

Wykład “Analiza przeżycia”
Praca domowa 3
Do oddania przed 28/05/2014
Punktów do zdobycia: 25

W raporcie (do czterech stron A4 + ekstra strona na syntaks) podaj niezbędne rezultaty oraz ich interpretację dla wszystkich zadań poniżej.

1. (5 pkt.) Model warunkowy (PWP) dla danych dotyczących raka pęcherza, przedstawiony na slajdzie 52 (sesja 46), zakłada ten sam efekt leczenia (zmienna rx) dla wszystkich typów zdarzeń. Skonstruuj i dopasuj do danych model, który dopuszcza możliwość, że efekt ten jest różny dla różnych typów zdarzeń. Przy użyciu tego modelu przeprowadź test hipotezy zerowej, że efekt leczenia jest taki sam dla wszystkich typów zdarzeń. Zakładamy, że efekt zmiennych size i number jest stały dla wszystkich typów zdarzeń. Podaj odpowiednie wyniki, ich interpretację, oraz użyty syntaks.

2. (5 pkt.) Pliki “leukemia.sas7bdat” i “leukemia.csv” zawierają dane dotyczące czasu do wznowy po uzyskaniu remisji dla par chorych na białaczkę:

patid - id pary
partial - typ remisji (0 - całkowita, 1 - częściowa)
treat - leczenie (0 - placebo, 1 - aktywny lek)
time - czas obserwacji
status - wskaźnik zdarzenia (1 - wznowa, 0 - cenzurowanie)

Pacjenci byli dobierani w pary tak, by pochodzili z tego samego ośrodka medycznego i mieli ten sam typu uzyskanej remisji. Losowo wybrany pacjent z każdej z par otrzymywał placebo, podczas gdy drugi był leczony nowym lekiem. Dopasuj do danych model PH pozwalający na ocenę efektu leczenia i uwzględniający efekt typu remisji. Uzasadnij wybór modelu; podaj odpowiednie wyniki, ich interpretację, oraz użyty syntaks.

3. (5 pkt.) Pliki “mammary.sas7bdat” i “mammary.csv” zawierają dane dotyczące czasów do pojawienia się guzów gruczołów mlekowych u 48 samic szczurów:

ratid - id zwierzęcia
time1 - początek obserwacji od poprzedniego zdarzenia (0 dla 1-ego zdarzenia)
time2 - koniec obserwacji dla kolejnego zdarzenia
status - wskaźnik zdarzenia (1 - pojawienie się guza, 0 - cenzurowanie)

Wszystkie samice były obserwowane przez okres 182 dni. Liczba guzów dla poszczególnych samic wahała się od 0 do 13. W niektórych przypadkach obserwowano pojawienie się dwóch i więcej guzów tego samego dnia. Dopasuj do danych model proporcjonalnych średnich. Podaj odpowiednie wyniki, ich interpretację, oraz użyty syntaks.

4. (10 points) Plik „hodgkin.csv” zawiera dane dla 865 chorych na ziarnicę złośliwą (chorobę Hodgkina), którzy byli leczeni radioterapią lub radioterapią łączoną z chemioterapią:

stnum - id pacjenta
male - wskaźnik płci (0 - kobieta, 1 - mężczyzna)
cmt - leczenie (0 - radioterapia, 1- radioterapia + chemioterapia)
mediast - zajęcie śródpiersia (0 - brak, 1 - małe, 2 - duże)
nodes - zajęcie narządu pozalimfatycznego (0 - nie, 1 - tak)
clinstg - stopień zaawansowania (1 - I, 2 - II)
age30 - wiek w chwili rozpoczęcia leczenia (0 >= 30, 1 - <30)
firsttime - czas do zdarzenia (miesiące)
event - wskaźnik zdarzenia (0 - cenzurowanie, 1 - zdarzenie)
event_cause - typ zdarzenia (0 - cenzurowanie, 1 - inny pierwotny nowotwór, 2 - zgon lub nawrót)

Obserwowano czas do pierwszego zdarzenia. Zdarzeniem mógł być nawrót choroby, zgon lub pojawienie się innego pierwotnego nowotworu (tzn. innego histologicznie niż ziarnica złośliwa). Interesuje nas efekt leczenia na czas do pojawienia się innego pierwotnego nowotworu. Wykonaj odpowiednią analizę danych, która uwzględni również efekt innych czynników klinicznych. Uzasadnij wybór modelu/modeli, podaj odpowiednie wyniki, ich interpretację, oraz użyty syntaks.