

## Quiz 7

Started: Nov 12 at 2:07pm

### Quiz Instructions

This quiz is open book, open notes, "open R". The expected duration is 60 minutes. Two attempts are allowed. If both attempts are taken, the score for the second attempt will "overwrite" that from the first attempt (regardless if it is higher or lower). Even though the quiz has 24 points, it will be graded out of 20 points (i.e., 4 points bonus).

You are allowed to use any of the class materials from our SML class, but no other materials (no internet browsing or communication with other parties online/offline).

Even if a question is asking for a numerical value or True/False answer, in order to receive full credit (if your "guess" is correct) or partial credit (if appropriate, if your "guess" is incorrect), please provide your rationale as comments in the uploaded file requested at the end of the quiz.

If you solved some of the questions analytically, provide a clear scan of your work (possibly, with a smartphone app, not a raw photo), in pdf format. In this case, create a folder, place all files in it, make an archive (zip or rar) and submit.

### Questions

- Question 1
- Question 2
- Question 3
- Question 4
- Question 5
- Question 6
- Question 7

Time Running: 

Hide Time

  
Attempt due: Nov 15 at 10am  
59 Minutes, 39 Seconds

Question 14 pts

Briefly discuss merits and drawbacks of (global) polynomial regression relative to (unpenalized) regression splines.

EditViewInsertFormatToolsTable

12ptParagraphBBIUAA<sup>2</sup><sub>2</sub>∑

p0 words</>

Question 24 pts

Briefly discuss how the model complexity is determined (i.e., measured/quantified in the model and chosen by a user when trying to optimize out-of-sample prediction error) using unpenalized (regression) splines.

p0 words</>

Question 34 pts

Briefly discuss how one can penalize nonlinearity of the fitted surface (as a measure of complexity) when using penalized splines, and why this is sensible.

p0 words</>

Question 44 pts

Briefly discuss how the model complexity is determined (i.e., measured/quantified in the model and chosen by a user when trying to optimize out-of-sample prediction error) using penalized splines.

p0 words</>

Question 54 pts

Briefly discuss the idea behind penalized splines (particularly, why the penalty is necessary) and their merits and drawbacks (relative to the unpenalized splines).

p



0 words

</>



#### Question 6

4 pts

In a certain (imaginary) population, scientists observed that the desired amount of sleep per night increases from the age of 20 to 40, then decreases from the age of 40 to 60, then increases again.

Without using specialized functions in R (such as `bs` or `ns`), setup a design matrix  $X$  for the linear spline regression for this problem (so that  $X$  may be passed to the `lm` function to fit the model). Include all your code here. Output the design matrix for the following vector of ages (our  $x$ ):

```
x = c(73.8, 35.9, 42.3, 54.4, 74.5, 32.1)
```

Edit View Insert Format Tools Table

12pt Paragraph **B** *I* U

p



0 words

</>



#### Question 7

0 pts

Upload your supplemental files (if any) here.

Upload

Choose a File

Quiz saved at 2:10pm

Submit Quiz