STAT 260 - Nonparametric Statistics
Ying Sun
Statistics (STAT) Program
Computer, Electrical and Mathematical Sciences & Engineering (CEMSE) Division
King Abdullah University of Science and Technology (KAUST)

HOMEWORK I

Henrique Aparecido Laureano

Spring Semester 2018

Contents

Problem 1

Question Q1.1-1.3 on page 15 of Topic 2.

Q1.1

$$\mathbb{V}\hat{\beta}_0 = \sigma^2 \frac{\sum_{i=1}^n x_i^2}{nS_{xx}} \quad ?$$

Solution:

$$\mathbb{V}\hat{\beta}_0 = \mathbb{V}(\bar{y} - \hat{\beta}_1 \bar{x}) = \mathbb{V}\bar{y} + \bar{x}^2 \mathbb{V}\hat{\beta}_1 - 2\bar{x}\mathrm{Cov}(\bar{y}, \hat{\beta}_1).$$

The variance terms are

$$\mathbb{V}\bar{y} = \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}y_i = \frac{\sigma^2}{n}, \qquad \mathbb{V}\hat{\beta}_1 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sigma^2}{S_{xx}}.$$

The covariance term is

$$\mathrm{Cov}(\bar{y}, \hat{\beta}_1) = \mathbb{E}\bar{y}\hat{\beta}_1 - \mathbb{E}\bar{y}\mathbb{E}\hat{\beta}_1 = \beta_1(\beta_0 + \beta_1\bar{x}) - (\beta_0 + \beta_1\bar{x})\beta_1 = 0.$$

So

$$\begin{aligned}
\mathbb{V}\hat{\beta}_0 = \mathbb{V}(\bar{y} - \hat{\beta}_1 \bar{x}) &= \mathbb{V}\bar{y} + \bar{x}^2 \mathbb{V}\hat{\beta}_1 - 2\bar{x}\mathrm{Cov}(\bar{y}, \hat{\beta}_1) \\
&= \frac{\sigma^2}{n} + \frac{\bar{x}^2\sigma^2}{S_{xx}} - 0 = \sigma^2 \frac{S_{xx} + n\bar{x}^2}{nS_{xx}} = \sigma^2 \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n\bar{x}^2}{nS_{xx}} \\
&= \sigma^2 \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2 + n\bar{x}^2}{nS_{xx}} = \sigma^2 \frac{\sum_{i=1}^n x_i^2}{nS_{xx}}.
\end{aligned}$$

Therefore

$$\boxed{\mathbb{V}\hat{\beta}_0 = \sigma^2 \frac{\sum_{i=1}^n x_i^2}{nS_{xx}}.}$$

$\square$

Q1.2

$$\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\sigma^2 \frac{\bar{x}}{S_{xx}} \quad ?$$

Solution:

$$
\begin{aligned}
\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \mathbb{E}(\hat{\beta}_0 - \mathbb{E}\hat{\beta}_0)(\hat{\beta}_1 - \mathbb{E}\hat{\beta}_1) &= \mathbb{E}(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1) \\
&= \mathbb{E}(\bar{y} - \hat{\beta}_1\bar{x} - \beta_0)(\hat{\beta}_1 - \beta_1) \\
&= \mathbb{E}(\beta_0 + \beta_1\bar{x} - \hat{\beta}_1\bar{x} - \beta_0)(\hat{\beta}_1 - \beta_1) \\
&= \mathbb{E}(-(\hat{\beta}_1 - \beta_1)\bar{x})(\hat{\beta}_1 - \beta_1) \\
&= -\bar{x}\mathbb{E}(\hat{\beta}_1 - \beta_1)^2 \\
&= -\bar{x}\mathbb{E}(\hat{\beta}_1^2 - 2\beta_1\hat{\beta}_1 + \beta_1^2) \\
&= -\bar{x}[\mathbb{E}\hat{\beta}_1^2 - 2\beta_1\mathbb{E}\hat{\beta}_1 + \beta_1^2] \\
&= -\bar{x}[\mathbb{V}\hat{\beta}_1 + \mathbb{E}^2\hat{\beta}_1 - \beta_1^2] \\
&= -\bar{x}\left[\frac{\sigma^2}{S_{xx}} + \beta_1^2 - \beta_1^2\right].
\end{aligned}
$$

Therefore,

$$
\boxed{\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\sigma^2\bar{x}}{S_{xx}}.}
$$

□

Q1.3

- **What does mean $\mathbb{V}\hat{\beta}_1 \neq 0$?**
- **What is the source of randomness in $\hat{\beta}_1$?**
- **How can you reduce $\mathbb{V}\hat{\beta}_1$?**

Solution:

□

Problem 2

Question Q2.1-2.2 on pages 16-17 of Topic 2.

Solution:

□

Problem 3

Question Q3 on page 19 of Topic 2.

Solution:

□

Problem 4

Question Q4 on page 20 of Topic 2.

Solution:

□

Problem 5

Write your own version of `anova()` function on page 14 of Topic 3 using `R`. Your function should not use `lm()` function. Must be able to compare two nested models. Replicate results on slides 14, 17, 18, 21 and 22 about the `gala` data set.

Solution:

□

Problem 6

Write an `R` function that can be used to automatically determine the order d of the polynomial regression and replicate the results on pages 31 and 32 of Topic 3 about the `savings` data. (The output should be the chosen order d and the plot of the fitted polynomial function to the data.)

Solution:

■