Tutorial 10

YANG YANG

The Australian National University

Week 11, 2017

Overview

Variance inflation factor (vif)

- vif() command need to use "faraway" package
- You can use the function given in Q1 solution
- Normally we care about independent variable with very large VIF.
- No exact cut-off value.

Influence statistics

- Present the relevant statistics. (Always interpret results you choose to include in the report. Statistics/plots without interpretation are useless to your clients.)
- Interpret influence statistics in relative terms.
- We are building an explanatory model to investigate the relationships between the variables (rather than a more precise predictive model).

Nested F test

- $H_0 : \frac{\sigma^2_{x.difference}}{\sigma^2_{error}} = 1$ vs $H_A : \frac{\sigma^2_{x.difference}}{\sigma^2_{error}} > 1$
- $x.difference$ in the above line refers to the difference between the full model and the nested model
- $TS = \frac{MSR_{x.difference}}{MSE}$
- Compare to $F_{n,error.df}$. Here n equals to the number of variables in $x.difference$

Deletion of outliers

- Delete outliers one at a time. After deleting an observation, re-fit the model and check if there is an improvement.
- Do not delete too many observations.
- Desired results: increase in R^2 and/or decrease in MSE.