

# STAT 341: Assignment 4 - Fall 2020

Ryan Browne

Due Friday November 13 at 10:am

## NOTES

Your assignment must be submitted by the due date listed at the top of this document, and it must be submitted electronically in .pdf format via Crowdmark/LEARN. This means that your responses for different questions should be in separate .pdf files. Your .pdf solution files must have been generated by R Markdown. Additionally:

- For mathematical questions: your solutions must be produced by LaTeX (from within R Markdown). Handwritten and scanned/photographed solutions will not be accepted and you will receive zero points.
- For computational questions: R code should always be included in your solution (via code chunks in R Markdown). If code is required and you provide none, you will receive zero points.
  - **Exception** any functions used in the notes or function glossary can be loaded using `echo=FALSE` but any other code chunks should have `echo=TRUE`. e.g. the code chunk loading `gradientDescent` can use `echo=FALSE` but chunks that call `gradientDescent` should have `echo=TRUE`.
- For interpretation question: plain text (within R Markdown) is fine.

Organization and comprehensibility is part of a full solution. Consequently, points will be deducted for solutions that are not organized and incomprehensible.

## Question 1 - 16 Marks - Estimating Brightness

For this question you will need the digit data from file “digitData.csv”. Use the sample below from the Digit data to answer the following questions. The sample was obtained by sampling without replacement.

```
digitSample <- c(294,133,95,265,154,1,289,232,121,99,129,83,30,56,249,134,46,68,165,279,105,91,248,285,1
```

- a) [2 Marks] Calculate the Horvitz-Thompson estimate of the average amount of brightness and provide the standard error of the estimate.
- b) [2 Marks] Calculate the Horvitz-Thompson estimate of the proportion of digits with brightness less than or equal to 45
- c) [2 Marks] Calculate the Horvitz-Thompson estimate of the proportion of digits with brightness in the interval [20, 25).
- d) [4 Marks] In three graphs in a row,
  - plot the Horvitz-Thompson estimate of the cumulative distribution function for brightness,

- plot the standard error of the estimate and
  - plot the Horvitz-Thompson estimate of the cdf and overlay the lines of  $\pm 2$  times the standard error.
  - **Note** Similar to the cdf, the standard error is also a function.
- e) [4 Marks] In three graphs in a row,
- plot the Horvitz-Thompson estimate of the histogram for brightness with bin width equal to five,
  - plot the standard error of the estimate and
  - plot the Horvitz-Thompson estimate of the histogram and overlay the lines of  $\pm 2$  times the standard error.
  - **Note** Similar to the cdf, the standard error is also a function.
- f) [2 Marks] Calculate the Horvitz-Thompson estimate of the proportion of ones and provide the standard error of the estimate.

---

### Question 2 - 16 Marks - Comparing Estimators

Using a population, compare two or more estimators of some population attribute. Your answer should include:

- description of the population and the context of the problem,
- the effect of changes in some parameters such as the sample size,
- numerical and graphical comparison and their descriptions,
- a conclusion.
- Your answer should be limited to 1 to 3 pages.

Your solution should be **in your own words**, but as **motivating** examples consider the following samples exercises:

- 1.6 Comparing the standard deviation and the Median Absolute Deviation,
- 1.7 Comparing two sampling designs,
- 1.10 Comparing measures of dispersion,
- 1.13 Comparing the Mean and Median and
- 1.17 Estimating the median for Radar Data.

### Rubric

Criteria	Descriptor	Marks
Population/Estimators	Description and Difficulty	/4
Format	Clarity, Organization and LaTeX	/4
Comparison	Description, Results and Graphic	/4
Discussion/Summary	Justification and Relevant Terminology used	/4

---

**Question 3 - 10 Marks - Summarize**

**In your own words** summarize the subsection **4.2.1-Anatomy\_of\_a\_Significance\_level**

- You are limited to 1 to 2 pages.
- You are recommended to use a combination of formulas, full sentences and example.

**Rubric**

Criteria	Descriptor	Marks
Format	Organization	/3
Writing	Clarity & Grammar	/2
Content	Coverage, Depth, Relevant Terminology used and Example	/5