

Kolokwium ze Statystycznych modeli liniowych i nieliniowych

Inżynieria i analiza danych III rok, grupa 3, sem. zimowy 2024/25

2025-01-29

Uwaga. Ten plik zawiera dwie strony.

ZADANIE 1. Pobierz dane *airquality* z biblioteki *datasets*. Zapoznaj się z opisem danych zamieszczonym w oknie Help programu R-Studio. Usuń ze zbioru braki danych. Możesz np. skorzystać z funkcji *base::na.omit()*. Następnie wykonaj polecenia.

(A) Zbuduj model liniowy za pomocą, którego można wyjaśnić zmienność stężenia ozonu w powietrzu w zależności od siły wiatru, temperatury powietrza i nasłonecznienia. Na podstawie zbudowanego modelu oceń wpływ każdej zmiennej objaśniającej na stężenie ozonu w powietrzu. (6 p.)

(B) Na podstawie modelu zbudowanego w punkcie (A) wyznacz stężenie ozonu w powietrzu dla największych wartości zmiennych objaśniających. (3 p.)

(C) Na podstawie poznanych testów na wykładzie sprawdź, czy w zbudowanym modelu w punkcie (A) zostało naruszone założenie o normalności i braku korelacji błędów. (6 p.)

ZADANIE 2. W zbiorze *Mroz* biblioteki *car* znajdują się dane dotyczące zatrudnienia zameężnych kobiet. Zbiór zawiera 8 zmiennych

lfp – zatrudnienie, czynnik z poziomami no i yes,

k5 – liczba dzieci do 5 roku życia,

k618 – liczba dzieci w wieku od 6 do 18 lat,

age – wiek w latach,

wc – uczęszczanie na studia, czynnik z poziomami no i yes,

hc – uczęszczanie męża na studia, czynnik z poziomami no i yes,

lwg – logarytm oczekiwanej stawki wynagrodzenia,

inc – dochód rodziny bez dochodu kobiety.

Budując odpowiedni model odpowiedz na pytania

(A) Jaki wpływ na zatrudnienie kobiety (lfp) ma to, czy ona uczęszcza na studia (wc)? (4 p.)

(B) Jaki wpływ na zatrudnienie kobiety (lfp) ma dochód rodziny (inc)? (4 p.)

Do modelu włącz wszystkie zmienne z rozważanego zbioru.

ZADANIE 3. Zbiorze *greenhouse* biblioteki *agricolae* znajdują się trzy zestawy danych *greehouse1*, *greehouse2* i *greehouse3*. Zapisz w zmiennej *gh1* dane znajdujące się zestawie *greehouse1*. Wszystkie polecenia wykonaj dla tak utworzonego zbioru danych.

(A) Odpowiedz na pytanie, czy średnia waga bulwy ziemniaka jest taka sama w każdych warunkach wzrostu (method)? Odpowiedź sformułuj na podstawie odpowiedniego testu. Pamiętaj o sprawdzeniu założeń. Nie wykonuj testów post-hoc. (6 p.)

(B) Sformułuj wnioski na podstawie kontrastów zbudowanych za pomocą funkcji *contr.sum*. (6 p.)

Rozwiązania wszystkich zadań zamieść w jednym pliku Rmd. Plikowi nadaj nazwę, którą jest twoje **nazwisko bez znaków diakrytycznych**. Tak nazwany plik prześlij jako odpowiedź do zadania na Teamsach. Pamiętaj, aby do każdego podpunktu podać odpowiedź.