

**Zadanie 1.** Dopuszczając możliwość, że efekt leczenia (zmienna  $rx$ ) jest różny dla różnych typów zdarzeń, konstruuje model warunkowy PWP uwzględniając typy zdarzeń, czyli uwzględniając zmienne:  $rx1$ ,  $rx2$ ,  $rx3$ ,  $rx4$ . Powoduje to, że zmienna warstwowująca określająca liczbę zdarzeń do obecnego czasu -  $enum$  nie jest potrzebna w modelu, gdyż jest liniowa zależna ze zmiennymi  $rx_i$ . Dodatkowo z racji na występowanie przedziałów czasowych stosuję klastrowanie ze względu na zmienną  $patid$ . Odpowiedni syntaks i wyniki dostępne są poniżej:

Call:

```
coxph(formula = Surv(time1, time2, status) ~ rx1 + rx2 + rx3 +
      rx4 + size + number + cluster(patid), data = bladder3)
```

	coef	exp(coef)	se(coef)	robust se	z	p
rx1	-0.6834	0.505	0.2202	0.2178	-3.1373	0.0017
rx2	-0.3061	0.736	0.2359	0.2228	-1.3739	0.1700
rx3	0.4652	1.592	0.2552	0.2531	1.8377	0.0660
rx4	0.2781	1.321	0.2846	0.2616	1.0629	0.2900
size	0.0052	1.005	0.0694	0.0631	0.0824	0.9300
number	0.1348	1.144	0.0518	0.0521	2.5881	0.0097

Likelihood ratio test=50.5 on 6 df, p=3.75e-09 n= 178, number of events= 112

Wynik testu dla statystyki największej wiarygodności ma  $p$ -wartość poniżej 0.05. Oznacza to, że są statystycznie istotne podstawy by uznać zmienne  $rx1$  i  $number$  za istotne w tym modelu, gdyż  $p$ -wartość dla tych zmiennych jest mniejsza niż 0.05. Oznacza to, że efekt leczenia jest różny dla zdarzenia typu  $rx1$  w stosunku do pozostałych zdarzeń  $rx2, rx3, rx4$ . Patrząc na oszacowanie wartości  $e^\beta$  (0.505) przy zmiennej  $rx1$  widzimy, że jest ona mniejsza od 1, co oznacza, że typ leczenia zmniejsza prawdopodobieństwo występowania zdarzenia  $rx1$ .

Podobne wyniki otrzymać można dla testu Walda, testu Score i statystyki odpornej (robust) :

Tabela 1: Podsumowanie modelu dla danych bladder3.

	Test	Stopnie swobody	p-wartość
Wald	52.900	6.000	0.000
Score	60.354	6.000	0.000
Robust	25.099	6.000	0.000

Sprawdzając założenia proporcjonalnych hazardów przeprowadzam formalny test Schoenfelda, którego wyniki umieszczone w tabeli (2) nie dają podstaw do odrzucenia hipotez o stałości w czasie współczynników przy zmiennych  $rx1$ ,  $rx2$ ,  $rx3$ ,  $rx4$ . Zatem założenia proporcjonalnych hazardów są spełnione. Dodatkowo estymator odporny dla oszacowania wariancji współczynnika przy  $rx1$  jest blisko zwykłemu oszacowaniu, co nie daje podstaw do niepokoju. Estymator odporny daje nawet lepsze oszacowanie!

Tabela 2: Podsumowanie testu Shoenfelda.

	rho	$\chi^2$	p-wartość
rx1	-0.094	0.923	0.337
rx2	0.017	0.029	0.865
rx3	-0.091	0.845	0.358
rx4	-0.016	0.024	0.876
size	-0.062	0.418	0.518
number	0.040	0.178	0.673
Global	NA	3.890	0.691

Zatem można stwierdzić, że w tym przypadku są silne statystyczne podstawy by uznać zmienną  $rx1$  za istotną w tym modelu i istotnie zmniejszane jest prawdopodobieństwo pojawienia się zdarzenia  $rx1$  w zależności od typu leczenia.

**Zadanie 2.** Dla danych dotyczących czasu do wznowy po uzyskaniu remisji dla par chorych na białaczkę, konstruuje model brzegowy WLW, uwzględniając to, że czasy podane są w postaci `total time` (dlatego model WLW). Dodatkowo wiedząc, że pacjenci byli dobierani w pary tak, by pochodzili z tego samego ośrodka medycznego i mieli ten sam typ uzyskanej remisji, stosuję `frailty(patid)` aby uwzględnić efekt sparowania pacjentów. Dodatkowo z racji na niespełnione założenia o proporcjonalności hazardu dla zmiennej `partial` warstwuję ją. Odpowiedni syntaks i wyniki dostępne są poniżej:

Call:

```
coxph(formula = Surv(time, status) ~ frailty(patid) + strata(partial) +
      treat, data = leukemia)
```

	coef	se(coef)	se2	Chisq	DF	p
frailty(patid)				0.0	0	9.5e-01
treat	-1.71	0.437	0.437	15.3	1	9.3e-05

Iterations: 6 outer, 35 Newton-Raphson

Variance of random effect= 5e-07 I-likelihood = -68.6

Degrees of freedom for terms= 0 1

Likelihood ratio test=17.8 on 1 df, p=2.53e-05 n= 42

Jak widać, patrząc na wynik testu dla statystyki największej wiarygodności, są głębokie podstawy by uznać zmienną `treat` za istotną w tym modelu, gdyż p-wartość tego testu jest mocno mniejsza niż 0.05. Oznacza to, że na czas do wznowy po remisji ma wpływ aktywny lek (`treat = 1`) i wydłuża on czas do wznowy, gdyż współczynnik przy tej zmiennej  $e^{\beta} = e^{-1.71} = 0.1808 < 1$ . Oszacowanie wariancji współczynnika i odpornego estymatora wariancji współczynnika są bliskie sobie i mało się różnią, co wskazuje na dobre dopasowanie modelu - a raczej nie daje podstaw do niepokoju.

Założenia proporcjonalnych hazardów są dla zmiennej `treat` w tym modelu spełnione, co ukazują Tabela (3):

Tabela 3: Podsumowanie testu Shoenfelda dla modelu WLW.

	rho	$\chi^2$	p-wartość
treat	-0.041	0.049	0.825

Zatem, ponieważ p-wartości w modelu WLW są mniejsze od zakładanego poziomu istotności badania 0.05, można stwierdzić, że w tym przypadku są istotnie statystyczne podstawy by uznać zmienną `treat` za istotną w tym modelu. Istnieją również istotne statystycznie efekty w parach - zmienna `frailty(patid)`.

**Zadanie 3.** Dla danych dotyczących czasów do pojawienia się guzów gruczołów mlekowych u 48 samic szczurów, konstruuje model proporcjonalnych średnich AG z racji na przedziałową strukturę danych:

Call:

```
coxph(formula = Surv(time1, time2, status) ~ cluster(ratid) +
      treat, data = mammary)
```

	coef	exp(coef)	se(coef)	robust se	z	p
treat	-0.685	0.504	0.152	0.212	-3.23	0.0012

Likelihood ratio test=21.9 on 1 df, p=2.85e-06 n= 252, number of events= 210

Sprawdziłem dane. Nie występują osobniki, u których odnotowano więcej niż 13 guzów. Model odrzucił hipotezę zerową i dał silne podstawy do stwierdzenia istotnych statystycznie wpływów zmiennej `treat` na wydłużenie czasu do pojawienia się guza. Zmniejszyło się prawdopodobieństwo wystąpienia zdarzenia, gdyż współczynnik przy zmiennej ma wartość  $e^\beta = e^{-0.685} = 0.504 < 1$ , która jest mniejsza od 1.

Podobne wyniki otrzymać można dla testu Walda, testu Score i statystyki odpornej (robust) :

Tabela 4: Podsumowanie modelu AG.

	Test	Stopnie swobody	p-wartość
Wald	10.450	1.000	0.001
Score	21.040	1.000	0.000
Robust	8.027	1.000	0.005

Założenia proporcjonalnych hazardów dla zmiennej `treat` są spełnione o czym można się przekonać przeprowadzając formalny test Shoenfelda, patrz Tabela (5):

Tabela 5: Podsumowanie testu Shoenfelda dla modelu AG.

	rho	$\chi^2$	p-wartość
treat	-0.038	0.604	0.437