

# Homework 1

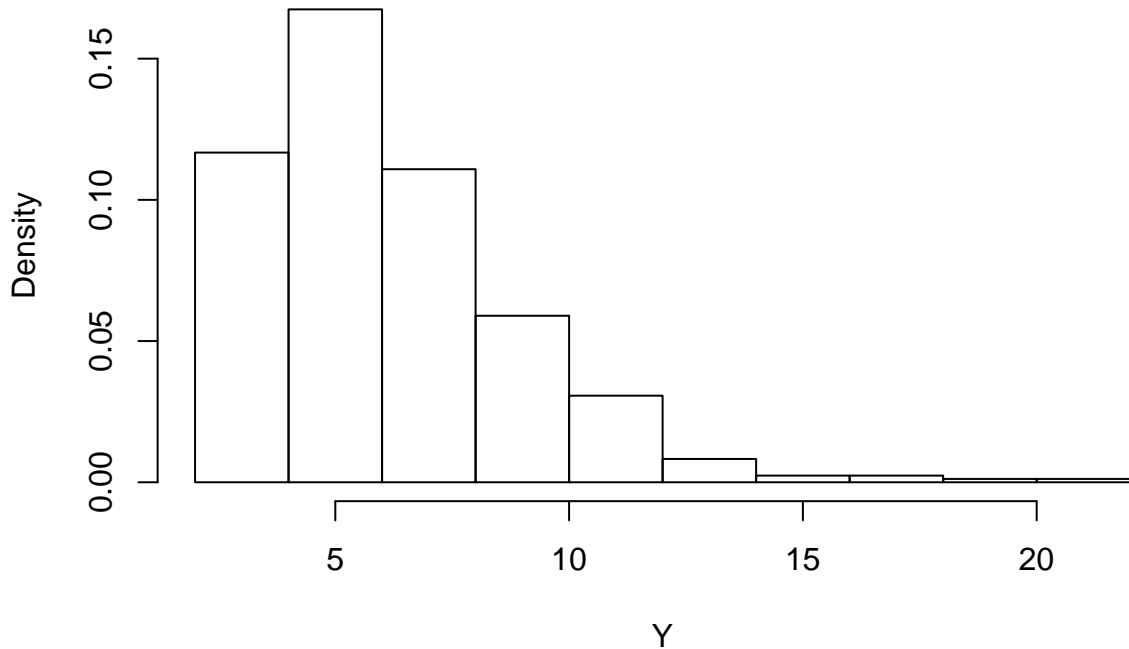
Due: Thursday 1/30/20 by 8:30am

The goal of this homework assignment is to ensure that you are comfortable with the review material, and to get you started using R. Future homework assignments will be made up of questions selected from the textbook.

Problems 1., 2., 3., and 4. will be graded. Problems 5., 6., and 7. will be done in class on Tuesday 1/28/20. Problem 8. is optional, but submissions after 8:30am on Thursday 1/30/20 will not be considered.

1. Install R and (if desired) RStudio. If you have successfully installed R by Thursday 1/30/20 by 8:30am, write “I have installed R” in response to this question to get full credit.
2.
  - (a) Obtain a 90 percent confidence interval for  $\mu$  when  $n = 424$ ,  $\bar{Y} = 6.09$ , and  $s = 2.78$ .
  - (b) Simulate 1,000 draws from the sampling distribution of  $\bar{Y}$ . Plot a histogram of  $\bar{Y}$  on the density scale. Add vertical lines at  $\bar{Y}$  and the upper and lower bounds of the interval derived in (a), and overlay the density of the sampling distribution of  $\bar{Y}$  on the histogram.
3. Choose between the alternatives  $H_0 : \mu \geq 5$  and  $H_a : \mu < 5$ , where  $\alpha$  is to be controlled at 0.03 and  $n = 84$ ,  $\bar{Y} = 5.08$ , and  $s = 9.51$ . Justify your answer with reference to the value of the test statistic, the decision rule, and the  $p$ -value.
4. Choose between the alternatives  $H_0 : \mu = -10$  and  $H_a : \mu \neq -10$  when  $\alpha$  is to be controlled at 0.15 and  $n = 13$ ,  $\bar{Y} = -6.81$ , and  $s = 1.55$ . Justify your answer with reference to the value of the test statistic, the decision rule, and the  $p$ -value.
5. (In class, not to be submitted)
  - (a) Obtain a 95 percent confidence interval for  $\mu_1 - \mu_2$  when  $n_1 = 49$ ,  $\bar{Y} = 1.1$ ,  $\sum (Y_i - \bar{Y})^2 = 96.04$ ,  $n_2 = 34$ ,  $\bar{Z} = 3.37$ , and  $\sum (Z_i - \bar{Z})^2 = 22.78$ .
  - (b) Choose between the alternatives  $H_0 : \mu_1 = \mu_2$  and  $H_a : \mu_1 \neq \mu_2$  when  $\alpha$  is to be controlled at 0.01. Justify your answer with reference to the value of the test statistic, the decision rule, and the  $p$ -value.
6. (In class, not to be submitted)
  - (a) Obtain a 85 percent confidence interval for  $\sigma^2$  using the data from 2.

## Histogram of Y



- (b) Suppose a peer who had access to the raw data showed you the above histogram of the values  $Y_1, \dots, Y_n$ . In at most one sentence, explain how this would affect your conclusions in (a).
7. (In class, not to be submitted)
- (a) Obtain a 99 percent confidence interval for  $\sigma_1^2/\sigma_2^2$  using the data from 5.
- (b) Choose between the alternatives  $H_0 : \sigma_1^2 = \sigma_2^2$  and  $H_a : \sigma_1^2 \neq \sigma_2^2$  when  $\alpha$  is to be controlled at 0.2. Justify your answer with reference to the value of the test statistic, the decision rule, and the  $p$ -value.
8. (Optional) Submit suggestions for additional project topics to add to the existing list of topics. To get full credit that counts towards your participation grade, submit:
- A very short (up to ten word) title of the dataset that describes its contents;
  - A more detailed one sentence description of the data including the number of observations, the number of covariates/predictors, and any other relevant information;
  - A one sentence description of at least one regression problem of interest that this data could be used to address.