Applied Statistical Methods – Exercise 2

Peter von Rohr

2019-03-04

Problem 1: Regression Model

During the lecture the regression model was explained using the dataset given in Table 1.

Table 1: Dataset for Regression of Body Weight on Breast Circumference for ten Animals

Animal	Breast Circumference	Body Weight
1	176	471
2	177	463
3	178	481
4	179	470
5	179	496
6	180	491
7	181	518
8	182	511
9	183	510
10	184	541

The same dataset is also available from the website at https://charlotte-ngs.github.io/GELASMSS2019/ex/w03/bw_bc_reg.csv.

Your Task

- Setup the linear regression model with an intercept for the data given in Table 1
- Compute the solution for the unknown parameter b
- Verify the result with the output from the function `lm()` in R

Problem 2: Prediction

Given the measurement of the trait `Breast Circumference` for two additional animals. The measurements are shown in the following table

Table 2: Breast Circumference Measurements For Two Animals Used To Predict Body Weight

Animal	Breast Circumference
Animal 11	181.2
Calf 12	99.5

We want to use the results of Problem 1 to computed the predicted values for `Body Weight` for the two animals. The observed value for `Breast Circumference` of "Calf 12" is outside of the range of the values used in Problem 1. Predicting values of response variables based on predictors that are outside of the range

of values used for the parameter estimation is called **extrapolation**. Based on the result of the predicted value of the trait `Body Weight` for "Calf 12" what can be said about the process of extrapolation?

Your Tasks

- Compute the predicted value of `Body Weight` for "Animal 11" using the results from Problem 1
- Compute the predicted value of `Body Weight` for "Calf 12" using the results from Problem 1
- Make a statement about the validity of the extrapolated value of `Body Weight` for "Calf 12"