Iniziato	mercoledì, 5 luglio 2023, 17:15
Stato	Completato
Terminato	mercoledì, 12 luglio 2023, 12:57
Tempo impiegato	6 giorni 19 ore
Valutazione	Non ancora valutato

Domanda 1

Risposta
corretta

Punteggio
ottenuto 1,00
su 1,00

Suppose that one is interested in fitting a flexible Gaussian regression model to describe the effect of a quantitative regressor on a dependent variable with P-splines, using linear splines, $K = 30$ knots and a penalty based on squared first-order differences. Which of the following R intructions should be used?

;
   a.   `mod1<-gam(y~s(x,bs="ps",k=32,m=c(0,1)),data=dataset)` ☐

'
   b.   `mod1<-gam(y~s(x,bs="ps",k=30,m=c(0,1)),data=dataset)`

'
   c.   `mod1<-gam(y~s(x,bs="ps",k=32,m=c(1,1)),data=dataset)`

'
   d.   `mod1<-gam(y~s(x,bs="ps",k=30,m=c(1,1)),data=dataset)`

'
   e.   none of them

Risposta corretta.

La risposta corretta è:

`mod1<-gam(y~s(x,bs="ps",k=32,m=c(0,1)),data=dataset)`

Domanda 2

Completo

Punteggio
max.: 2,00

The effects of a set of 7 candidate quantitative regressors on a dependent variable have been investigated by fitting 5 Gaussian linear regression models (all including the intercept) on the same sample of n units. The following table summarizes the findings, reporting which regressors have been included in each model, along with the values of some relevant summary quantities:

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	$\sum_{i=1}^{n} e_i^2$	R_a^2
M1	X	X	X					599.821	0.505
M2				X	X	X		489.345	0.505
M3		X	X	X	X			577.222	0.519
M4			X	X	X			569.689	0.278
M5		X				X		621.004	0.493

Assume that the values of R_a^2 are correct. Explain why the reported values for $\sum_{i=1}^{n} e_i^2$ contain some mistakes.

If two models have the same R^2_a and the same number of parameters, also their sum of squared residuals should be the same, which is here not the case for M1 and M2.

Since the number of sample units and the total sum of squares are unknown, in principle it is impossible to check if the residual sums of squares are correct, starting from the values of the adjusted R^2.

However, since M4 is nested in M3, the residual sum of squares for M4 cannot be smaller than the residual sum of squares for M3. Thus, one of these two residual sums of squares must be wrong.

In addition, since M1 and M2 have the same number of regressors and the same value for the adjusted R^2, their residuals sums of squares cannot be different. So, also one of these two residuals sums of squares must be wrong.
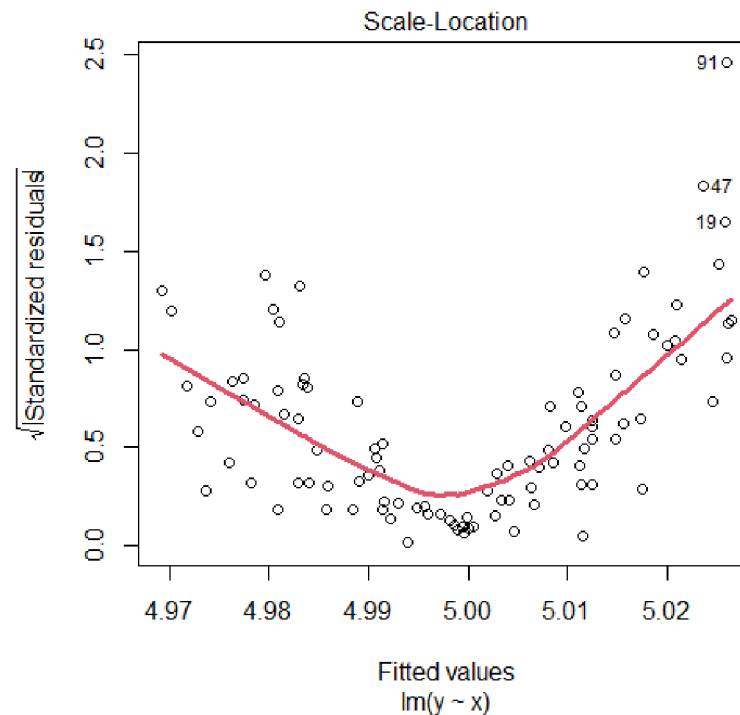
file:///G:/My Drive/Master/Unibo-StatStudents/Statistical-Models-and-Applications/Old exams/Exam question examples 3.html

4/6

Domanda 3

Completo

Punteggio
max.: 2,00

Suppose that the following plot has been obtained, after fitting a Gaussian linear model to study the effect of the regressor x on the dependent variable y.

Explain why the Box-Cox transformation would not be effective in accommodating the violation of the homoscedastisity assumption for this specific situation.

Because it tends to reduce the variance with increasing values for y. In this case, the variance of y is higher for values much higher and much lower than ca. 5. Thus it would not fix the violation of homoscedasticity assumption, because it does not reduce the variance lower than 5 and higher than 5 at the same time. This would only work if the variance would become higher with higher y (higher fitted values), but not for higher variance with lower y.

The Box-Cox transformation can be effective in accomodating the violation of the homoscedasticity assumption if the conditional variance is a proportional to a power transformation of the conditional expected value.

The scale-location plot highlights a non-monotonic pattern in the magnitude of the residuals, suggesting that the conditional variance is a non-monotonic function of the conditional expected value. Thus, in this specific situation the conditional variance cannot be proportional to a power transformation of the conditional expected value.

Precedente   Vai a...   Successivo