STAT302 Lec5101, Methods of Data Analysis 1
Module 1: Problem set

September 3, 2024

1 Basic learning objective practice

Problem 1.

A researcher randomly sampled 20 men from a population and measured their Age (X, in years) and their Percent Body Fat (Y). They found the following numerical summaries for their data:

$$\sum x_i = 628.9866 \qquad \sum x_i^2 = 21562.01 \qquad \sum y_i = 836.8006 \qquad \sum x_i y_i = 28075$$

- Suppose we seek to use a simple linear regression model for this problem. State each part of the simple linear regression model.
- Which components of the linear regression model are random? Which components are fixed? Which components are unobserved from the data?
- Using this information, calculate the slope of a simple regression relationship between Percent Body Fat and Age.
- Using this information, calculate the intercept of a simple regression relationship between Percent Body Fat and Age.
- Interpret your result in the context of the problem.
- Compute the residual for $y_{20} = 19, x_{20} = 24$.

1.1 Problem 2.

Sheather: Chapter 2, 2.8 Exercise 4 (a)

1.2 Problem 3.

Rencher: Chapter 6, Problem 6.14 (a) (dataset can be accessed in RStudio by running data(faithful), which loads the dataset named "faithful"). Use R to get the summary values needed to compute the estimates by hand and then check you get the same thing with the "lm" function.

2 Advanced learning objective practice

2.1 Problem 4.

Weisberg: Chapter 2, Problem 2.1.2

2.2 Problem 5.

Weisberg: Chapter 2, Problem 2.3