

Lab 1 – MATH 240 – Computational Statistics

First Author
Affiliation
Department
email@domain

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 L^AT_EX and to see the types of thing 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* helpful.

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 column 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 Submissions

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

$$\hat{y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + \epsilon_i.$$

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

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

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

The 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

Codechunks can be entered into Sweave; e.g., 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 produce the code without evaluating it.

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

Alternatively, we can produce 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] 41.13271
```

The result is 41.132712. Note that I did not type the result, but I used the `\Sexpr{}` command.

2.7 Plotting

We can also plot with R.

```
#Plot a histogram of random normally distributed data
hist(rexp(100))
```

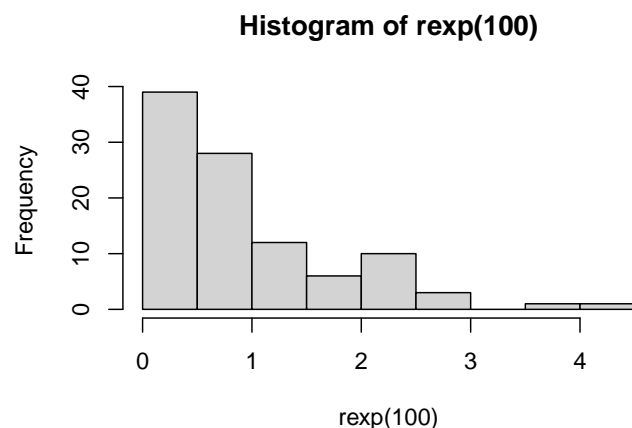


Figure 1: A histogram of random normally distributed data, $n = 100$.

2.8 Tables

Below, we load and take a peek at some data about the death rates per 1000 in Virginia in 1940 (Molyneaux et al., 1947).

```
data(VADeaths)
head(VADeaths) # Take a peek of the data
```

```
##      Rural Male Rural Female Urban Male Urban Female
## 50-54      11.7       8.7     15.4       8.4
## 55-59      18.1      11.7     24.3      13.6
## 60-64      26.9      20.3     37.0      19.3
## 65-69      41.0      30.9     54.6      35.1
## 70-74      66.0      54.3     71.1     50.0
```

If we want to print this nicely we can do so using the `xtable` package (Dahl et al., 2019), which we can reference using the label.

```
library(xtable)
sleep.table = xtable(VADeaths,
  label = "VADeaths.tab",
  caption = "Death rates per 1000 in Virginia (1940).")
```

Rural Male	Rural Female	Urban Male	Urban Female
11.70	8.70	15.40	8.40
18.10	11.70	24.30	13.60
26.90	20.30	37.00	19.30
41.00	30.90	54.60	35.10
66.00	54.30	71.10	50.00

Provide an overarching summary of what you're talking about. In this section, you introduce the idea to the reader, and your goal is to pull them in. What's the mystery you aim to solve?

You want to provide enough background to understand the context of the work. Specifically, what is the question you are addressing? If it applies, describe what information currently exists about this problem, including citations, and explain how the question you're answering complements this work.

Provide a roadmap of the structure of the paper.

2.9 Intro Subsection

You might need/want to discuss the topics in subsections. Or, you may have multiple questions.

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 phrase 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.

References

Dahl, D. B., Scott, D., Roosen, C., Magnusson, A., and Swinton, J. (2019). *xtable: Export Tables to LaTeX or HTML*. R package version 1.8-4.
Molyneux, L., Gilliam, S. K., and Florant, L. (1947). Differences in virginia death rates by color, sex, age and rural or urban residence. *American Sociological Review*, 12(5):525–535.

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.