Iniziato	mercoledì, 5 luglio 2023, 17:11
Stato	Completato
Terminato	martedì, 11 luglio 2023, 13:47
Tempo impiegato	5 giorni 20 ore
Valutazione	Non ancora valutato

undefinedG

Domanda 1

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Consider the following Gaussian polynomial regression model:

M1: $Y_i = \sum_{p=0}^{5} \beta_p x_i^p + \varepsilon_i, \quad \varepsilon_i | x_i \sim N(0, \sigma^2) \;\; IID$

Which of the following hypotheses should be tested in order to compare M1 with a Gaussian cubic regression model?

a. $H_0 : \beta_4 = \beta_5$

b. $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$

c. $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3$

d. $H_0 : \beta_4 = \beta_5 = 0$

Risposta corretta.

La risposta corretta è:
$H_0 : \beta_4 = \beta_5 = 0$

Domanda 2

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Suppose that the following Gaussian regression models have been fitted in R using the following instructions:

```
M1: lm(y~poly(x,degree=6), data=data1)
M2: lm(y~bs(x,knots=quantile(data1$x,probs=(1:3)/4),
degree=3),data=data1)
M3: lm(y~bs(x,knots=quantile(data1$x,probs=(1:4)/5),
degree=1),data=data1)
M4: lm(y~bs(x,knots=quantile(data1$x,probs=(1:2)/3),
degree=2),data=data1)
```

Which model has the largest number of parameters?

Note: more than one option can be selected.

Scegli una o più alternative:

- [x] a. M1
- [x] b. M2
- [ ] c. M3
- [ ] d. M4

Risposta corretta.

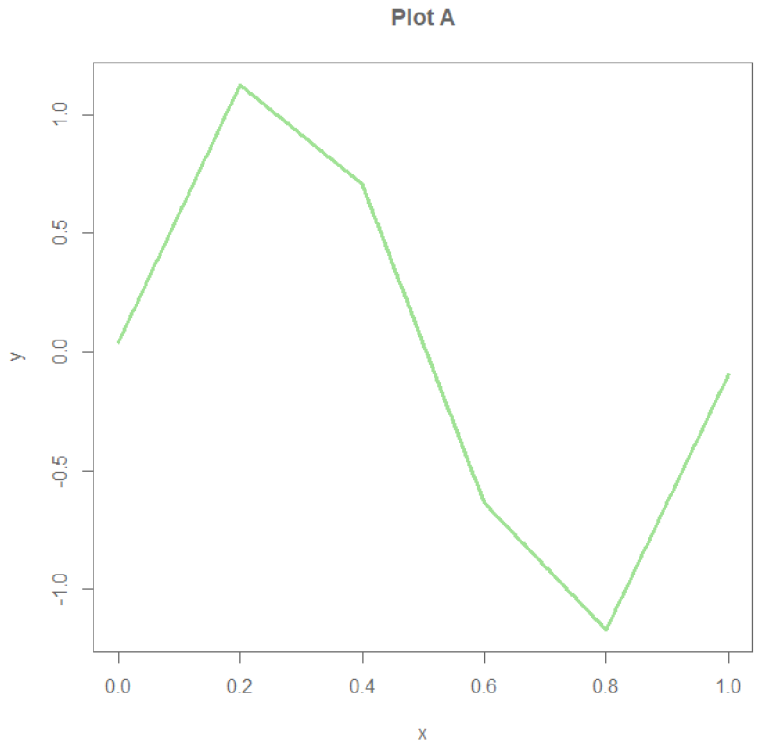
Le risposte corrette sono: M1, M2

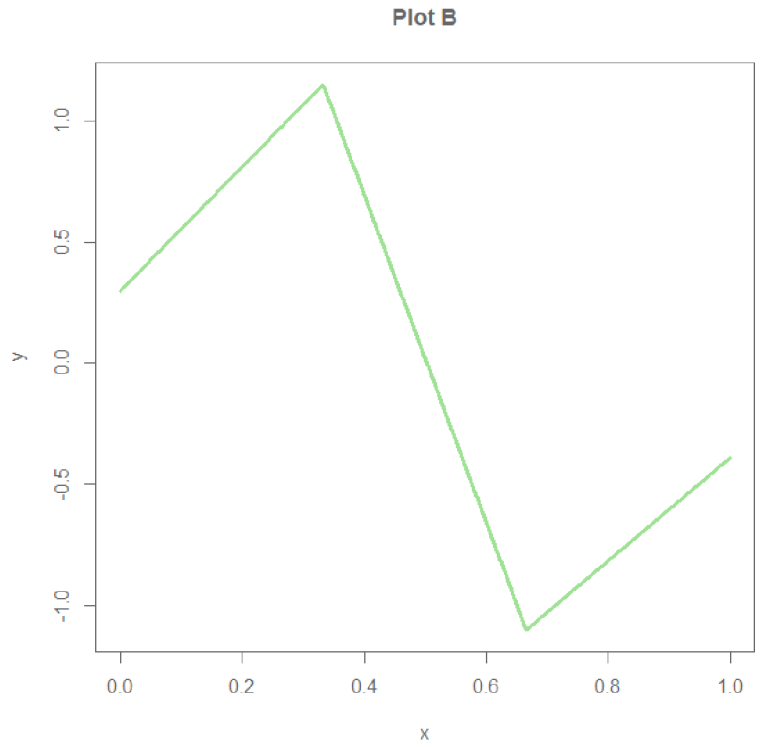
Domanda 3

Completo

Punteggio max.: 2,00

Suppose that one is interested in studying the relationship between a dependent variable and a regressor using linear spline functions. Consider the following plots, showing two estimated regression functions:


Plot A
y
1.0
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
x



Which of these two estimated regression functions cannot be obtained using the following set of bases?

$$\begin{cases} b_1(x) = 1 \\ b_2(x) = x \\ b_3(x) = (x - 0.333)_+ \\ b_4(x) = (x - 0.666)_+ \end{cases}$$

Motivate your answer.

Only Model B could be obtained by that model with the specified bases, since this is the truncated power basis for a piecewise linear model, which should be linear except for the knots that are here at 0.333 and 0.666.

The regression function depicted in Plot A cannot be obtained using the given set of bases. Namely, the set of bases is associated with linear spline functions with two knots (located at 0.333 and 0.666), while Plot A depicts a linear spline function with four knots.

Domanda 4

Completo

Punteggio max.: 2,00

Suppose that one is studying the relationship between a dependent variable and a regressor using smoothing splines. In particular, two regression functions have been estimated on the same sample of units. The first one is characterised by $\mathrm{edf}_\lambda = 12.511$, the second one by $\mathrm{edf}_\lambda = 5.768$. Which of the two estimated regression function has been obtained using the largest smoothing parameter? Motivate your answer.

The second one had the larger smoothing parameter since the larger the smooting parameter, the less flexible the model, and the less flexible the model, the less effective degrees of freedom (edf) it has.

The estimated regression function obtained with the largest smoothing parameter is the second one (the one with the smallest effective degrees of freedom). The effective degrees of freedom are obtained by computing the trace of the smoothing matrix, and this trace is a monotonic decreasing function of the smoothing parameter. In particular, with in the smoothing spline approach, when the smoothing parameter approaches zero, the trace of the smoothing matrix tends to n (the number of unique values of the regressor). On the other hand, when the smoothing parameter approaches $+\infty$, the trace of the smoothing matrix tends to 2.

Precedente

Vai a...

Successivo