

Lab 10 – MATH 240 – Computational Statistics

Caroline Devine
Colgate University
Mathematics Department
cdevine@colgate.edu

04/08/2025

Abstract

This document provides a basic template for the 2-page labs we will complete each week. Here, briefly summarize what you did and why it might be helpful. Provide all the top-line conclusions, but avoid providing *all* the details. Results should be limited to “we show X, Y, and Z.”

Keywords: What topics does the lab cover concerning class? List 3-4 key terms here, separated by semicolons.

1 Introduction

Gallup published a document called “*How are Polls Conducted?*”. It outlines how Gallup selects people to include in its poll as well as other details. Focusing on two excerpts describing sample sizes of 1,000 and then 2,000, Gallup highlights how the results are accurate within a certain margin of error ($\pm 4\%$ for $n = 1000$ and $\pm 2\%$ for $n = 2000$). This margin of error tells researchers how much they can expect the sample proportion to differentiate from the proportion for the entire population. The goal of statisticians and quantitative researchers is to report a margin of error that provides 95% confidence meaning they are 95% confident that if they were to conduct the poll repeatably, they would contain the actual proportion 95% of the time on the interval. The goal of this lab is to explore how sample size affects the margin of error as well as how to estimate this accurately.

2 Methods

We will explore different simulations of different sample sizes and look at the differences in margins of errors through various mediums. To do this, we will look at basic simulation, resampling, simulations over different sample sizes and proportions, and calculate the actual margin of error.

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.

2.1 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.

3 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.

3.1 Results Subsection

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

4 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.

5 Appendix