

SAR 2020/2021 Laboratorium 6

10.11.2020

6.1 (Metoda Boxa-Coxa)

Zbiór danych *cellular.txt* zawiera informacje dotyczące liczby czytelników pewnego czasopisma (zmienna *Subscribers*).

- Wykonaj wykres pokazujący jak zmienia się liczba czytelników (zmienna *Subscribers*) w funkcji czasu (zmienna *Period*).
- Zaproponować przekształcenie zmiennej *Subscribers* które daje lepszy model.
- Zastosować metodę Boxa-Coxa do znalezienia optymalnego przekształcenia zmiennej celu.

6.2 (częściowe wykresy regresji/rezyduów, współliniowość)

Dane w pliku *savings.txt* dotyczą sytuacji ekonomicznej mieszkańców 50 krajów. Poszczególne kolumny zawierają wartości średnie z lat 1960-1970:

Country - nazwa kraju,

Savings - łączne oszczędności przypadające na osobę podzielone przez dochód netto,

dpi - dochód netto przypadający na jednego mieszkańca, *ddpi* - tempo wzrostu dochodu (w %),

Pop15, *Pop75* - procent obywateli w wieku, odpowiednio, mniejszym niż 15 lat i powyżej 75 lat.

- Dopasować model liniowy opisujący zależność *Savings* od *dpi*, *ddpi*, *Pop15* i *Pop75*. Zidentyfikować obserwacje o dużych studentyzowanych rezyduach modyfikowanych i dużych wpływach (h_{ii}). Sporządzić diagram Cooke'a. Usunąć z modelu obserwację o największej odległości Cooke'a. Sprawdzić, że jest ona wpływowa.
- Dlaczego *dpi* nie jest istotne w modelu dopasowanym w punkcie (a)? Sporządzić częściowe wykresy regresji i częściowe wykresy rezyduów dla zmiennych *dpi* i *ddpi* i zinterpretować je.
- Obliczyć współczynnik korelacji dla zmiennych *Pop15* i *Pop75*. Porównać go ze współczynnikiem korelacji dla estymatorów β_{Pop15} i β_{Pop75} . Jak możemy wytłumaczyć zaobserwowane zjawisko? Policzyc współczynnik podbicia wariancji VIF dla wszystkich zmiennych w tym modelu.
- Sporządzić wykres częściowych rezyduów dla zmiennej *Pop15*. Jaką dodatkową relację w danych możemy zauważyć na tym wykresie?