

ST227 - Introductory exercise

Submission format

You must submit a pdf that contains the output from your code, along with your explanations and other responses. Additionally, you must submit a file that you used to generate these responses. If you use RMarkdown, submit your rendered pdf and the corresponding `.rmd` file. If you generated your response file pdf in another way, you must also submit a `.R` script that contains all the code that you used.

Your work must be replicable.

Exercise 1

Suppose we have two factories, say *Factory A* and *Factory B*, that use certain machines to produce goods. The probability of failure of such a machine in any given week is 0.1 in Factory A and 0.2 in Factory B.

1.a

In R, write a procedure that simulates the time to failure for a machine in Factory A

1.b

Estimate the expected time to failure by repeating your procedure from 1.a 30 times and taking the average of your results.

1.c

The expected lifespan in this case is inversely proportional to the probability of failure. Provide an argument why this is the case and compare your simulation result from 1.b with the expected lifespan

1.d

Repeat steps 1.a and 1.b for Factory B. Using a boxplot, or any other appropriate visualisation method of your choosing, visually compare the distribution of the simulated times to failure of Factories A and B.

1.e (OPTIONAL)

Use a Gamma GLM to test for difference between the time to failure data from Factory A and B. What are your conclusions?

Exercise 2

Load the `ST227_example_data.xlsx` using the library `readxl`. Run a linear regression using `avocado_toast_searches` as your response and `year` as covariate. Interpret the results.