

Homework 1

Math 445, Spring 2017

Due Wednesday, April 5 by 4:30 p.m.

Instructions

This homework is intended to provide you space to practice EDA in R (problems 1 and 2) and to review a few concepts about working with probability density functions from probability (problems 3-5). If you need to review concepts from probability, Appendix A is a good place to start.

Please complete problems 1 and 2 using R markdown. Problems 3, 4, and 5 may be handwritten. Please knit your .Rmd file to either Word or PDF and print it for submission with the remainder of the problems.

Remember that the textbook data sets are contained in the `resampled` R package. Once it is installed, you can load the package to gain access to the data sets using the naming conventions presented in the textbook.

Problem 1

Note: This is a revised version of Chapter 2, Exercise 5.

Import data from the General Social Survey Case Study in Section 1.6 into R.

(a)

Create a table and a bar chart summarizing the responses to the question about the death penalty.

(b)

Create a table and a bar chart summarizing the responses to the question about gun ownership.

(c)

Create a table comparing the responses to the death penalty and gun ownership questions.

(d)

What proportion of gun owners favor the death penalty? Does it appear to be different from the proportion among those who do not own guns?

Problem 2

Note: This is a revised/expanded version of Chapter 2, Exercise 6.

Import data from the Black Spruce Case Study in Section 1.9 into R.

(a)

Compute the following numeric summaries for the height changes (Ht.change): minimum, .25 quantile (Q1), median, .75 quantile (Q3), mean, standard deviation, and the count.

(b)

Create a histogram and normal quantile plot for the height changes of the seedlings. Is the distribution approximately normal?

(c)

Compute the following numeric summaries for the height changes (**Ht.change**) by whether or not they were fertilized plots (**Fertilizer**): minimum, .25 quantile (Q1), median, .75 quantile (Q3), mean, standard deviation, and the count.

(d)

Create a plot to compare the distribution of the change in diameters of the seedlings (**Di.change**) by whether or not they were fertilized plots. What does the plot reveal?

(e)

Compute the following numeric summaries for the height changes (**Ht.change**) by whether or not they were fertilized plots (**Fertilizer**) and competition (**Competition**) status: minimum, .25 quantile (Q1), median, .75 quantile (Q3), mean, standard deviation, and the count.

(f)

Create a plot to compare the distribution of the change in diameters of the seedlings (**Di.change**) by whether or not they were fertilized plots and competition (**Competition**) status. What does the plot reveal?

(g)

Create a scatter plot of the height changes against the diameter changes and describe the relationship.

Problem 3

Chapter 2, Exercise 8

Problem 4

Chapter 2, Exercise 10

Problem 5

Chapter 2, Exercise 12