1 Laboratorium 4. Regresja liniowa.

Wczytaj do pakietu dane ze zbioru 'crime.txt'.

1. Dla każdej z $\binom{14}{2}$ par utworzonych ze zmiennych $Rate, \ldots, X$ wykonaj wykres rozrzutu (to polecenie można wykonać za pomocą jednej komendy).

2. Wyznacz macierz współczynników korelacji próbkowych Pearsona dla zmiennych $Rate, \ldots, X$ (to polecenie również można wykonać za pomocą jednej komendy).

3. Na podstawie tych dwóch analiz odpowiedz na następujące pytania:

   (a) które ze zmiennych objaśniających $Age, \ldots, X$ wydają się mieć najmocniejszy liniowy wpływ na zmienną objaśnianą *Rate*?

   (b) czy pojawia się problem współliniowości, tzn. czy istnieje choć jedna para silnie ze sobą skorelowanych zmiennych objaśniających?

4. Skonstruuj model regresji liniowej opisujący zależność między zmienną *Rate* a trzynastoma zmiennymi objaśniającymi $Age, \ldots, X$. Wyznacz wartości współczynników determinacji R^2 i $adjR^2$.

   Podaj **równanie regresji** opisujące zależność zmiennej *Rate* od zmiennych objaśniających występujących w tym modelu.

   Oblicz prognozowaną przez ten model wartość współczynnika *Rate*, gdy zmienne objaśniające $Age, \ldots, X$ mają wartość

   $$150, 1, 90, 50, 60, 500, 950, 30, 300, 100, 40, 400, 200.$$

5. Powtórz analizę z poprzedniego punktu, przyjmując za zmienne objaśniające:

   (a) *Ex1,X,Ed, Age,U2*,

   (b) *EX0,LF,M,N,NW*.

6. Spośród tych trzech modeli regresji liniowej wybierz „najlepszy", za kryterium wyboru przyjmując wartość **skorygowanego współczynnika determinacji**.

   Podaj **równanie regresji** opisujące zależność zmiennej *Rate* od zmiennych objaśniających występujących w tym modelu.

7. Oblicz prognozowaną przez „najlepszy" model wartość współczynnika *Rate* dla podanych powyżej wartości zmiennych objaśniających. Porównaj obie prognozy.

8. Za pomocą wykresu kwantylowego dla rezyduów zbadaj (nieformalnie), czy spełnione jest założenie o normalności wektora błędów: $\boldsymbol{\varepsilon} \stackrel{D}{=} N(\mathbf{0}, \sigma^2 \boldsymbol{I}_n)$.

Maciej Wilczyński