MATH205, Fall 2021, Homework 1, Due Wednesday, Sep 22, 3:30 pm

Instructions

You can submit the homework either in paper or online

- Online: Take pictures of the written (theory) part; send them (along with the simulation part) to me on Slack or through Canvas before the lecture on Wednesday
- Paper: Print out the result of the simulation part and staple it with the written work; hand it in at the beginning of the lecture on Wednesday

1 Theory

1. Problem 1: Assume we have a dataset $\{x\}$ of N data items $x_1, x_2, \ldots, x_N$. Recall the definition of mean and standard deviation

$$mean(\{x\}) = \frac{1}{N}\sum_{i=1}^{N} x_i$$

and

$$std(\{x\}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - mean(\{x\}))^2}$$

For any number c,

   - Show that $mean(\{x + c\}) = mean(\{x\}) + c$ by substituting into the definition
   - Show that $std(\{x+c\}) = std(\{x\})$ by substituting into the definition

2. Problem 2: In a population, the correlation coefficient between weight and adiposity is 0.9. The mean weight is 150 lb. The standard deviation in weight is 30 lb. Adiposity is measured on a scale such that the mean is 0.8, and the standard deviation is 0.1.

   (a) Using this information, predict the expected adiposity of a subject whose weight is 170 lb

   (b) Using this information, predict the expected weight of a subject whose adiposity is 0.75

2 Simulations

From subsection "Problems" of Section 2, *simpleR - Using R for Introductory Statistics*:

- Problem 2.2
- Problem 2.3
- Problem 2.6

For each problem, attach the script containing the commands, and a screenshot of the result of the script.