

1. Liittyy esimerkkiin 5.1. Rakenna malli, jossa on mukana sekä pitkät että lyhyet versot. Perustelee valintasi.
2. Jatkoa ed. kerran tehtäviin 1 ja 2. a) Sovita malleja

$$\text{earn}^\lambda = \beta_0 + \beta_1 \text{height} + \beta_2 \text{age} + \beta_3 \text{sex} + \varepsilon.$$

arvoilla $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ ($\lambda = 0$ tarkoittaa logaritmia) ja valitse se λ :n arvo, joka tuottaa mielestäsi parhaan qqnorm-plotin. b) Valitsemallasi λ :n arvolla plottaa jäännökset sovituksen ja selittäjien `height` ja `age` suhteen. Näkyykö epälineaarisuutta?

3. Jatkoa edelliseen. a) Lisää malliin muuttujien `height` ja `age` neliöt. Onko syytä pitää molemmat tai jompikumpi neliö mallissa? b) Lisää ed. kohdassa valitsemaasi malliin sukupuoleen liittyvät interaktio termit. c) Missä iässä ansiot ovat suurimmillaan miehillä ja naisilla?
4. Aineistossa `salary` on erään amerikkalaisen collegen vakinaisen opetushenkilökunnan palkka- ym. tietoja 1980-luvun alkupuolelta. Rakenna sopiva malli, jonka avulla on tarkoituksena ennustaa palkkaa taustamuuttujien sopivalla kombinaatiolla. Muuttujat ovat

`Sex` = Sukupuoli, 1 = nainen, 0 = mies,

`Rank` = Virka-asema alimmasta ylimpään, 1 = assistant professor, 2 = associate professor, 3 = full professor,

`Year` = Virkavuodet nykyisessä asemassa,

`Degree` = Korkein oppiarvo, 1 = tohtori, 0 = maisteri,

`YSdeg` = Korkeimman oppiarvon saavuttamisesta kuluneet vuodet,

`Salary` = Virasta saatu vuosipalkka dollareissa.

5. Aineisto `spirit`. Tämäkin on alkoholin kulutusaineisto 1870-1938 (UK). Luvut ovat 10-kantaisia logaritmeja. Muuta ne luonnollisiksi logaritmeiksi ja tee myös muuttuja `time`, joka saa arvot 1, 2, ..., 69. Toista esimerkin 5.2 ajot ja tee tulkinnot. Totea graafisesti, että tässä riittää AR(1)-malli.