.2 Rozwiązanie

	estimate
age	0.0023530
factor(sex)2	-0.5613368
factor(ph.ecog)1	0.3199014
factor(ph.ecog)2	1.1477675
factor(ph.ecog)3	1.8507250
ph.karno	-0.0022423

Korzystając z poniższych wzorów możemy podać interpretację:

$$\theta_0(t) = \frac{1 - S_0(t)}{S_0(t)}$$

$$\ln\left(\frac{\theta_{z_1}(t)}{\theta_{z_2}(t)}\right) = \beta^T(z_1 - z_2)$$

- Zmienna `age` ma współczynnik
$$\gamma > 0,$$
co oznacza, że wraz ze wzrostem wieku rosną **szanse (odds) wystąpienia zdarzenia do czasu**
$$t$$
(przy stałych pozostałych zmiennych). Równoważnie:
$$\exp(\gamma) > 1.$$

- Zmienna `sex` (kategoryczna) ma współczynnik
$$\gamma < 0$$
dla poziomu porównywanego do poziomu referencyjnego (bazowego). Oznacza to, że w tej grupie **szanse (odds) zajścia zdarzenia do czasu**
$$t$$
są mniejsze niż w grupie bazowej, tj.
$$\exp(\gamma) < 1.$$

- Zmienna `ph.ecog` (kategoryczna) ma współczynniki
$$\gamma > 0$$
dla poszczególnych poziomów w porównaniu do poziomu referencyjnego (np.
$$1 \text{ vs } 0,\ 2 \text{ vs } 0,\ 3 \text{ vs } 0$$
). Wskazuje to, że osoby z wyższym `ph.ecog` mają **większe szanse (odds) wystąpienia zdarzenia do czasu**
$$t$$
niż osoby z poziomu bazowego, czyli
$$\exp(\gamma) > 1.$$

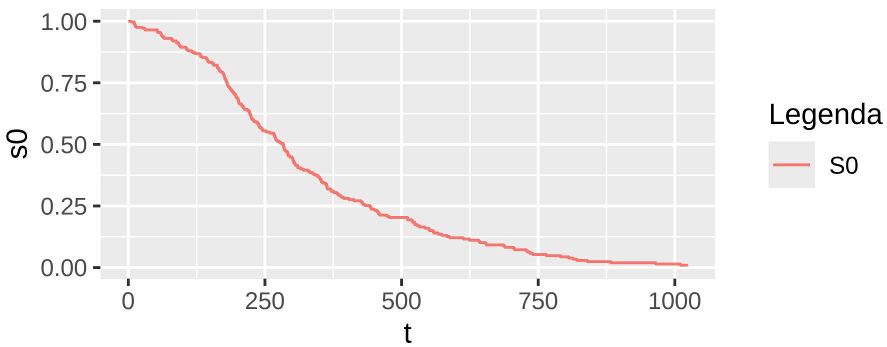
- Dla zmiennej ciągłej `ph.karno` otrzymano
$$\gamma < 0,$$
co oznacza, że wraz ze wzrostem `ph.karno` **maleją szanse (odds) zajścia zdarzenia do czasu**
$$t$$
(tj.
$$\exp(\gamma) < 1$$
). Jeśli jednak w poprzedniej analizie (Lista 10) współczynnik dla `ph.karno` był **nieistotny statystycznie**, to również tutaj wniosek o kierunku efektu należy traktować ostrożnie: przy braku istotności znak współczynnika może zmieniać się między dopasowaniami/modelami i nie powinien być interpretowany jako stabilny efekt.

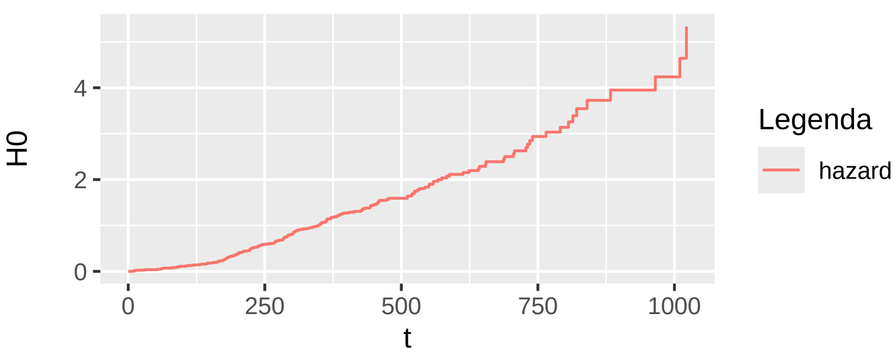
3 Zadanie 3

3.1 Treść

Wyznaczyć oszacowanie bazowej skumulowanej funkcji hazardu i bazowej funkcji przeżycia odpowiadającej rozkładowi czasu życia (przyjmując model opisany w zadaniu 1). Uwaga. To zadanie można wykonać korzystając np. z funkcji nltm biblioteki o tej samej nazwie.

3.2 Rozwiązanie



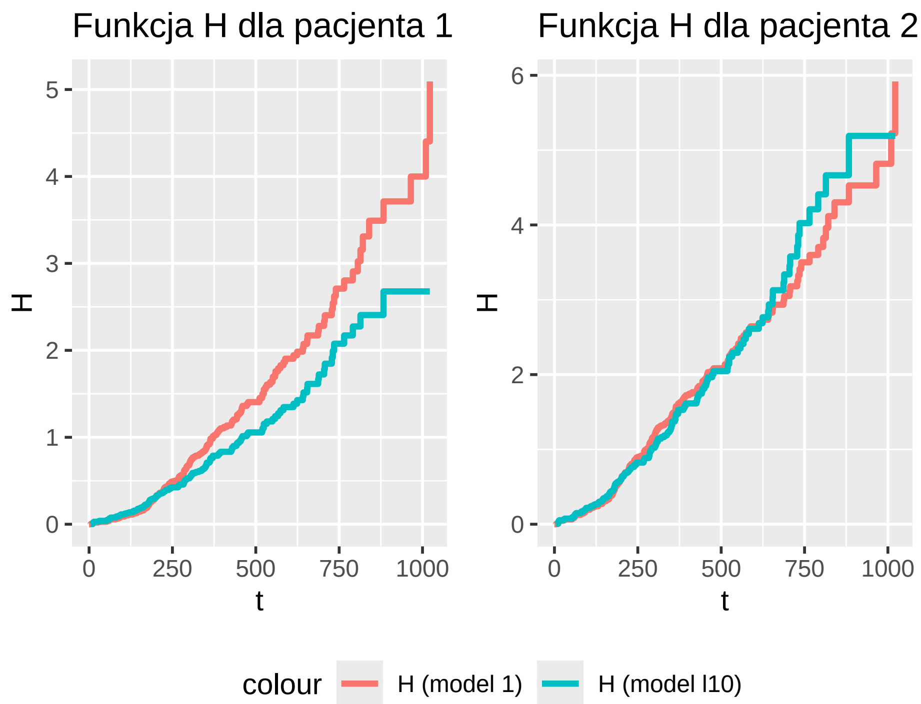


4 Zadanie 4

4.1 Treść

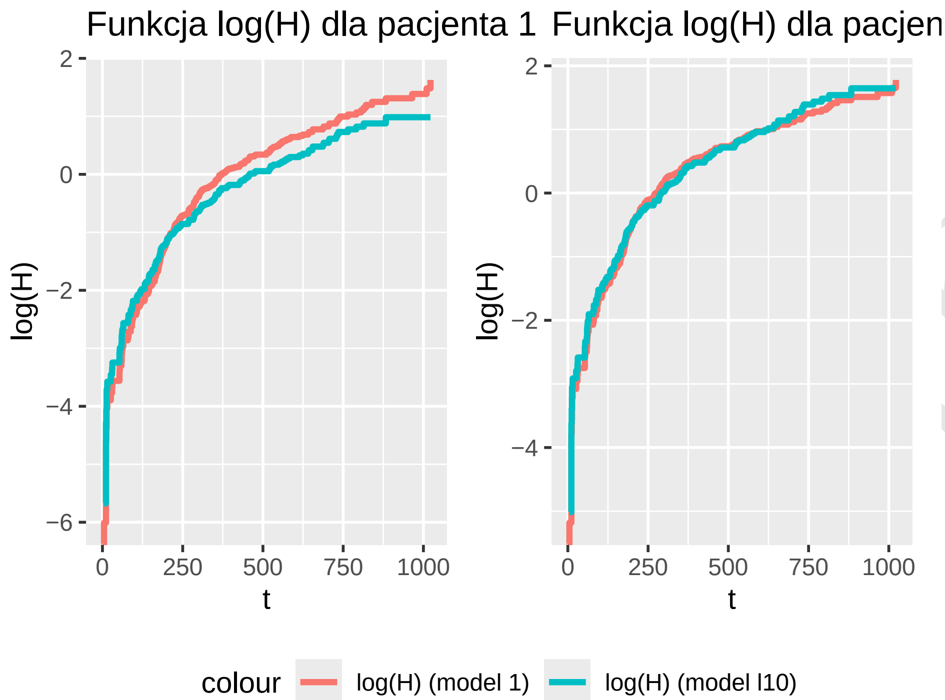
Wyznaczyć oszacowanie skumulowanej funkcji hazardu odpowiadającej rozkładowi czasu życia (przyjmując model opisany w zadaniu 1)

4.2 Rozwiązanie



Widzimy, że w przypadku pacjenta 1 skumulowany hazard rośnie wyraźnie szybciej, gdy użyjemy modelu proporcjonalnych szans, niż wtedy, gdy korzystamy z modelu proporcjonalnych hazardów Coxa. Oznacza to, że model proporcjonalnych szans przewiduje dla tego pacjenta gorsze przeżycie (większe skumulowane ryzyko w czasie) niż model Coxa.

W przypadku pacjenta 2 krzywe są bardzo podobne mniej więcej do 625 dnia, a później zaczynają się rozjeżdżać. W dalszej części obserwacji różnice między modelami rosną, jednak interpretację samej końcówki należy traktować ostrożnie, ponieważ w ogonie czasu często zostaje niewiele obserwacji i pojedyncze zdarzenia potrafią powodować wyraźne „skoki” krzywej.

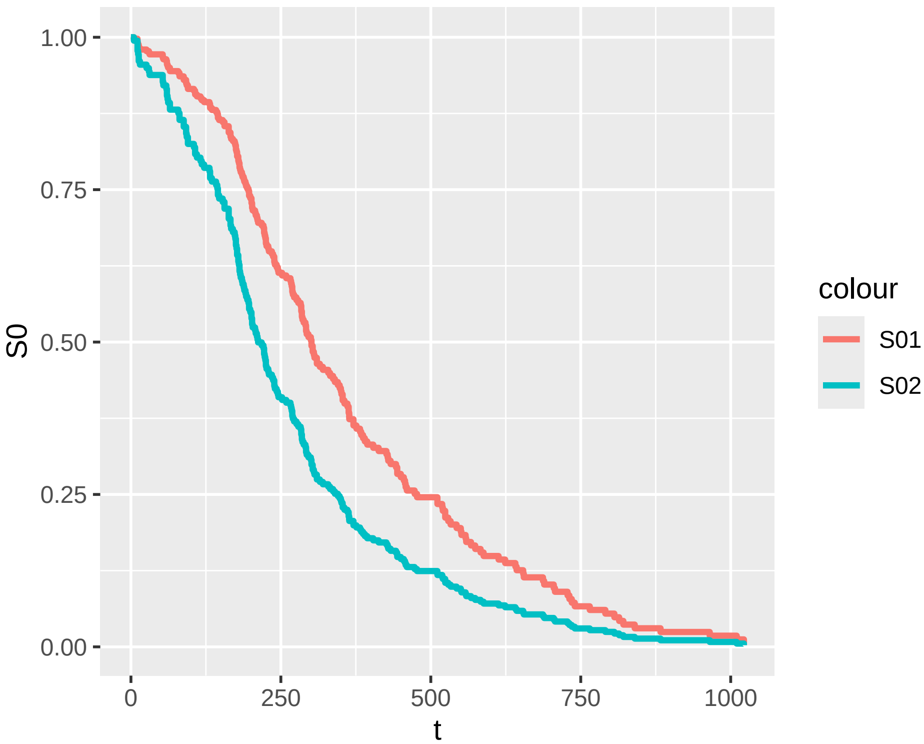


Analogiczne wnioski możemy wysnuć także na podstawie wykresów logarytmów tych funkcji, ponieważ logarytm jest funkcją ściśle rosnącą i dla dodatnich wartości zachowuje relacje. # Zadanie 5

4.3 Treść

Wyznaczyć oszacowanie funkcji przeżycia (w dniach) odpowiadającej rozkładowi czasu życia

4.4 Rozwiązanie



Z wykresów widać, że funkcja przeżycia dla pacjenta 2 maleje znacznie szybciej, co jest zgodne z charakterystyką danych.

Prawdopodobieństwa tego ,że czas życia będzie większy od 300 , to

$$S_02(300)$$

i

$$S_01(300)$$

Dla pacjenta 1 to prawdopodobieństwo wynosi:

```
## [1] 0.5033501
```

Jest to spora szansa.

A na liście 10 otrzymaliśmy:

```
## [1] 0.5901344
```

Prawdopodobieństwo dla modelu proporcjonalnych hazardów coxa jest większe.

Dla pacjenta 2 to prawdopodobieństwo wynosi:

```
## [1] 0.3067809
```

W stosunku do pacjenta pierwszego widać spadek aż o 20 punktów procentowych na liście 11.

```
## [1] 0.3597682
```

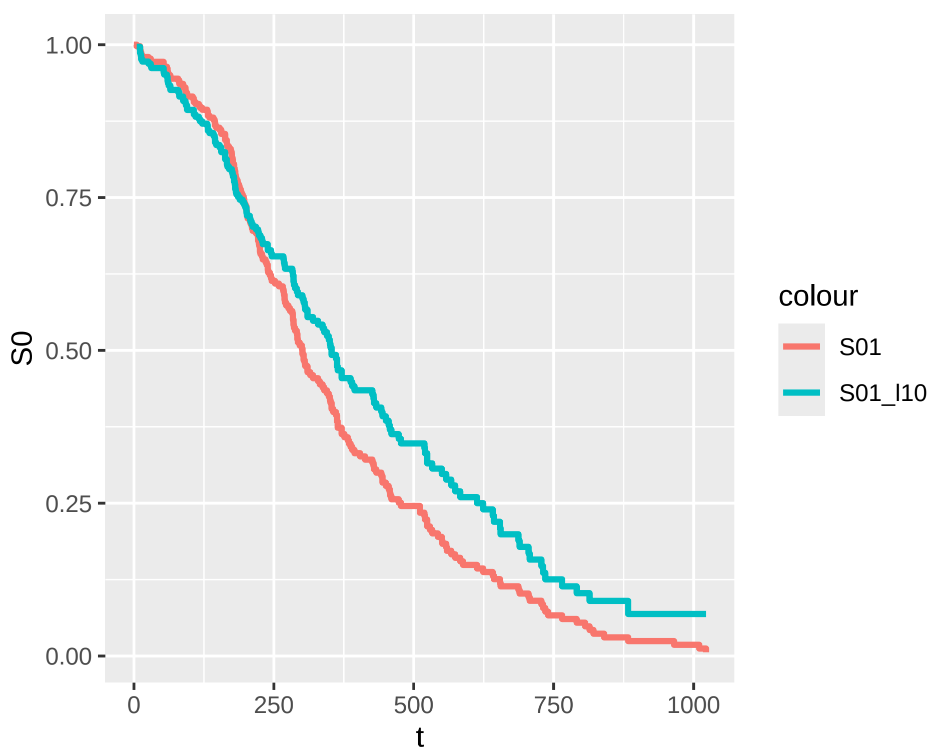
Prawdopodobieństwo dla modelu proporcjonalnych hazardów coxa jest większe.

5 Zadanie 6

5.1 Treść

Narysować wykres oszacowanej w zadaniu 5 punkt (a) funkcji przeżycia i porównać go z wykresem z zadania 6 z listy 10.

5.2 Rozwiązanie



Widać, że funkcja przeżycia jest większa dla modelu proporcjonalnych hazardów coxa, niż w obecnym modelu proporcjonalnych szans.