Lab 1 – MATH 240 – Computational Statistics

Ben Horner
Colgate University
Math
bhorner@colgate.edu

Abstract

This document provides a basic template for the 2-page labs we will complete each week. Here, you should provide a succinct summary about what you did and why it might be helpful.

Keywords: What topics does the lab cover with respect to class?

1 Instructions

For this lab, you will

1. Install R and RStudio
2. Install tinytex (if necessary):
   `install.packages(‘‘tinytex")`
3. Create a GitHub account here, and email me your username
4. Install GitHub desktop.
5. Accept the LAB 1 assignment here.
6. Recreate this document (except put your name/info at the top) to get used to writing in LaTeX and to see the types of things we can do when creating a document to convey statistical information. Make sure to commit and push your work using GitHub desktop as you finish each section.

Remark: You will find the class Sweave cheatsheet to be *incredibly* (`\emph{incredibly}`) helfpul.

2 Word Processing Tasks

2.1 Centering Text

We can center text in Sweave

2.2 Bold, Italics, and Underlining

We can **bold**, *italicize*, underline, and *emphasize* text in Sweave.

Note, I did a clumn break here so that the list wasn't broken across columns.

2.3 Lists, and Numbered Lists

We can write an unordered list in Sweave.

- first item
- second item
- third item

We can write a numbered list in Sweave.

1. first item
2. second item
3. third item

We can write a lettered list in Sweave.

a. first item
b. second item
c. third item

2.4 Submissions

This part of the midterm is due Sunday November 14 by 5p. I will not accept late submissions. Note that you may use this template to help build your introduction and methods sections, and you can use the work you did as a group during the datathon. Still, I expect this submission to be your own summary and extension of that work without collaboration.

2.5 Typing Mathematical Equations

We can write a one line equation that is centered like this

$$\widehat{y_i} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_1 i x_2 i + \epsilon_i.$$

This can be written in the text, as $\widehat{y_i} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + \epsilon_i$ as well.

When we need to show multiple steps, we can create a multi-line equation that is centered like this:

$$\begin{aligned} 8(x-5)+x &= 9(x-5)+5 \\ 8x-40+x &= 9x-45+5 && \text{(Distributing)} \\ 9x-40 &= 9x-40 && \text{(Combining like terms)} \\ 9x &= 9x && \text{(Adding 40 to both sides)} \\ x &= x && \text{(Dividing both sides by 9)} \end{aligned}$$

This equality holds for any x.

Note, I did a page break here so that the next section started on a clean page.

2.6 Running R Code

Code chunks can be entered into Sweave; eg., here are some comments.

```
# R code goes here
# Output is automatically printed in the pdf
```

Below, you can see that we can do algebra with R.

```
8*(9-5) + 9  #8(x-5) + x for x=9

## [1] 41
```

Below, we show we can produced the code without evaluating it.

```
8*(9-5) + 9  #8(x-5) + x for x=9
```

Alternatively, we can produced the output without the code.

```
## [1] 41
```

We can also call object values from `R`. directly.

```
result <- 8*(9-5) + 9   #8(x-5) + x for x=9
result.with.error <- result + rnorm(1, mean = 0, sd = 0.1)
result.with.error

## [1] 40.94736
```

The result is 40.9473612. Note that I did not type the result, but I used the `\Sexpr{}` command.

3 Methods

Describe the data you are working with, if applicable. Describe the specific process you will follow to answer the question at hand. This does not mean you should write something like this.

> *I did this and then I did that and then I did this other thing and then..., and then..., and then...*

Instead, it should provide a clear and concise narrative that flows from the problem specification in the Introduction to how you will approach answering it. This is where I would expect to see some citations for `R` packages you will use to conduct the statistical analysis reported in the Results section.

3.1 Methods Subsection

Much like the Introduction, subsections can be helpful for the Methods section. For example, you might describe data collection and the statistical analyses of the collected data in different subsections. Or, you may have different questions that require distinct methods.

4 Results

Tie together the Introduction – where you introduce the problem at hand – and the methods – what you propose to do to answer the question. Present your data, the results of your analyses, and how each reported aspect contributes to answering the question. This section should include table(s), statistic(s), and graphical displays. Make sure to put the results in a sensible order and that each result contributes a logical and developed solution. It should not just be a list. Avoid being repetitive.

4.1 Results Subsection

Subsections can be helpful for the Results section, too. This can be particularly helpful if you have different questions to answer.

5 Discussion

You should objectively evaluate the evidence you found in the data. Do not embellish or wish-terpet (my made-up phase for making an interpretation you, or the researcher, wants to be true without the data *actually* supporting it). Connect your findings to the existing information you provided in the Introduction.

Finally, provide some concluding remarks that tie together the entire paper. Think of the last part of the results as abstract-like. Tell the reader what they just consumed – what's the takeaway message?

Bibliography: Note that when you add citations to your bib.bib file *and* you cite them in your document, the bibliography section will automatically populate here.

6 Appendix

If you have anything extra, you can add it here in the appendix. This can include images or tables that don't work well in the two-page setup, code snippets you might want to share, etc.